

IONIZABLE CATIONIC COMPOUND

FIELD

[0001] The present disclosure relates to cationic and/or ionizable lipid compounds
5 that can be used, in combination with other lipid molecules, to form lipid nanoparticles
for delivery of a polynucleotide to a subject.

BACKGROUND

[0002] Nucleic acid-based therapies have shown substantial promise in a range of
10 therapeutic applications. The delivery of polynucleotides such as messenger RNA
(mRNA), small interfering RNA (siRNA), antisense oligonucleotides, plasmids, DNA
and the like does, however, present a number of challenges. Free nucleic acids, such as
RNAs, are subject to rapid enzymatic degradation and so generally do not persist
systemically. Additionally, due to their negative charge the nucleic acids may not be able
15 to effectively cross the cellular barriers to enter the necessary intracellular compartment,
for example, for translation or to otherwise achieve their effect.

[0003] Lipid particles, such as lipid nanoparticles (LNPs), have therefore been used
to formulate nucleic acids so as to protect them from degradation and improve cellular
uptake and intracellular delivery. LNPs are commonly formed from ionizable cationic
20 lipids and other lipid components such as neutral lipids, sterols such as cholesterol and
PEGylated lipids. Ionizable cationic lipids are amphiphilic molecules having a lipophilic
region containing one or more hydrocarbon groups and a hydrophilic region containing
at least one positively charged or ionizable polar head group. Such cationic lipids are
ionized at an appropriate pH and can then form a positively charged complex with nucleic
25 acids, making it easier for the nucleic acids to pass through the plasma membrane of the
cell and enter the cytoplasm.

[0004] The first siRNA therapeutic to be approved, Onpattro (patisiran), entered
the market just a few years ago for treatment of hereditary amyloidogenic transthyretin
(TTR) amyloidosis. Patisiran's therapeutic effect relies on siRNA-mediated TTR gene
30 silencing, preventing mutant protein production to at least prevent disease progression.

The efficient delivery of the siRNA depends upon the LNP technology. Even more recently, nucleic acid vaccines have emerged as a promising approach to the treatment and prevention of various diseases, including against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) responsible for causing the on-going worldwide pandemic of the severely infectious coronavirus disease 2019 (COVID-19). mRNA vaccines rely on the delivery of the mRNA into the cytoplasm of host cells, where it is transcribed into antigenic proteins to trigger the production of an immune response. The large size and negative charge of mRNA prevents cellular uptake and so LNPs are again necessary for appropriate delivery.

10 [0005] Different ionizable cationic lipids may present not only different physicochemical profiles, including their acid dissociation constant (pKa) value, thereby affecting their ability to complex with the nucleic acid, but also different toxicity profiles *in vivo* and so it will be apparent to the skilled person that there is an on-going need for improved ionizable cationic lipid compounds which are suitable to form lipid particles,

15 such as lipid LNPs, for delivery of nucleic acids and polynucleotides.

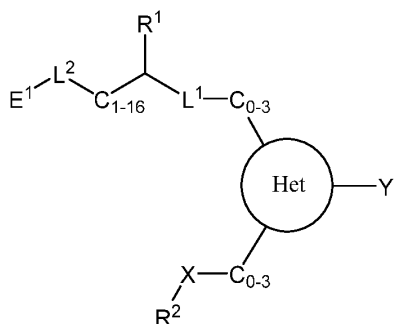
BRIEF SUMMARY

[0006] Embodiments of the present invention provide for novel ionizable cationic lipid compounds, and pharmaceutically acceptable salts, prodrugs and stereoisomers thereof, which can form lipid particles, for example LNPs, in the presence of additional lipids including one or more of neutral lipids, charged lipids, structural lipids, PEGylated lipids and analogs thereof, and which can be used for the delivery of a polynucleotide. Compositions comprising such lipid particles, methods of forming the lipid particles, their use in delivering a polynucleotide and methods of using the lipid particles in the treatment of a range of diseases, disorders and conditions are provided.

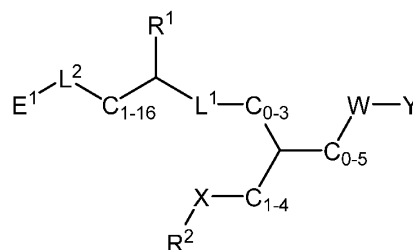
20

25

[0007] In one broad form, the present disclosure provides for a compound of Formula I-Het or a compound of Formula II:



Formula I-Het;



Formula II;

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

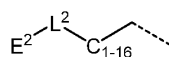
Het is a nitrogen heterocycle;

X is selected from the group consisting of -S-, -O- and -C-;

E¹ is a linear or branched -C₁₋₃₀-alkyl;

5 R¹ is selected from the group consisting of -H and Formula IA;

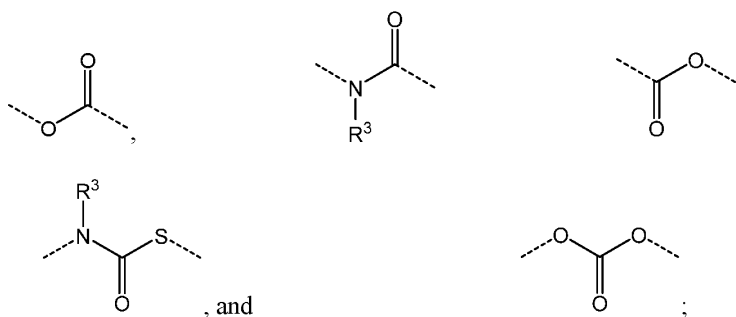
R² is selected from the group consisting of linear or branched -C₁₋₃₀-alkyl and Formula IA;



Formula IA;

10 E², if present, is a linear or branched -C₁₋₃₀-alkyl;

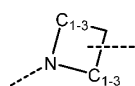
L¹ is selected from the group consisting of:



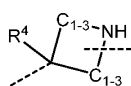
R³ is selected from the group consisting of -H and linear or branched -C₁₋₈-alkyl;

L^2 is selected from $-OC(O)-$ and $-C(O)O-$;

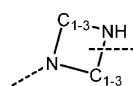
W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



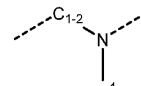
Formula IIA;



Formula IIB;



Formula IIC;



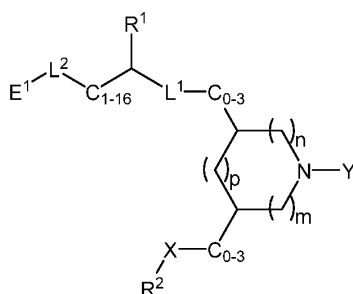
Formula IID;

R^4 , if present, is selected from the group consisting of $-H$ and linear or branched $-C_{1-5}$ -alkyl;

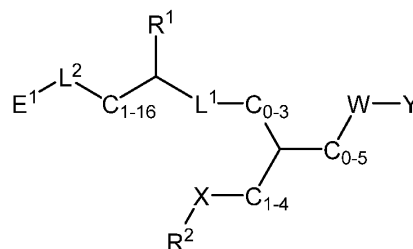
Y is selected from the group consisting of $-H$, linear or branched $-C_{1-5}$ -alkyl, and linear or branched $-C_{1-5}$ -alkanol; and

dashed lines represent a bond to an adjacent atom in the compound of Formula I-Het or the compound of Formula II.

10 [0008] Accordingly, in one embodiment, the present disclosure provides for a compound of Formula I or a compound of Formula II:



Formula I;



Formula II;

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

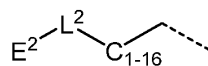
X is selected from the group consisting of $-S-$, $-O-$ and $-C-$;

E^1 is a linear or branched $-C_{1-30}$ -alkyl;

15 R^1 is selected from the group consisting of $-H$ and Formula IA;

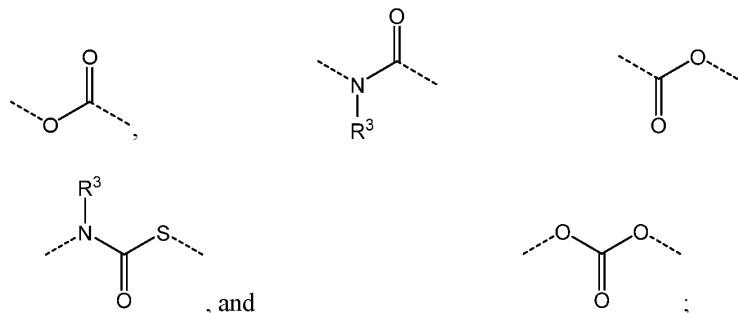
5

R^2 is selected from the group consisting of linear or branched $-C_{1-30}$ -alkyl and Formula IA;



Formula IA;

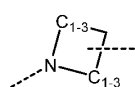
- 5 E^2 , if present, is a linear or branched $-C_{1-30}$ -alkyl;
 m and n are each independently an integer from 0 to 3;
 p is an integer from 0 to 2;
 L^1 is selected from the group consisting of:



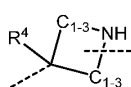
R^3 is selected from the group consisting of $-H$ and linear or branched $-C_{1-8}$ -alkyl;

- 10 L^2 is selected from $-OC(O)-$ and $-C(O)O-$;

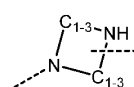
W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



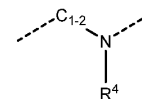
Formula IIA;



Formula IIB;



Formula IIC;



Formula IID;

R^4 , if present, is selected from the group consisting of $-H$ and linear or branched $-C_{1-5}$ -alkyl;

Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol; and

dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

5 [0009] In embodiments, lipid nanoparticles (LNPs) are provided which comprise a compound of Formula I or a compound of Formula II.

[0010] In embodiments, the LNPs further comprise a polynucleotide.

[0011] In embodiments, a pharmaceutical composition comprising such LNPs and at least one pharmaceutically acceptable carrier, diluent or excipient is provided.

10 [0012] In other embodiments, a method of forming an LNP comprising a compound of Formula I or a compound of Formula II is provided.

[0013] In additional embodiments, a method of delivering a polynucleotide, within the foregoing LNPs, to a cell is provided.

[0014] In further embodiments, a method of producing a polypeptide of interest in
15 a cell is provided.

[0015] In still further embodiments, a method of treating a disease, disorder or condition in a subject is provided by administering one or more of the foregoing LNPs comprising a polynucleotide, or a pharmaceutical composition comprising same, to a subject in need of such treatment.

20 [0016] The LNPs comprising a polynucleotide, or a pharmaceutical composition comprising same, may be delivered to the subject as a component of a vaccine.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Figure 1 shows the results of a FACS potency assay of an LNP formulation
25 of the invention relative to an LKY750 (control) ionizable lipid LNP.

[0018] Figure 2 shows the results of a Cytation 5 potency assay of an LNP formulation of the invention relative to an LKY750 (Control) ionizable lipid LNP.

[0019] Figures 3A to D show a graphical representation of *in vivo* data comparing LKY750 as a known lipid comparator and an adjuvanted inactivated virus vaccine (aH5N1) against two lipids of the invention (SL57 and SL60) when used as ionizable lipids within LNPs in the following assays: (A) hemagglutination inhibition (HAI) assay;
5 (B) pseudovirus microneutralization assay; and (C) neuraminidase inhibition enzyme-linked lectin assay (ELLA); (D) IgG enzyme linked immunosorbent assay (IgG ELISA).

DETAILED DESCRIPTION

[0020] The present disclosure is based on the use of certain novel biodegradable
10 lipid compounds as a component of an LNP for the delivery of a polynucleotide. The novel lipid compounds present biodegradable groups which may assist in reducing toxicity or improving clearance, *in vivo*.

[0021] In one embodiment, the LNP may be a component of a vaccine although the therapeutic use of the compounds described herein and the use of LNPs which they form
15 a component of, is not so limited.

[0022] In some embodiments the LNPs formed may be suitable for the delivery of messenger RNA (mRNA).

[0023] In some embodiments, the LNPs formed may be suitable for the delivery of mRNA as a component of a mRNA vaccine.

20

General

[0024] Throughout this specification, unless specifically stated otherwise or the context requires otherwise, reference to a single step, composition of matter, group of steps or group of compositions of matter shall be taken to encompass one and a plurality
25 (i.e., one or more) of those steps, compositions of matter, groups of steps or groups of compositions of matter.

[0025] Those skilled in the art will appreciate that the present disclosure is susceptible to variations and modifications other than those specifically described. It is to be understood that the disclosure includes all such variations and modifications. The

disclosure also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

[0026] The present disclosure is not to be limited in scope by the specific
5 embodiments described herein, which are intended for the purpose of exemplification only. Functionally equivalent products, compositions, and methods are clearly within the scope of the present disclosure.

[0027] Any embodiment of the present disclosure herein shall be taken to apply
10 *mutatis mutandis* to any other embodiment of the disclosure unless specifically stated otherwise.

[0028] Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (for embodiments, in organic synthetic chemistry, cell culture, molecular genetics, immunology, immunohistochemistry, protein chemistry, and
15 biochemistry).

[0029] Unless otherwise indicated, any recombinant protein, cell culture, and immunological techniques utilized in the present disclosure are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular Cloning, John
20 Wiley and Sons (1984), J. Sambrook *et al.* Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), and F.M. Ausubel *et al.* (editors), Current Protocols in Molecular
25 Biology, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present), Ed Harlow and David Lane (editors) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, (1988), and J.E. Coligan *et al.* (editors) Current Protocols in Immunology, John Wiley & Sons (including all updates until present).

[0030] The term “and/or”, e.g., “X and/or Y” shall be understood to mean either “X and Y” or “X or Y” and shall be taken to provide explicit support for both meanings or for either meaning.

[0031] The terms “from” and “to”, when indicating a range, shall be understood to mean the range is inclusive of the recited lower and upper values. For example, “n is an integer from 0 to 3” shall be understood as including the situation in which n is not present (n is 0), that in which n is 3, as well as each whole number integer value in between i.e., n is 1 or 2.

[0032] Throughout this specification the word “comprise”, or variations such as “comprises” or “comprising”, will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

[0033] As used herein the term “derived from” shall be taken to indicate that a specified integer may be obtained from a particular source albeit not necessarily directly from that source.

Selected Definitions

[0034] As used herein, the terms “lipid particle”, “lipid nanoparticle” or “LNP” shall be understood to refer to lipid-based particles having at least one dimension on the order of nanometers (e.g., 1-1,000 nm) and which comprises a compound of any formulae described herein. In embodiments, LNPs are formulated in a composition for delivery of a polynucleotide to a desired target such as a cell, tissue, organ, tumor, and the like. The LNPs generally comprise an ionizable cationic compound of the present disclosure and one or more of a neutral lipid, charged lipid, sterol and PEGylated lipid. In embodiments, the lipid particle or LNP may be selected from liposomes or vesicles, where an aqueous volume is encapsulated by amphipathic lipid bilayers (e.g., single; unilamellar or multiple; multilamellar), micelle-like lipid nanoparticles having a non-aqueous core and solid lipid nanoparticles. In embodiments, the lipid nanoparticle or LNP may have a structure that includes a single monolayer or bilayer of lipids that

encapsulates a solid phase. In preferred embodiments, the lipid nanoparticle or LNP does not have an aqueous phase or other liquid phase in its interior.

[0035] A "cationic compound", "ionizable cationic compound", "cationic lipid compound", "ionizable cationic lipid compound", or like terms, refer to a lipid compound of any structural formulae described herein and which is capable of bearing a positive charge. Ionizable cationic lipids disclosed herein include one or more nitrogen-containing groups which may bear the positive charge. They are ionizable such that they can exist in a positively charged or neutral form, depending on pH. The ionization of the cationic lipid affects the surface charge of the lipid nanoparticle under different pH conditions.

[0036] The term "neutral lipid" refers to any of a number of lipid species that exist either in an uncharged or neutral zwitterionic form at a selected pH. At physiological pH, such lipids include, but are not limited to, phosphatidylcholines such as 1,2-Distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-Dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), and phosphatidylethanolamines such as 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), sphingomyelins (SM).

[0037] The term "charged lipid" refers to any of a number of lipid species that exist in either a positively charged or negatively charged form independent of the pH within a useful physiological range e.g. pH ~3 to pH ~9. Non-limiting examples of charged lipids include phosphatidylserines, phosphatidic acids, phosphatidylglycerols, phosphatidylmositols, sterol hemisuccinates, dialkyl trimethylammonium-propanes, (including DOTAP and DOTMA), dialkyl dimethylaminopropanes, ethyl phosphocholines, and dimethylaminoethane carbamoyl sterols.

[0038] The term "polynucleotide" as used herein refers to a polymer containing at least two deoxyribonucleotides or ribonucleotides in either single- or double-stranded form and includes DNA, RNA, and hybrids thereof. DNA may be in the form of antisense molecules, plasmid DNA, cDNA, PCR products, or vectors. RNA may be in the form of small hairpin RNA (shRNA), messenger RNA (mRNA), self-amplifying messenger

RNA (samRNA or saRNA), antisense RNA, miRNA, micRNA, multivalent RNA, dicer substrate RNA or viral RNA (vRNA), and combinations thereof. Polynucleotides include those containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, and which have
5 similar binding properties as the reference polynucleotide. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2'-O-methyl ribonucleotides, and peptide-nucleic acids (PNAs). Unless specifically limited, the term encompasses polynucleotides containing known analogues of natural nucleotides that have similar binding properties as the
10 reference polynucleotide. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, single nucleotide polymorphisms, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the
15 third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzner et al., Nucleic Acid Res., 19:5081 (1991); Ohtsuka et al., J. Biol. Chem., 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes, 8:91-98 (1994)).

[0039] An "effective amount" or "therapeutically effective amount" of a
20 therapeutic polynucleotide is an amount sufficient to produce the desired effect, such as an increase or inhibition of expression of a target sequence in comparison to the normal expression level detected in the absence of the polynucleotide. Suitable assays for measuring expression of a target gene or target sequence include, examination of protein or RNA levels using techniques known to those of skill in the art such as dot blots,
25 northern blots, in situ hybridization, ELISA, immunoprecipitation, enzyme function, fluorescence or luminescence of suitable reporter proteins, as well as phenotypic assays.

[0040] As used herein, "prodrug" is meant to indicate a compound that may be converted under physiological conditions or by solvolysis to a compound of any one or more of the formulae described herein. Thus, the term "prodrug" refers to a metabolic
30 precursor of such a compound that is pharmaceutically acceptable. A prodrug may be inactive when administered to a subject in need thereof, but is converted *in vivo* to an

active form. The term may also include any covalently bonded carriers, which release the active compound *in vivo* when such prodrug is administered to a mammalian subject. Prodrugs of a compound of formula I, or other formulae described herein, may be prepared by modifying functional groups present in the compound in such a way that the
5 modifications are cleaved, either in routine manipulation or *in vivo*, to the parent compound.

[0041] "Pharmaceutically acceptable carrier, diluent or excipient", or like terms, refers to any ingredient other than the compounds described herein (for example, a vehicle capable of suspending, complexing, or dissolving the active compound) and
10 having the properties of being substantially nontoxic and non-inflammatory in a patient. Excipients may include, for example: anti-adherents, antioxidants, binders, coatings, compression aids, disintegrants, dyes (colors), emollients, emulsifiers, fillers (diluent), film formers or coatings, flavors, fragrances, glidants (flow enhancers), lubricants, preservatives, printing inks, sorbents, suspending or dispersing agents, sweeteners, and
15 waters of hydration. Exemplary excipients include, but are not limited to: butylated hydroxy toluene (BHT), calcium carbonate, calcium phosphate (dibasic), calcium stearate, croscarmellose, crosslinked polyvinyl pyrrolidone, citric acid, crospovidone, cysteine, ethylcellulose, gelatin, hydroxypropyl cellulose, hydroxypropyl methylcellulose, lactose, magnesium stearate, maltitol, mannitol, methionine,
20 methylcellulose, methyl paraben, microcrystalline cellulose, polyethylene glycol, polyvinyl pyrrolidone, povidone, pregelatinized starch, propyl paraben, retinyl palmitate, shellac, silicon dioxide, sodium carboxymethyl cellulose, sodium citrate, sodium starch glycolate, sorbitol, starch (corn), stearic acid, sucrose, talc, titanium dioxide, vitamin A, vitamin E (alpha-tocopherol), vitamin C, xylitol, and other species disclosed herein.

25 [0042] "Pharmaceutically acceptable salt" includes both acid and base addition salts. Lists of suitable salts may be found in *Remington's Pharmaceutical Sciences*, 18th ed., Mack Publishing Company, Easton, PA, 1990, p. 1445, and *Journal of Pharmaceutical Science*, 66, 2-19 (1977). Acid addition salts are those which retain the biological effectiveness and properties of the free bases, which are not biologically or
30 otherwise undesirable, and which are formed with inorganic acids such as, but are not limited to, hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid

and the like, and organic acids such as, but not limited to, acetic acid, 2,2-dichloroacetic acid, adipic acid, alginic acid, ascorbic acid, aspartic acid, benzenesulfonic acid, benzoic acid, 4-acetamidobenzoic acid, camphoric acid, camphor-10-sulfonic acid, capric acid, caproic acid, caprylic acid, carbonic acid, cinnamic acid, citric acid, cyclamic acid, 5 dodecylsulfuric acid, ethane-1,2-disulfonic acid, ethanesulfonic acid, 2-hydroxyethanesulfonic acid, formic acid, fumaric acid, galactaric acid, gentisic acid, glucoheptonic acid, gluconic acid, glucuronic acid, glutamic acid, glutaric acid, 2-oxo-glutaric acid, glycerophosphoric acid, glycolic acid, hippuric acid, isobutyric acid, lactic acid, lactobionic acid, lauric acid, maleic acid, malic acid, malonic acid, mandelic acid, 10 methanesulfonic acid, mucic acid, naphthalene-1,5-disulfonic acid, naphthalene-2-sulfonic acid, 1-hydroxy-2-naphthoic acid, nicotinic acid, oleic acid, orotic acid, oxalic acid, palmitic acid, pamoic acid, propionic acid, pyroglutamic acid, pyruvic acid, salicylic acid, 4-aminosalicylic acid, sebacic acid, stearic acid, succinic acid, tartaric acid, thiocyanic acid, p-toluenesulfonic acid, trifluoroacetic acid, undecylenic acid, and 15 the like. Base addition salts are those which retain the biological effectiveness and properties of the free acids, which are not biologically or otherwise undesirable. These salts are prepared from addition of an inorganic base or an organic base to the free acid. Salts derived from inorganic bases include, but are not limited to, the sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts 20 and the like. Preferred inorganic salts are the ammonium, sodium, potassium, calcium, and magnesium salts. Salts derived from organic bases include, but are not limited to, salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as ammonia, isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, 25 diethanolamine, ethanolamine, deanol, 2-dimethylaminoethanol, 2-diethylaminoethanol, dicyclohexylamine, lysine, arginine, histidine, caffeine, procaine, hydrabamine, choline, betaine, benethamine, benzathine, ethylenediamine, glucosamine, methylglucamine, theobromine, triethanolamine, tromethamine, purines, piperazine, piperidine, N-ethylpiperidine, polyamine resins and the like.

30 [0043] As used herein, "stereoisomer" refers to a compound made up of the same atoms bonded by the same bonds but having different three-dimensional structures,

which are not interchangeable. The present invention contemplates various stereoisomers and mixtures thereof and includes "enantiomers", which refers to two stereoisomers whose molecules are nonsuperimposable mirror images of one another.

[0044] As used herein, the term "biodegradable group" is a group that may facilitate faster metabolism of a lipid in a mammal. An ester is a suitable biodegradable group and compounds of the present disclosure present two such groups for improved *in vivo* biodegradability.

[0045] As used herein, "encapsulation efficiency" refers to the amount of a polynucleotide that becomes part of an LNP composition, relative to the initial total amount of polynucleotide used in the preparation of the LNP composition. For example, if 92 mg of polynucleotides are encapsulated in an LNP composition out of a total 100 mg of polynucleotide initially provided to the composition, the encapsulation efficiency may be given as 92%. As used herein, "encapsulation" may refer to complete, substantial, or partial enclosure, confinement, surrounding, or encasement.

[0046] As used herein, the term "subject" shall be taken to mean any animal, such as a mammal, and including humans. Exemplary subjects include but are not limited to humans and non-human primates. For example, the subject is a human.

[0047] As used herein, the term "mammal" includes humans and both domestic animals such as laboratory animals and household pets (e.g., cats, dogs, swine, cattle, sheep, goats, horses, rabbits), and non-domestic animals such as wildlife and the like.

[0048] As used herein, as depicted in chemical structures, C_{A-B} will represent the number of carbons in that location within the chemical structure. In an embodiment, C₁₋₃ will be taken to mean that one to three carbon atoms (i.e., a carbon chain) are in that location within the chemical structure. In an embodiment, C₁₋₃ is taken to mean that one (-CH₂-), two (-CH₂-CH₂-), or three carbon atoms (-CH₂-CH₂-CH₂-) (i.e., carbon chain) are present in that location within the chemical structure. In the instance where C₀ is provided for, it will be understood that the carbon atom is absent, and instead a bond is formed between adjacent atoms within the chemical structure.

[0049] As used herein, the term "alkyl" encompasses both straight chain (i.e., linear) and branched chain hydrocarbon groups. Examples of alkyl groups include

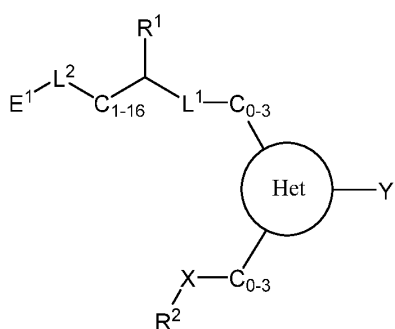
methyl, ethyl, n-propyl, iso-propyl, n-butyl, t-butyl, i-butyl, sec-butyl, pentyl, and hexyl groups. In one example, the alkyl group is of one to thirty carbon atoms (i.e., C₁₋₃₀alkyl) or one to 25 carbon atoms (i.e., C₁₋₂₅alkyl) or one to 20 carbon atoms (i.e., C₁₋₂₀alkyl).

[0050] As used herein, the term “alkanol” encompasses both straight chain (i.e., linear) and branched chain hydrocarbon groups to which an -OH (alcohol) group is substituted. Examples of alkanol groups include methanol, ethanol, and propanol groups. In one example, the alkanol group is of one to five carbon atoms (i.e., C₁₋₅alkanol).

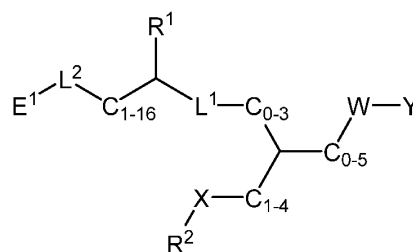
Compounds

[0051] In embodiments, the compounds of the present disclosure may provide for advantages over other select prior art ionizable cationic lipid compounds including one or more of: improved complexation with a polynucleotide; beneficial pK_a properties; improved encapsulation efficiency as part of an LNP; reduced toxicity; improved biodegradability; improved *in vivo* clearance; desirable N:P ratio when complexed with a polynucleotide; desirable polydispersity index for the LNPs comprising them; and improved LNP formation.

[0052] In one broad form, the present disclosure provides for a compound of Formula I-Het or a compound of Formula II:



Formula I-Het;



Formula II;

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

Het is a nitrogen heterocycle;

X is selected from the group consisting of -S-, -O- and -C-;

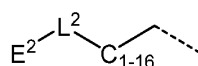
16

E^1 is a linear or branched $-C_{1-30}$ -alkyl;

R^1 is selected from the group consisting of $-H$ and Formula IA;

R^2 is selected from the group consisting of linear or branched $-C_{1-30}$ -alkyl and Formula IA;

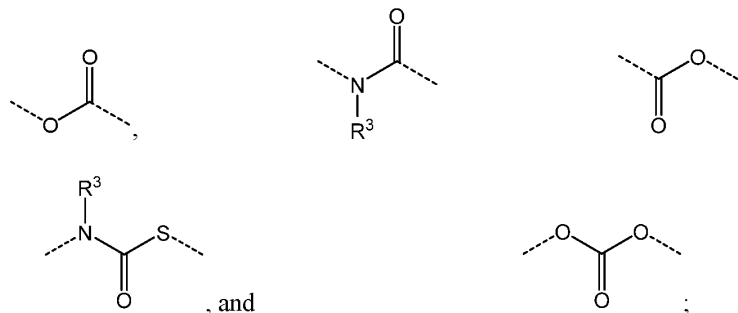
5



Formula IA;

E^2 , if present, is a linear or branched $-C_{1-30}$ -alkyl;

L^1 is selected from the group consisting of:

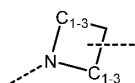


R^3 is selected from the group consisting of $-H$ and linear or branched $-C_{1-8}$ -alkyl;

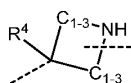
10

L^2 is selected from $-OC(O)-$ and $-C(O)O-$;

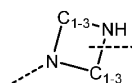
W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



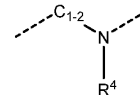
Formula IIA;



Formula IIB;



Formula IIC;



Formula IID;

R^4 , if present, is selected from the group consisting of $-H$ and linear or branched $-C_{1-5}$ -alkyl;

Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol; and

dashed lines represent a bond to an adjacent atom in the compound of Formula I-Het or the compound of Formula II.

5 [0053] In embodiments of Formula I-Het, Het is selected from a 4-, 5-, and 6-membered nitrogen heterocycle.

[0054] In embodiments of Formula I-Het, Het is selected from a 4-, 5-, and 6-membered nitrogen heterocycle comprising one or two ring nitrogen atoms as the only ring heteroatoms.

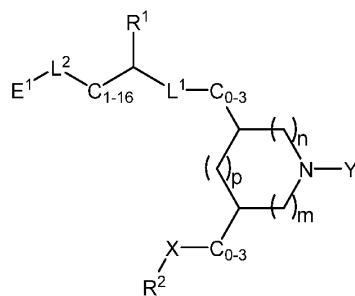
10 [0055] In embodiments of Formula I-Het, Het is selected from a 4-, 5-, and 6-membered nitrogen heterocycle comprising one ring nitrogen atom as the only ring heteroatom.

[0056] In embodiments of Formula I-Het, Het is selected from a 5-, or 6-membered nitrogen heterocycle comprising one ring nitrogen atom as the only ring heteroatom.

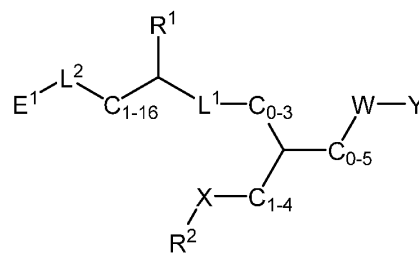
15 [0057] In embodiments of Formula I-Het, Het is a 6-membered nitrogen heterocycle comprising one ring nitrogen atom as the only ring heteroatom.

[0058] The further moieties of Formula I-Het may be selected from any of those described herein in relation to Formula I which represents a subset of Formula I-Het compounds.

20 [0059] In embodiments, the present disclosure provides for a compound of Formula I or a compound of Formula II:



Formula I;



Formula II;

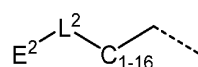
or a pharmaceutically acceptable salt, prodrug, or stereoisomer thereof, wherein:

X is selected from the group consisting of -S-, -O- and -C-;

E¹ is a linear or branched -C₁₋₃₀-alkyl;

R¹ is selected from the group consisting of -H and Formula IA;

- 5 R² is selected from the group consisting of linear or branched -C₁₋₃₀-alkyl and Formula IA;



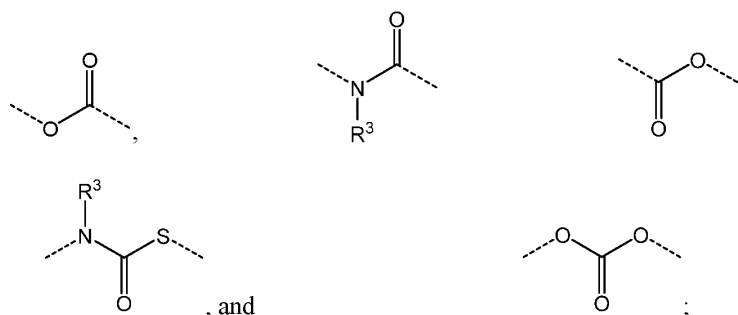
Formula IA;

E², if present, is a linear or branched -C₁₋₃₀-alkyl;

- 10 m and n are each independently an integer from 0 to 3;

p is an integer from 0 to 2;

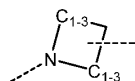
L¹ is selected from the group consisting of:



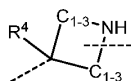
R³ is selected from the group consisting of -H and linear or branched -C₁₋₈-alkyl;

L² is selected from -OC(O)- and -C(O)O-;

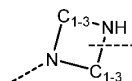
- 15 W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



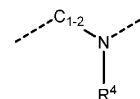
Formula IIA;



Formula IIB;



Formula IIC;



Formula IID;

R^4 , if present, is selected from the group consisting of -H and linear or branched -C₁₋₅-alkyl;

Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol; and

- 5 dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

[0060] In the compound of Formula I and the Compound of Formula II, X is selected from the group consisting of -S-, -O- and -C-. In an embodiment, X is -S-. In an embodiment, X is -O-. In an embodiment, X is -C-.

- 10 [0061] Preferably, in the compound of Formula I and the Compound of Formula II, X is -S- or -O-. Most preferably, in the compound of Formula I and the Compound of Formula II, X is -S-.

- [0062] In the compound of Formula I and the compound of Formula II, E¹ is a linear or branched -C₁₋₃₀-alkyl. In some embodiments, E¹ is a linear -C₃₀-alkyl, -C₂₅-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, or -C₈-alkyl. In an embodiment, E¹ is a linear -C₁₋₃₀-alkyl. In an embodiment, E¹ is a linear -C₁₋₂₀-alkyl. In an embodiment, E¹ is a linear -C₁₋₁₈-alkyl. In an embodiment, E¹ is a linear -C₁₋₁₆-alkyl. In an embodiment, E¹ is a linear -C₁₋₁₅-alkyl. In an embodiment, E¹ is a linear -C₁₋₁₂-alkyl. In an embodiment, E¹ is a linear -C₁₋₁₀-alkyl. In an embodiment, E¹ is a linear -C₁₋₈-alkyl. In some embodiments, E¹ is a linear -C₁₀₋₃₀-alkyl. In some embodiments, E¹ is a linear -C₁₅₋₃₀-alkyl. In some embodiments, E¹ is a linear -C₂₀₋₃₀-alkyl. In some embodiments, E¹ is a branched -C₃₀-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, or -C₈-alkyl. In an embodiment, E¹ is a branched -C₁₋₃₀-alkyl. In an embodiment, E¹ is a branched -C₁₋₂₀-alkyl. In an embodiment, E¹ is a branched -C₁₋₁₈-alkyl. In an embodiment, E¹ is a branched -C₁₋₁₆-alkyl. In an embodiment, E¹ is a branched -C₁₋₁₅-alkyl. In an embodiment, E¹ is a branched -C₁₋₁₂-alkyl. In an

embodiment, E^1 is a branched -C₁₋₁₀-alkyl. In an embodiment, E^1 is a branched -C₁₋₈-alkyl. In some embodiments, E^1 is a branched -C₁₀₋₃₀-alkyl. In some embodiments, E^1 is a branched -C₁₅₋₃₀-alkyl. In some embodiments, E^1 is a branched -C₂₀₋₃₀-alkyl.

[0063] In the compound of Formula I and the compound of Formula II, R^1 is selected from the group consisting of -H (i.e., hydrogen) and Formula IA. In one embodiment, R^1 is -H. In one embodiment, R^1 is Formula IA.

[0064] In the compound of Formula I and the compound of Formula II, R^2 is selected from the group consisting of -linear or branched C₁₋₃₀-alkyl and Formula IA. In an embodiment, R^2 is a linear or branched -C₁₋₃₀-alkyl. In some embodiments, R^2 is a linear -C₃₀-alkyl, -C₂₅-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, or -C₈-alkyl. In an embodiment, R^2 is a linear -C₁₋₃₀-alkyl. In an embodiment, R^2 is a linear -C₁₋₂₀-alkyl. In an embodiment, R^2 is a linear -C₁₋₁₈-alkyl. In an embodiment, R^2 is a linear -C₁₋₁₆-alkyl. In an embodiment, R^2 is a linear -C₁₋₁₅-alkyl. In an embodiment, R^2 is a linear -C₁₋₁₂-alkyl. In an embodiment, R^2 is a linear -C₁₋₁₀-alkyl. In an embodiment, R^2 is a linear -C₁₋₈-alkyl. In some embodiments, R^2 is a linear -C₁₀₋₃₀-alkyl. In some embodiments, R^2 is a linear -C₁₅₋₃₀-alkyl. In some embodiments, R^2 is a linear -C₂₀₋₃₀-alkyl. In some embodiments, R^2 is a branched -C₃₀-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, or -C₈-alkyl. In an embodiment, R^2 is a branched -C₁₋₃₀-alkyl. In an embodiment, R^2 is a branched -C₁₋₂₀-alkyl. In an embodiment, R^2 is a branched -C₁₋₁₈-alkyl. In an embodiment, R^2 is a branched -C₁₋₁₆-alkyl. In an embodiment, R^2 is a branched -C₁₋₁₅-alkyl. In an embodiment, R^2 is a branched -C₁₋₁₂-alkyl. In an embodiment, R^2 is a branched -C₁₋₁₀-alkyl. In an embodiment, R^2 is a branched -C₁₋₈-alkyl. In some embodiments, R^2 is a branched -C₁₀₋₃₀-alkyl. In some embodiments, R^2 is a branched -C₁₅₋₃₀-alkyl. In some embodiments, R^2 is a branched -C₂₀₋₃₀-alkyl. In an embodiment, R^2 is Formula IA.

[0065] In the compound of Formula I and the compound of Formula II, R^1 and R^2 may each independently be Formula IA. In an embodiment, R^1 is Formula IA and R^2 is a linear or branched -C₁₋₃₀-alkyl. In an embodiment, R^1 is Formula IA and R^2 is a linear or branched -C₁₋₃₀-alkyl. In an embodiment, R^1 is Formula IA and R^2 is a linear or branched -C₁₋₁₆-alkyl. In an embodiment, R^1 is -H and R^2 is Formula IA. In an embodiment, R^1 is -H and R^2 is -C₁₋

₁₆-alkyl. In an embodiment, R¹ is Formula IA and R² is Formula IA. In an embodiment, at least one of R¹ and R² is Formula IA.

[0066] In the compound of Formula I and the compound of Formula II, Formula IA has the structure:



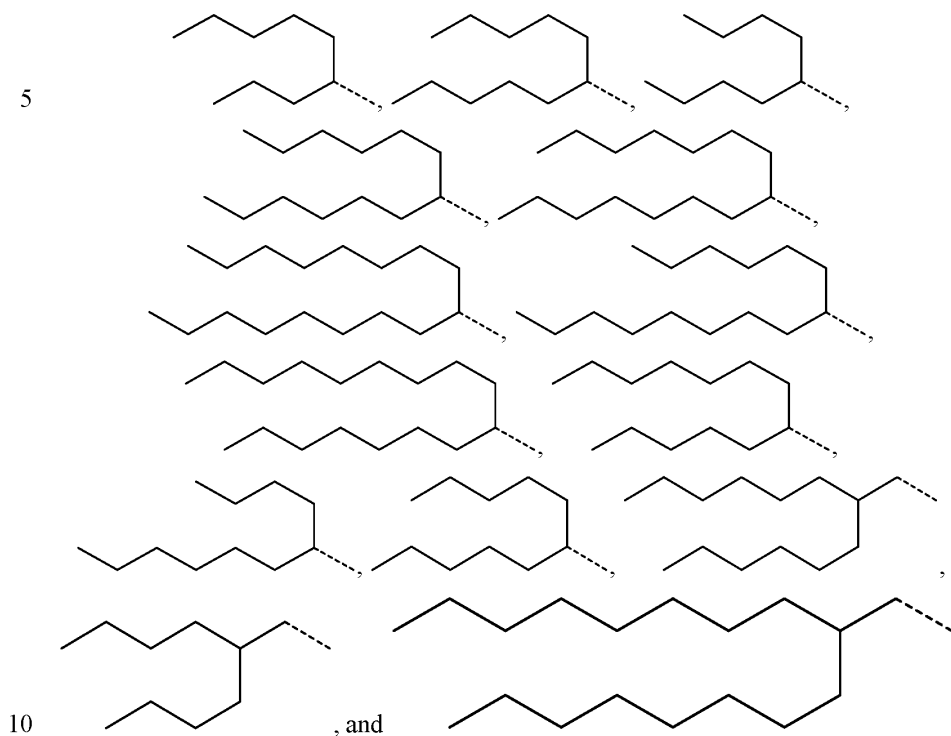
Formula IA.

[0067] The dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

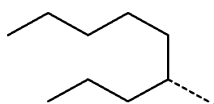
[0068] It will be appreciated that in the compound of Formula I and the compound
10 of Formula II, Formula IA may not be present (i.e., R¹ is -H and R² is -C₁₋₃₀-alkyl). In such an instance, it will therefore follow that in the compound of Formula I and the compound of Formula II, E² will not be present. Alternatively, in the compound of Formula I and the compound of Formula II, where Formula IA is present, E² will also be present.

[0069] In the compound of Formula IA, E², if present, is a linear or branched -C₁₋₃₀-alkyl. In an embodiment, E² is a linear or branched -C₁₋₃₀-alkyl. In some embodiments, E² is a linear -C₃₀-alkyl, -C₂₅-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, or -C₈-alkyl. In an embodiment, E² is a linear -C₁₋₃₀-alkyl. In an embodiment, E² is a linear -C₁₋₂₀-alkyl. In an embodiment, E² is a linear -C₁₋₁₈-alkyl. In an embodiment, E² is a linear -C₁₋₁₆-alkyl. In an embodiment, E² is a linear -C₁₋₁₅-alkyl. In an embodiment, E² is a linear -C₁₋₁₂-alkyl. In an embodiment, E² is a linear -C₁₋₁₀-alkyl. In an embodiment, E² is a linear -C₁₋₈-alkyl. In some embodiments, E² is a linear -C₁₀₋₃₀-alkyl. In some embodiments, E² is a linear -C₁₅₋₃₀-alkyl. In some embodiments, E² is a linear -C₂₀₋₃₀-alkyl. In some embodiments, E² is a branched -C₃₀-alkyl, -C₂₀-alkyl, -C₁₈-alkyl, -C₁₆-alkyl, -C₁₅-alkyl, -C₁₂-alkyl, -C₁₀-alkyl, -C₉-alkyl, or -C₈-alkyl. In an embodiment, E² is a branched -C₁₋₃₀-alkyl. In an embodiment, E² is a branched -C₁₋₂₀-alkyl. In an embodiment, E² is a branched -C₁₋₁₈-alkyl. In an embodiment, E² is a branched -C₁₋₁₆-alkyl. In an embodiment, E² is a branched -C₁₋₁₅-alkyl. In an embodiment, E² is a branched -C₁₋₁₂-alkyl. In an embodiment, E² is a branched -C₁₋₁₀-
25

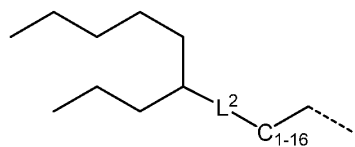
alkyl. In an embodiment, E² is a branched -C₁₋₈-alkyl. In some embodiments, E² is a branched -C₁₀₋₃₀-alkyl. In some embodiments, E² is a branched -C₁₅₋₃₀-alkyl. In some embodiments, E² is a branched -C₂₀₋₃₀-alkyl. In embodiments, E² has the structure selected from the group consisting of:



[0070] In an embodiment, E^2 has the structure:



[0071] Accordingly, in an embodiment, Formula IA has the structure:

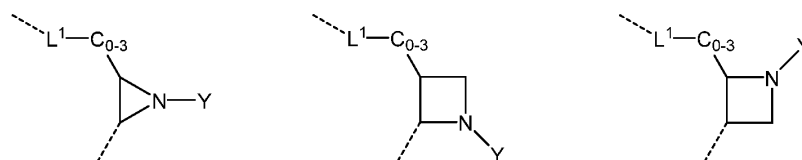


[0072] In the compound of Formula I, n and m are each independently an integer from 0 to 3 (i.e., 0, 1, 2, or 3). It will be understood that the integer describes the number of carbon atoms present in the structure within the corresponding parentheses. In the instance where m or n are 0, it will be appreciated that a direct bond is instead made
 5 between adjacent atoms. In an embodiment, m is an integer from 0 to 3. In an embodiment, m is 0, 1, 2, or 3. In an embodiment, n is an integer from 0 to 3. In an embodiment, n is 0, 1, 2, or 3.

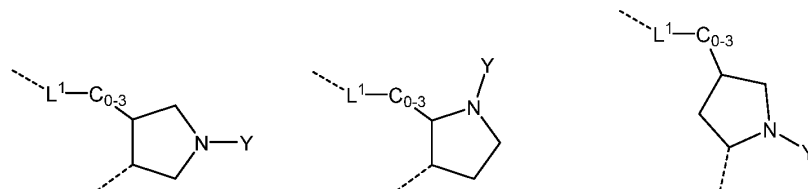
[0073] In the compound of Formula I, p is an integer from 0 to 2 (i.e., 0, 1, or 2). It will be understood that the integer describes the number of carbon atoms present in the structure within the corresponding parentheses. In an embodiment, p is 0. In the instance
 10 where p is 0, it will be appreciated that a directed bond is instead made between adjacent atoms. In an embodiment, p is 0, 1, or 2. In an embodiment, p is 1. In an embodiment, p is 2.

[0074] It will therefore be appreciated that, taken together, the integers of m, n, and
 15 p dictate the size of the heterocyclic group in the structure of the compound of Formula I. In an embodiment, the heterocyclic group in the structure of the compound of Formula I is a saturated heterocyclic group. In an embodiment, at least one of any of m, n, and p is an integer greater than 0. In an embodiment, at least two of any one of m, n, and p are an integer greater than 0. In an embodiment, p is 0, m is 0, n is 0, and the heterocyclic
 20 group is a 3-membered heterocyclic group. In an embodiment, p is 0, m is 1, n is 1, and the heterocyclic group is a 5-membered heterocyclic group. In an embodiment, p is 0, m is 1, n is 2, and the heterocyclic group is a 6-membered heterocyclic group. In an embodiment, p is 1, m is 1, n is 1, and the heterocyclic group is a 6-membered heterocyclic group.

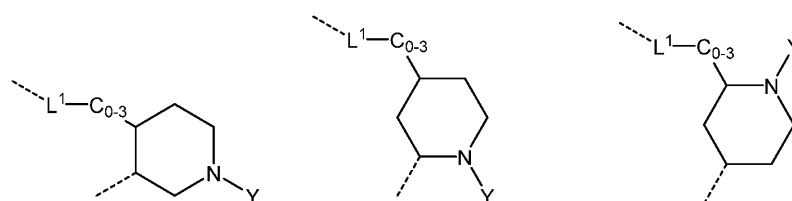
25 [0075] In an embodiment, the heterocyclic group of the compound of Formula I has a structure selected from the group consisting of:



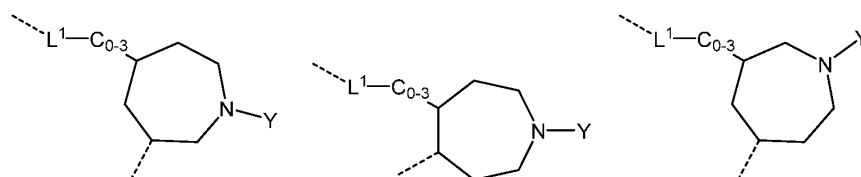
(i.e., m is 0, n is 0, p is 0); (i.e., m is 0, n is 1, p is 0); (i.e., m is 1, n is 0, p is 0);



(i.e., m is 1, n is 1, p is 0); (i.e., m is 2, n is 0, p is 0); (i.e., m is 0, n is 1, p is 1);



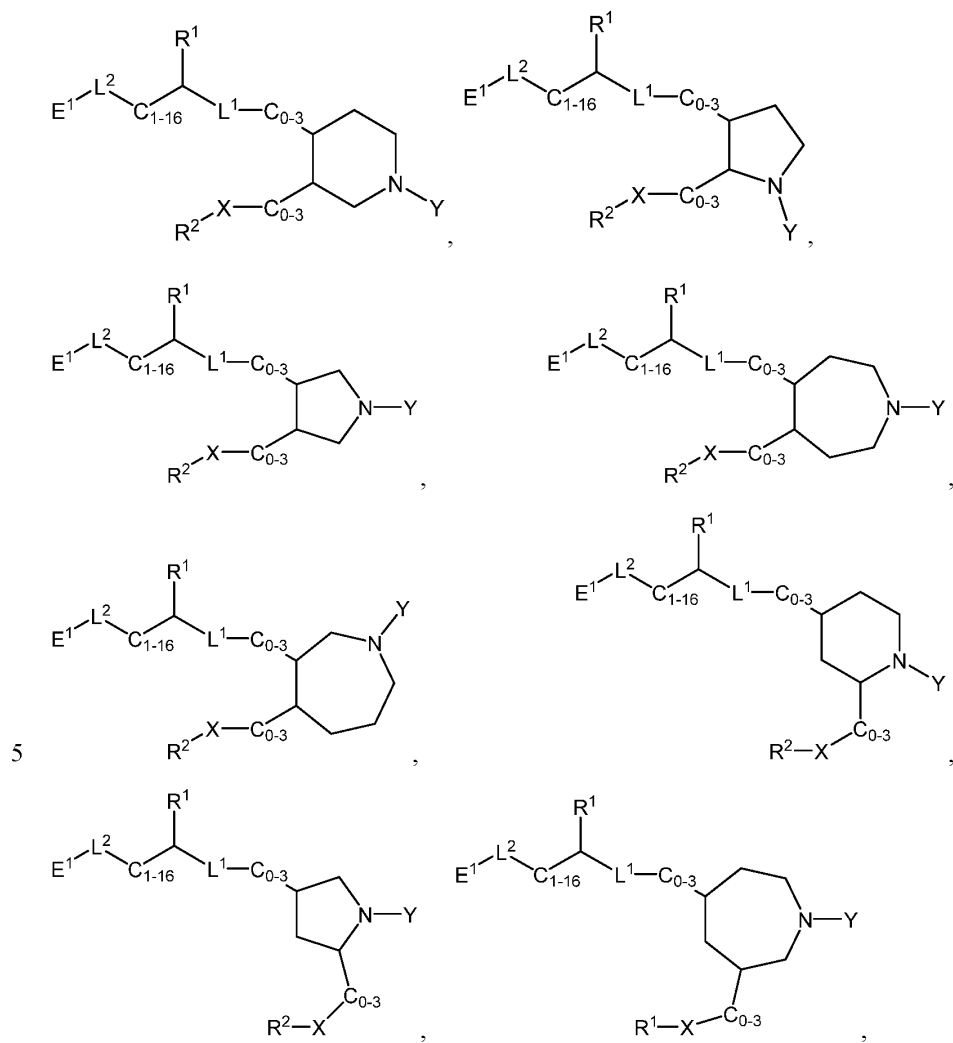
(i.e., m is 1, n is 2, p is 0); (i.e., m is 0, n is 2, p is 1); (i.e., m is 2, n is 0, p is 1);

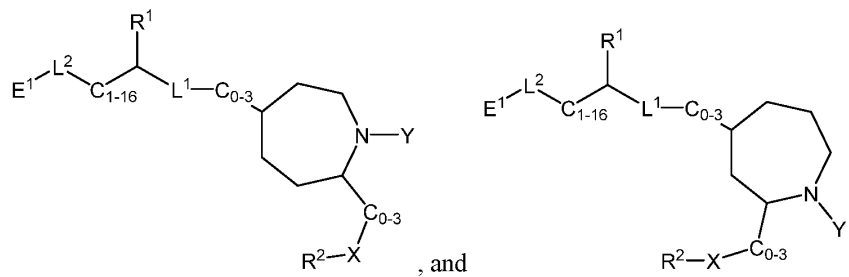


(i.e., m is 1, n is 2, p is 1); (i.e., m is 2, n is 2, p is 0); (i.e., m is 2, n is 1, p is 0).

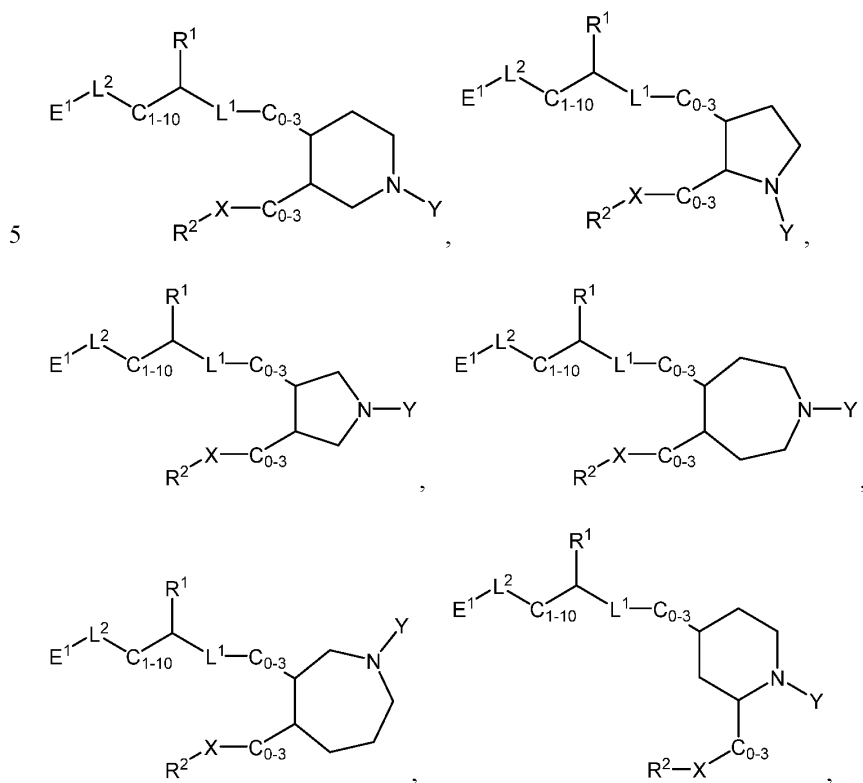
[0076] In the immediately above structures, shown to depict various heterocyclic groups of the compound of Formula I as contemplated by the present disclosure, the adjacent C₀₋₃ and L¹, while not considered to constitute the heterocyclic structure *per se*, are also shown to clarify the orientation of the heterocyclic group within the structure of the compound of Formula I. The dashed lines represent a bond to an adjacent atom in the remaining structure of the compound of Formula I. It will be appreciated that the integers of m, n, and p provide for numerous iterations of the heterocyclic group of the compound of Formula I, and those shown immediately-above are merely exemplary of a broader group of heterocycles contemplated by the present disclosure.

[0077] In an embodiment, the compound of Formula I is selected from the group consisting of:

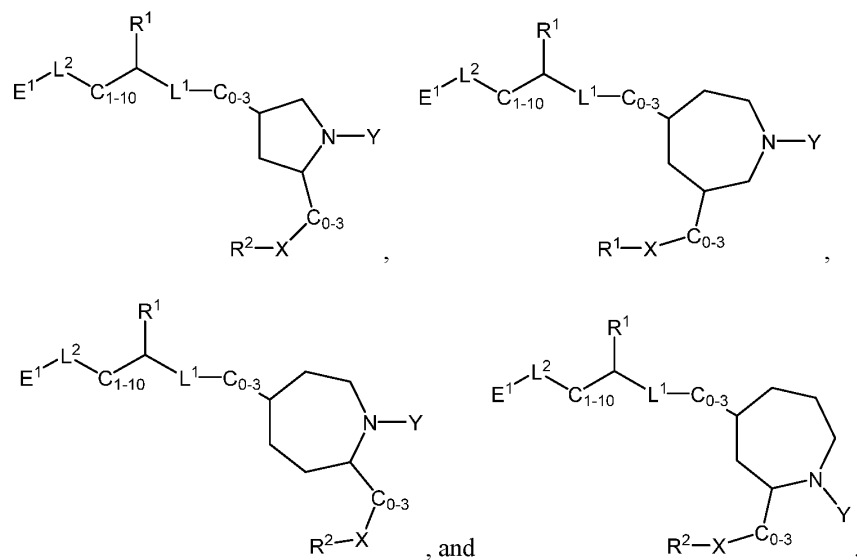




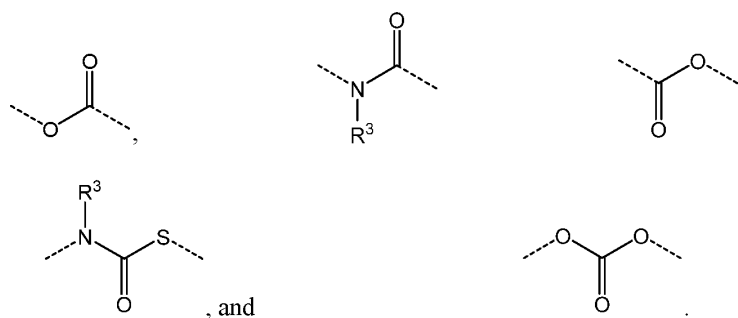
[0078] In an embodiment, the compound of Formula I is selected from the group consisting of:



27



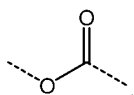
[0079] In the compound of Formula I and the compound of Formula II, L^1 is selected from the group consisting of:



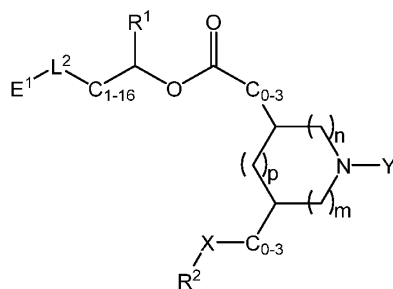
5

[0080] The dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

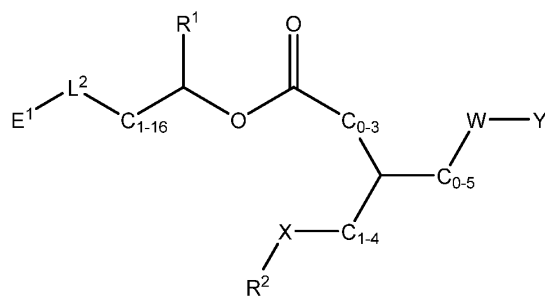
[0081] In an embodiment, L^1 is:



[0082] Accordingly, it follows that in an embodiment, the compound of Formula I has the structure:

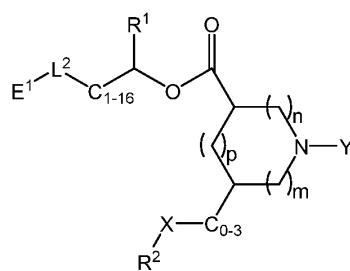


[0083] In an embodiment, the compound of Formula II has the structure:



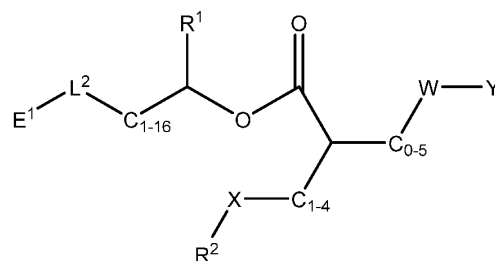
5

[0084] In an embodiment, the compound of Formula I has the structure:

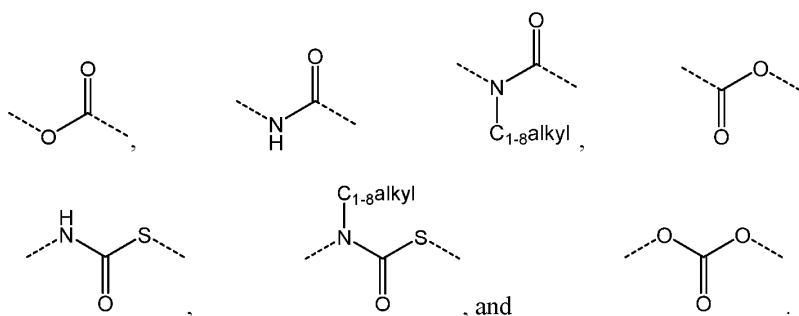


[0085] In an embodiment, the compound of Formula II has the structure:

29



- [0086] In the compound of Formula I and the compound of Formula II, R^3 , if present, is selected from the group consisting of -H and linear or branched - C_{1-8} -alkyl. In an embodiment, R^3 is -H. In an embodiment, R^3 is a linear - C_{1-8} -alkyl. In an embodiment, R^3 is a branched - C_{1-8} -alkyl. Accordingly, in an embodiment, the structure of L^1 in the compound of Formula I or the compound of Formula II is selected from the group consisting of:



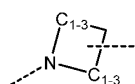
- [0087] In an embodiment, R^3 is a linear - C_{1-5} -alkyl. In an embodiment, R^3 is a linear - C_{1-3} -alkyl. In an embodiment, R^3 is a C_1 -alkyl (i.e., - CH_3).

[0088] In the compound of Formula I and the compound of Formula II, L^2 may be selected from -OC(O)- and -C(O)O-. That is, L^2 is an ester link which may be present in either orientation.

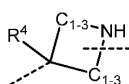
- [0089] In certain embodiments of the compound of Formula I and the compound of Formula II, L^2 is -OC(O)-. That is, the oxygen which is not the carbonyl oxygen is directly attached to E^1 .

[0090] In certain embodiments of the compound of Formula I and the compound of Formula II, L^2 is $-\text{C}(\text{O})\text{O}-$. That is, the carbonyl carbon is directly attached to E^1 .

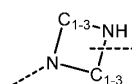
[0091] In the compound of Formula II, W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



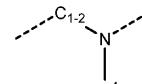
Formula IIA;



Formula IIB;



Formula IIC;

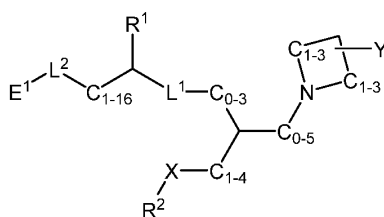


Formula IID;

5

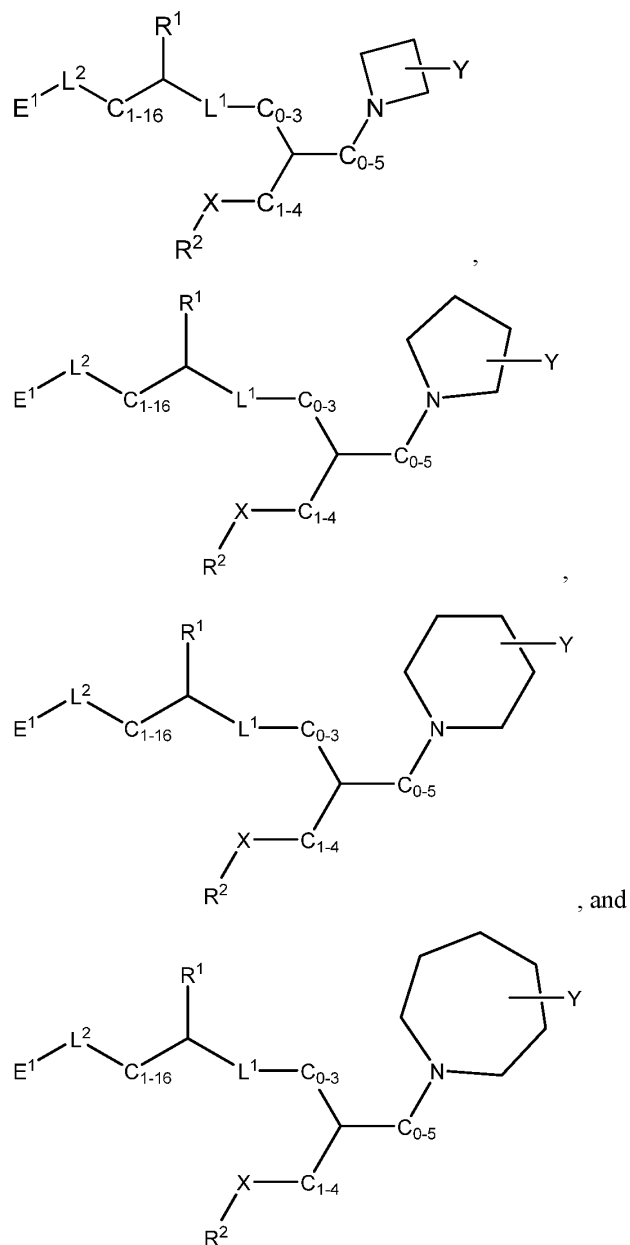
[0092] In each of Formula IIA, Formula IIB, Formula IIC, and Formula IID, dashed lines represent a bond to an adjacent atom in the compound of Formula II.

[0093] In an embodiment, in the compound of Formula II, W is Formula IIA. Accordingly, in an embodiment, the compound of Formula II has the structure:



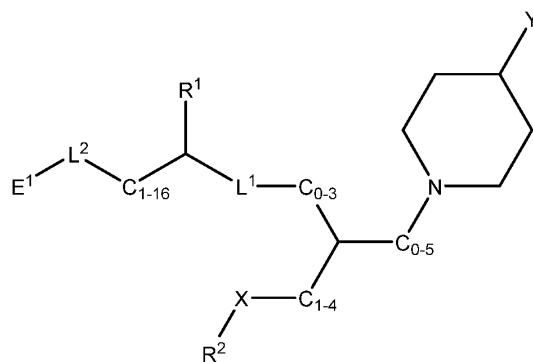
10

[0094] In an embodiment, the compound of Formula II has a structure selected from the group consisting of:



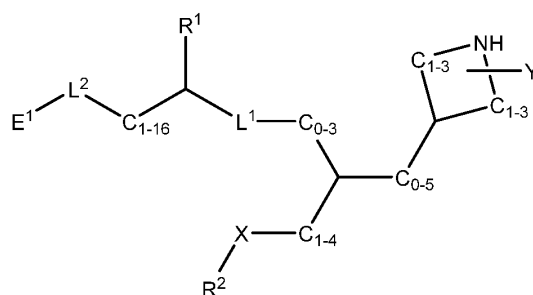
- 5 [0095] In the immediately-above structures, it will be understood that substituent - Y may be substituted at any suitable position, being a carbon atom, on the heterocyclic group. It will be appreciated that further heterocyclic groups of the Formula II are

contemplated by the present disclosure, and those depicted immediately-above are merely exemplary. In an embodiment, the compound of Formula II has the structure:

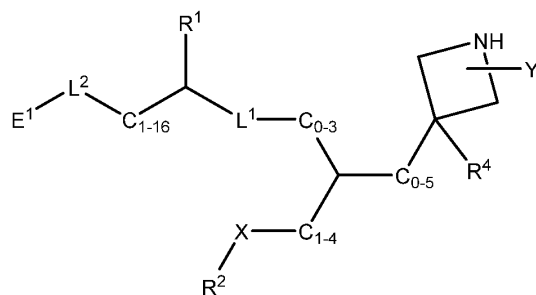


[0096] In an embodiment, in the compound of Formula II, W is Formula IIB.

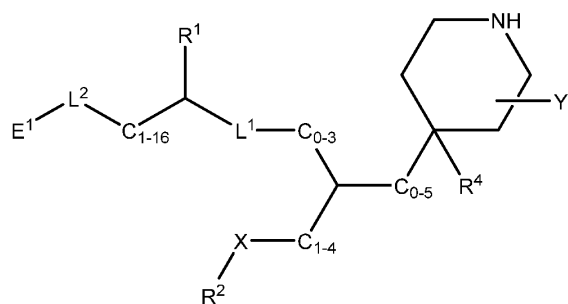
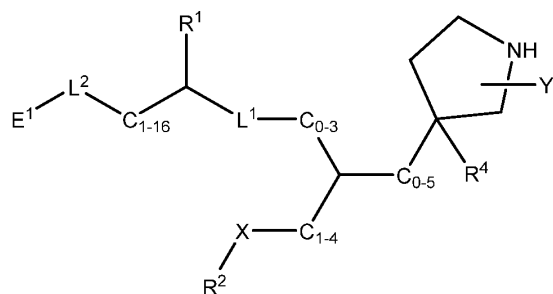
5 Accordingly, in an embodiment, the compound of Formula II has the structure:



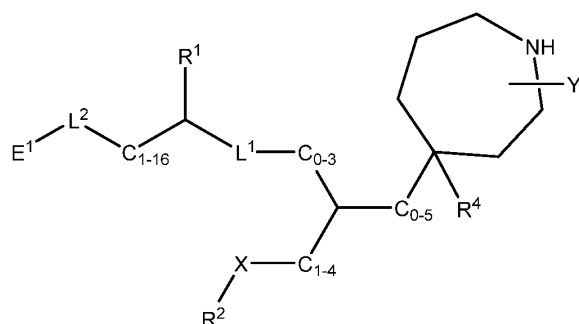
[0097] In an embodiment, the compound of Formula II has a structure selected from the group consisting of:



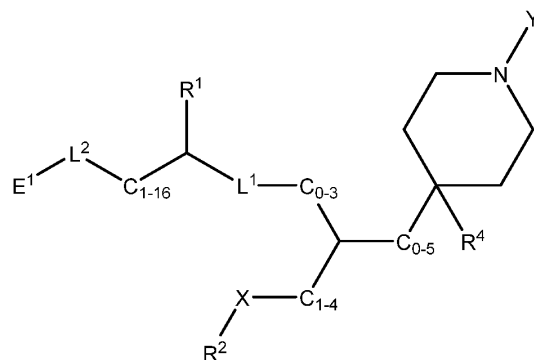
33



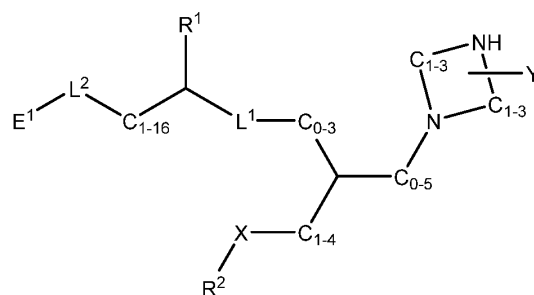
, and



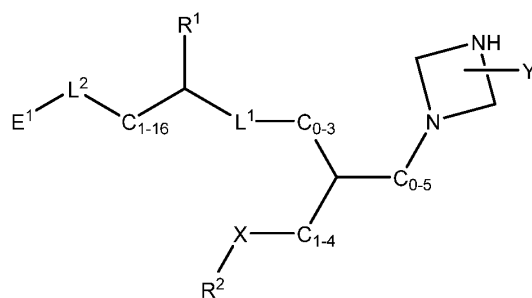
[0098] In the immediately above structures, it will be understood that substituent - Y may be substituted at any suitable position, being a carbon atom or the nitrogen atom, on the heterocyclic group. It will be appreciated that further heterocyclic groups of the Formula IIB are contemplated by the present disclosure, and those depicted immediately-above are merely exemplary. In an embodiment, the compound of Formula II has the structure:



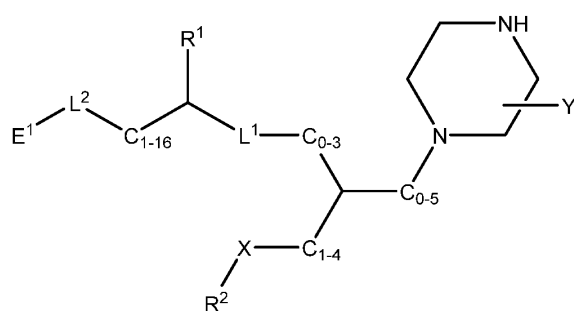
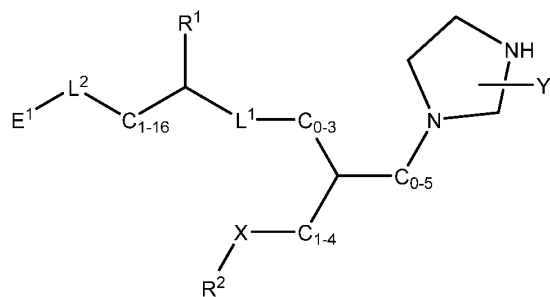
[0099] In an embodiment, in the compound of Formula II, W is Formula IIC. Accordingly, in an embodiment, the compound of Formula II has the structure:



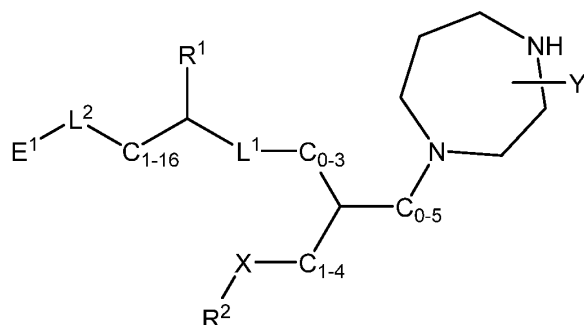
- 5 [0100] In an embodiment, the compound of Formula II has a structure selected from the group consisting of:



35

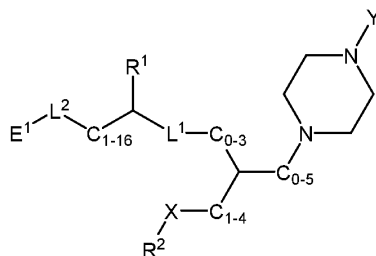


and

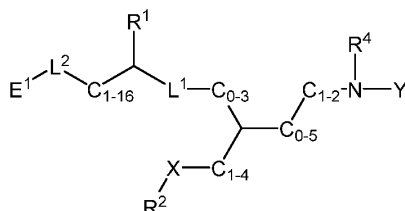


[0101] In the immediately above structures, it will be understood that substituent -
 5 Y may be substituted at any suitable position, being a carbon atom or the nitrogen atom,
 on the heterocyclic group. It will be appreciated that further heterocyclic groups of the
 Formula IIB are contemplated by the present disclosure, and those depicted immediately-
 above are merely exemplary. In an embodiment, the compound of Formula II has the
 structure:

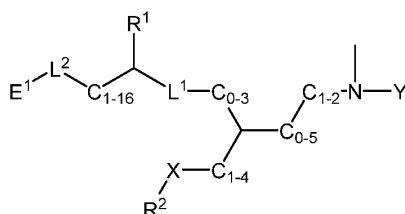
36



[0102] In an embodiment, in the compound of Formula II, W is Formula IID. Accordingly, in an embodiment, the compound of Formula II has the structure:



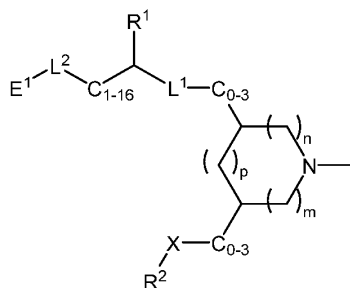
- 5 [0103] In the compound of Formula II, R^4 , if present, is selected from the group consisting of -H and linear or branched -C₁₋₅-alkyl. In an embodiment, R^4 is -H. In an embodiment, R^4 is linear -C₁₋₅-alkyl. In an embodiment, R^4 is branched -C₁₋₅-alkyl. In an embodiment, R^4 is a -C₁-alkyl (i.e., -CH₃). Accordingly, in an embodiment, the compound of Formula II has the structure:



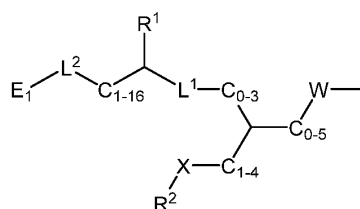
10

- [0104] In the compound of Formula I and the compound of Formula II, Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol. In an embodiment, Y is -H. In an embodiment, Y is a linear or branched -C₁₋₅-alkyl. In an embodiment, Y is a linear or branched -C₁₋₅-alkanol.
- 15 Accordingly, in an embodiment, Y is C₁-alkyl (i.e., -CH₃), and the compound of Formula I has the structure:

37

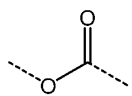


[0105] In an embodiment, Y is C₁-alkyl (i.e., -CH₃), and the compound of Formula II has the structure:

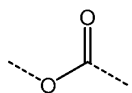


- 5 [0106] The compound of Formula I and the compound of Formula II may comprise any combination of the above-defined variables, namely X, E¹, R¹, R², E² (if present), m, n, p, L¹, R³ (if present), W, R⁴ (if present), and Y, as would be reasonably contemplated by the person skilled in the art. Preferably, in all of the structures shown above of Formula I or Formula II or substructures thereof, X is -S- or -O- or -C-, and most
- 10 preferably X is -S-.

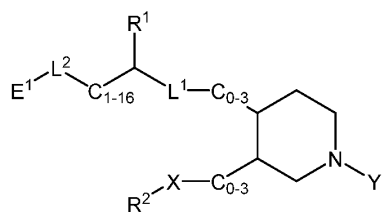
[0107] In an embodiment, in the compound of Formula I and the compound of Formula II, X is -S- and L¹ is:



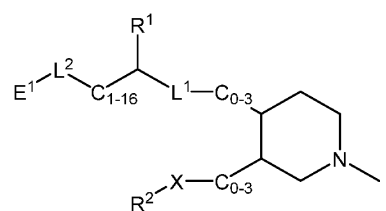
- [0108] In an embodiment, in the compound of Formula I and the compound of
- 15 Formula II, X is -O- and L¹ is:



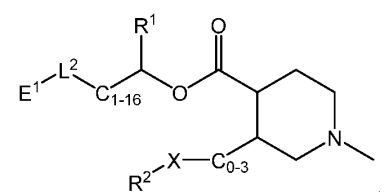
[0109] In an embodiment, the compound of Formula I is:



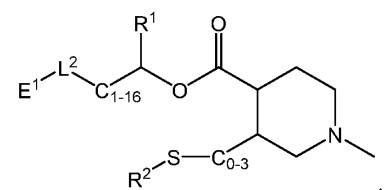
[0110] In an embodiment, the compound of Formula I is:



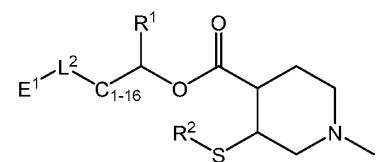
5 [0111] In an embodiment, the compound of Formula I is:



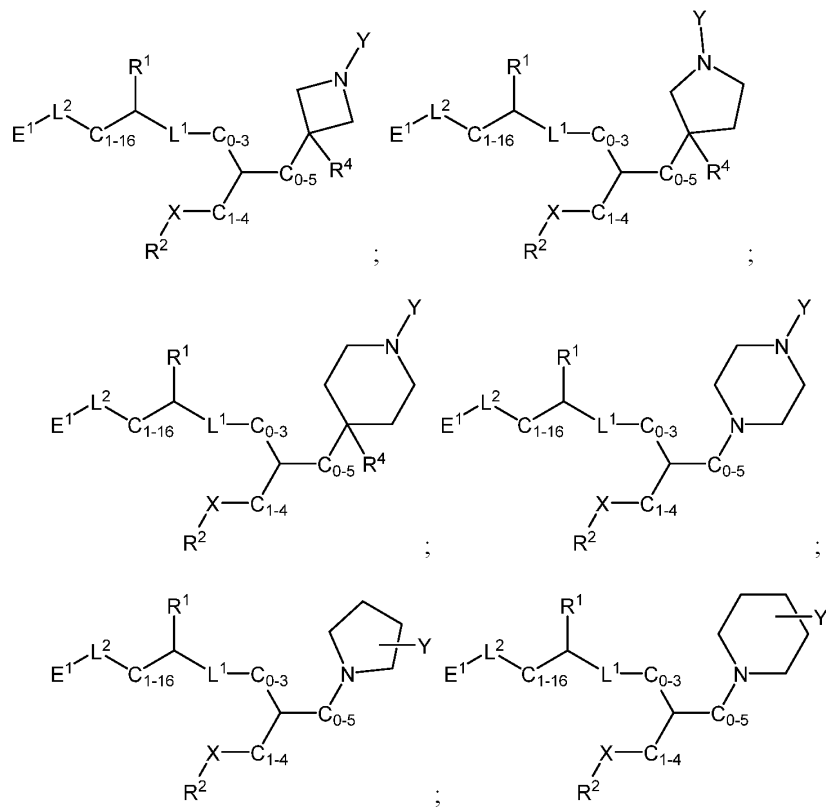
[0112] In an embodiment, the compound of Formula I is:



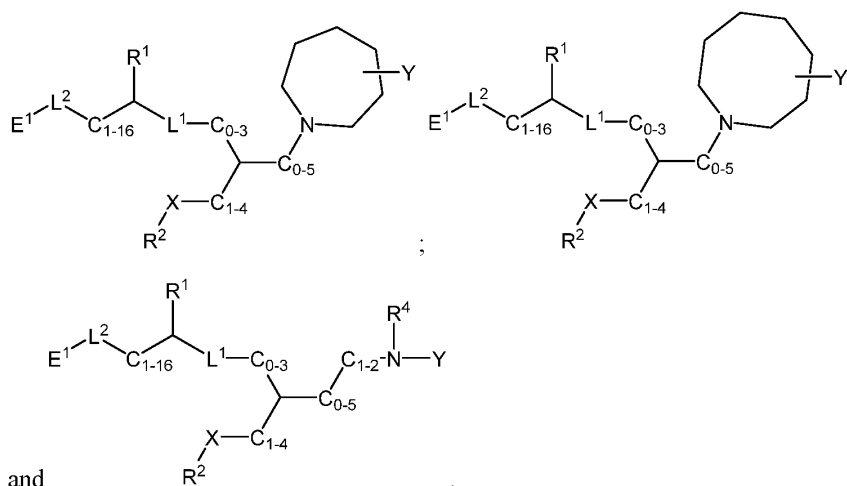
[0113] In an embodiment, the compound of Formula I is:



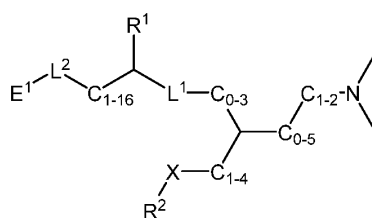
[0114] In an embodiment, the compound of Formula II is selected from the group consisting of:



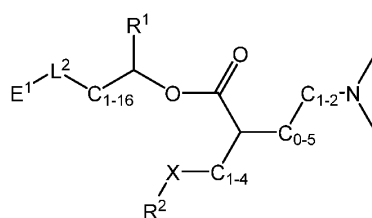
40



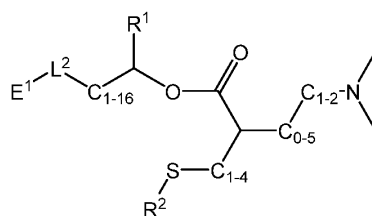
[0115] In an embodiment, the compound of Formula II is:



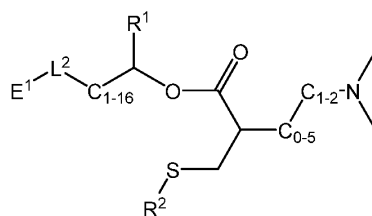
5 [0116] In an embodiment, the compound of Formula II is:



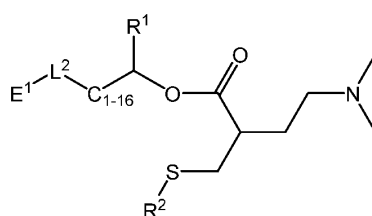
[0117] In an embodiment, the compound of Formula II is:



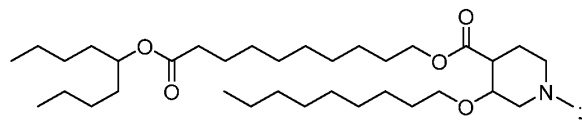
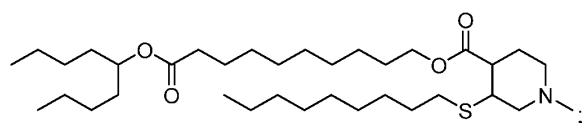
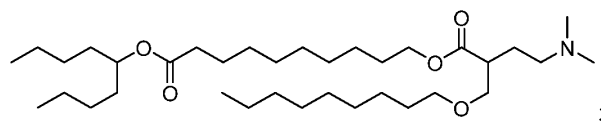
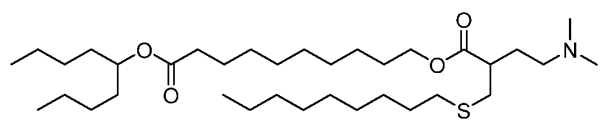
[0118] In an embodiment, the compound of Formula II is:



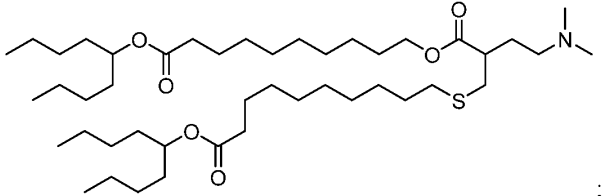
[0119] In an embodiment, the compound of Formula II is:



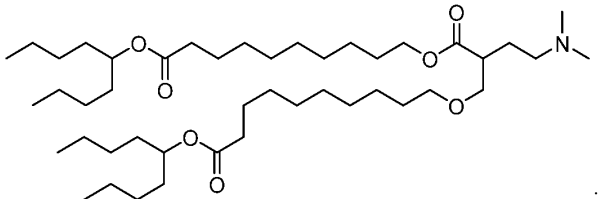
5 [0120] In an embodiment, the compound of Formula I or the compound of Formula II is selected from the group consisting of:



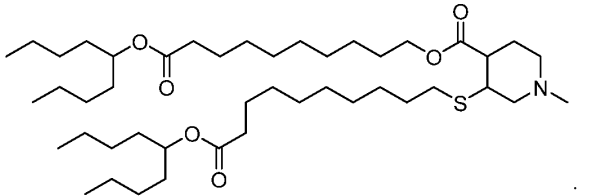
42



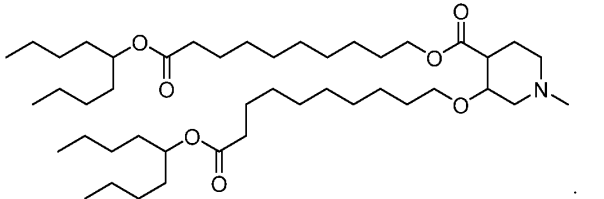
;



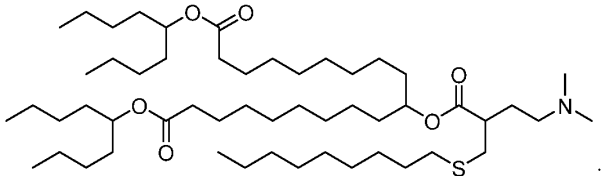
;



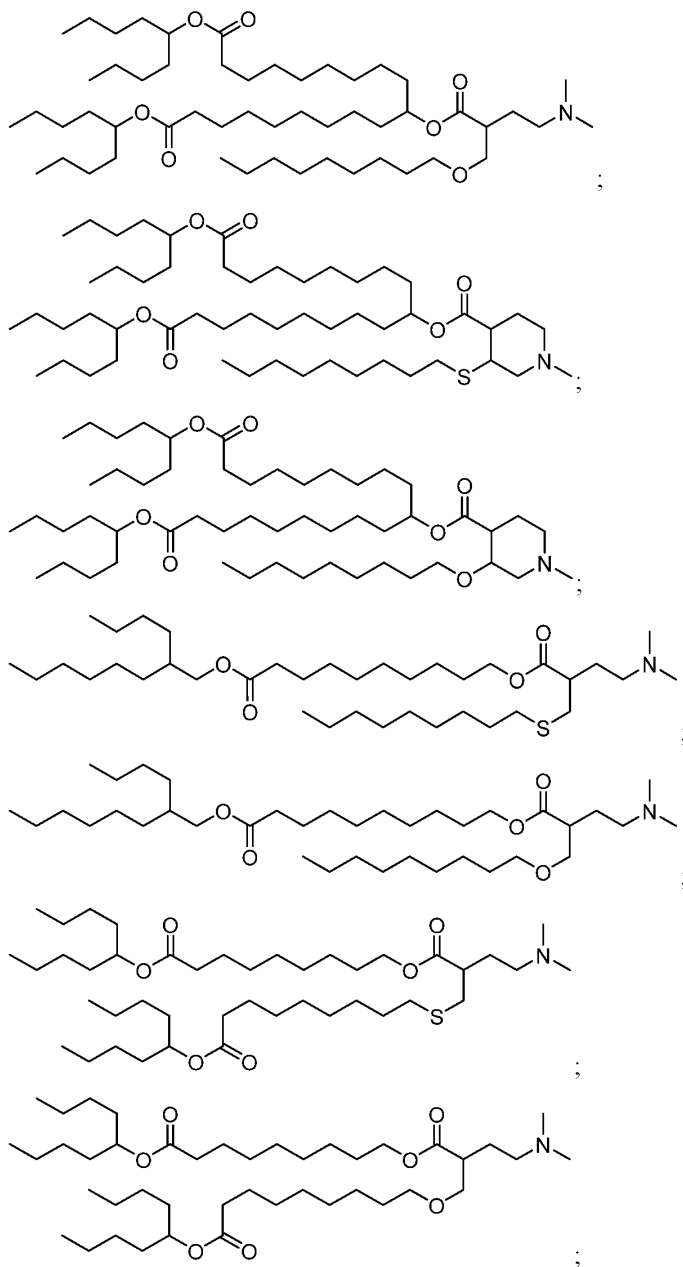
;

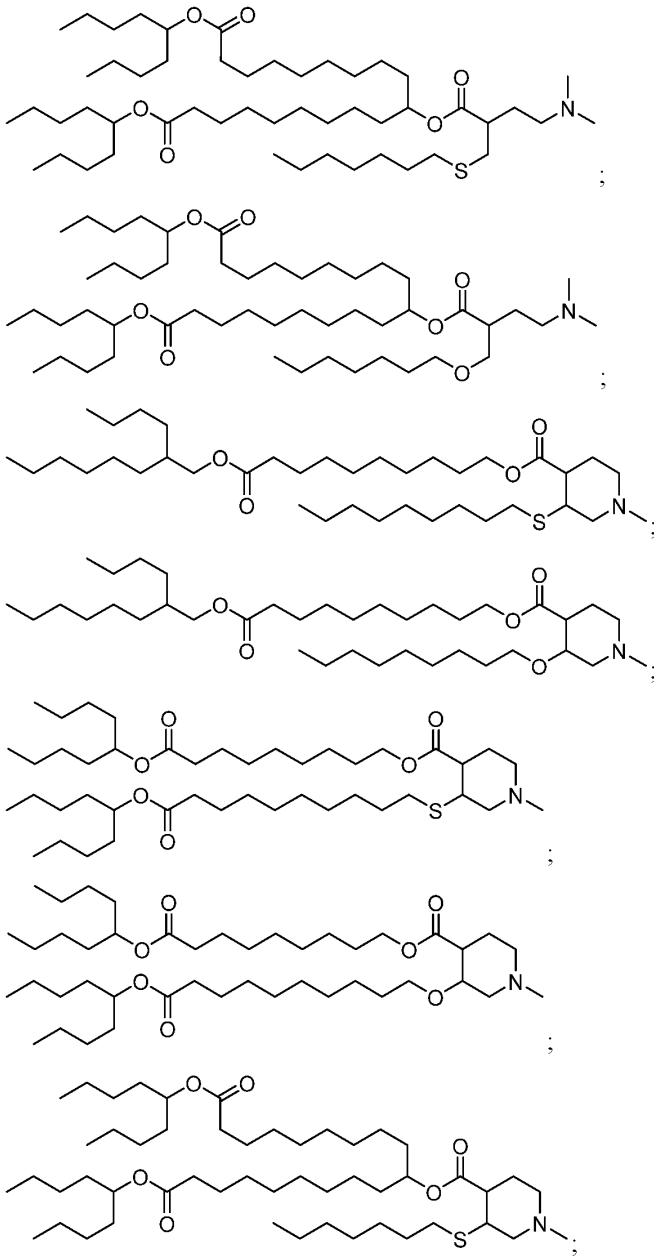


;

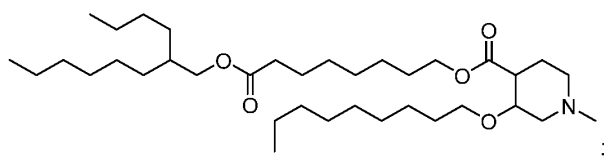
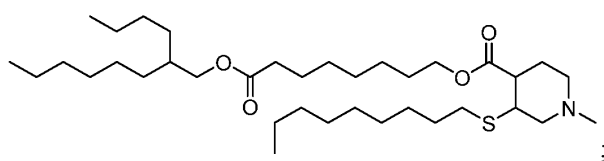
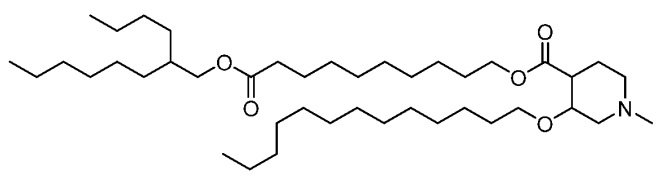
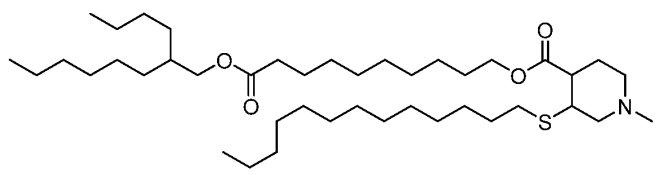
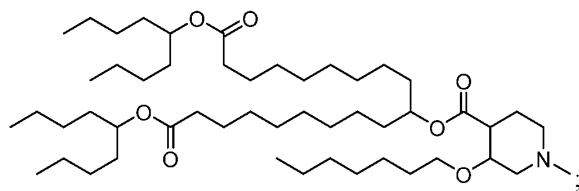


;

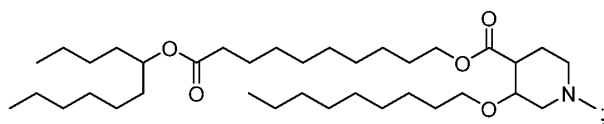
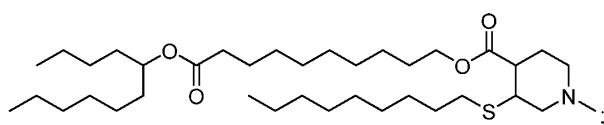




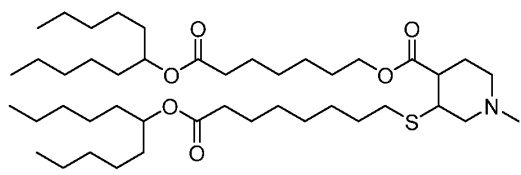
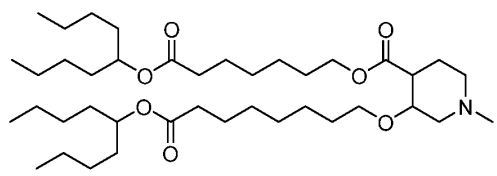
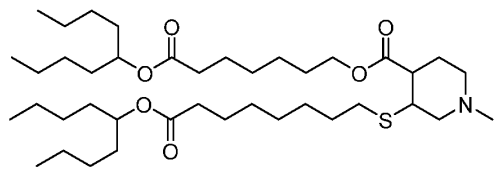
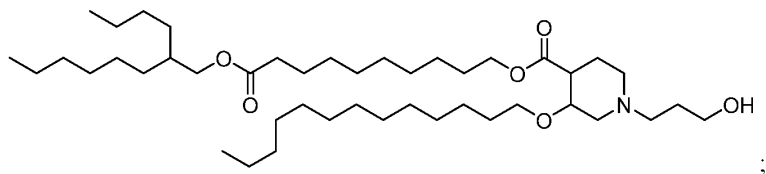
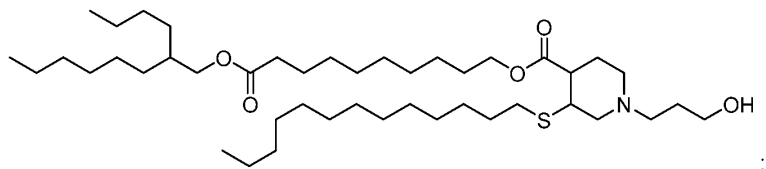
45



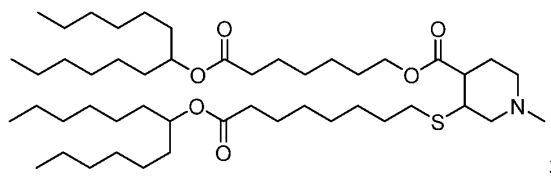
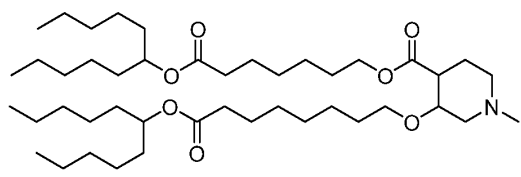
5



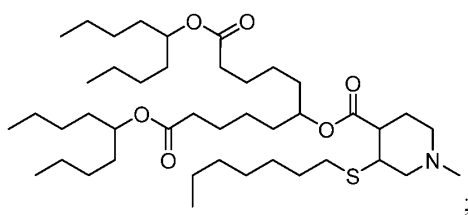
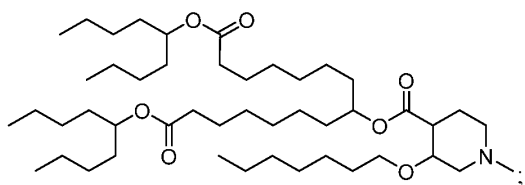
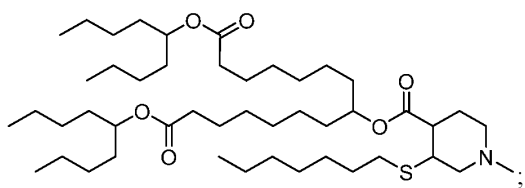
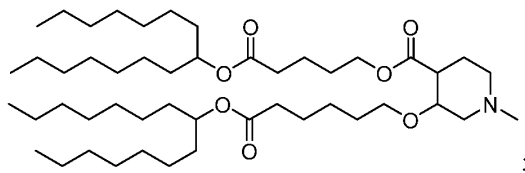
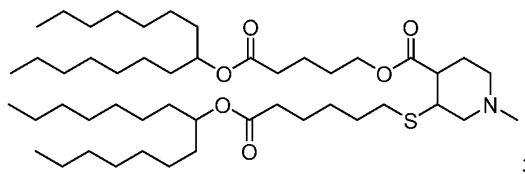
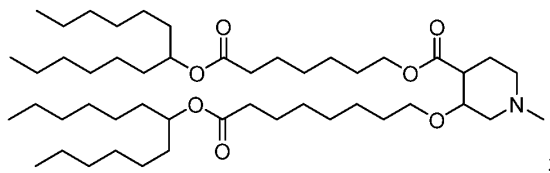
46



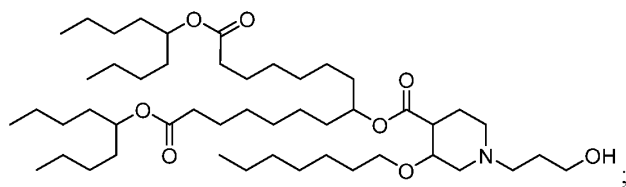
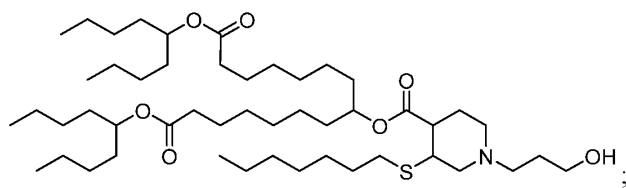
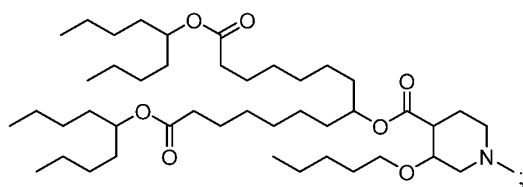
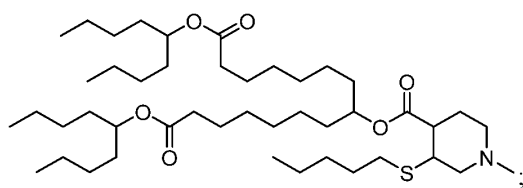
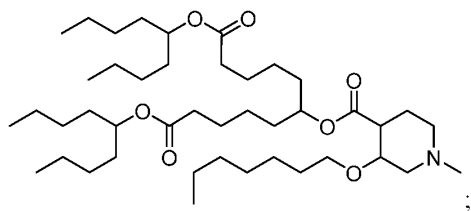
5



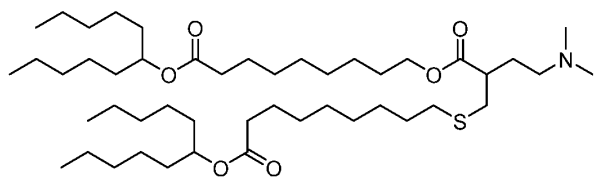
47



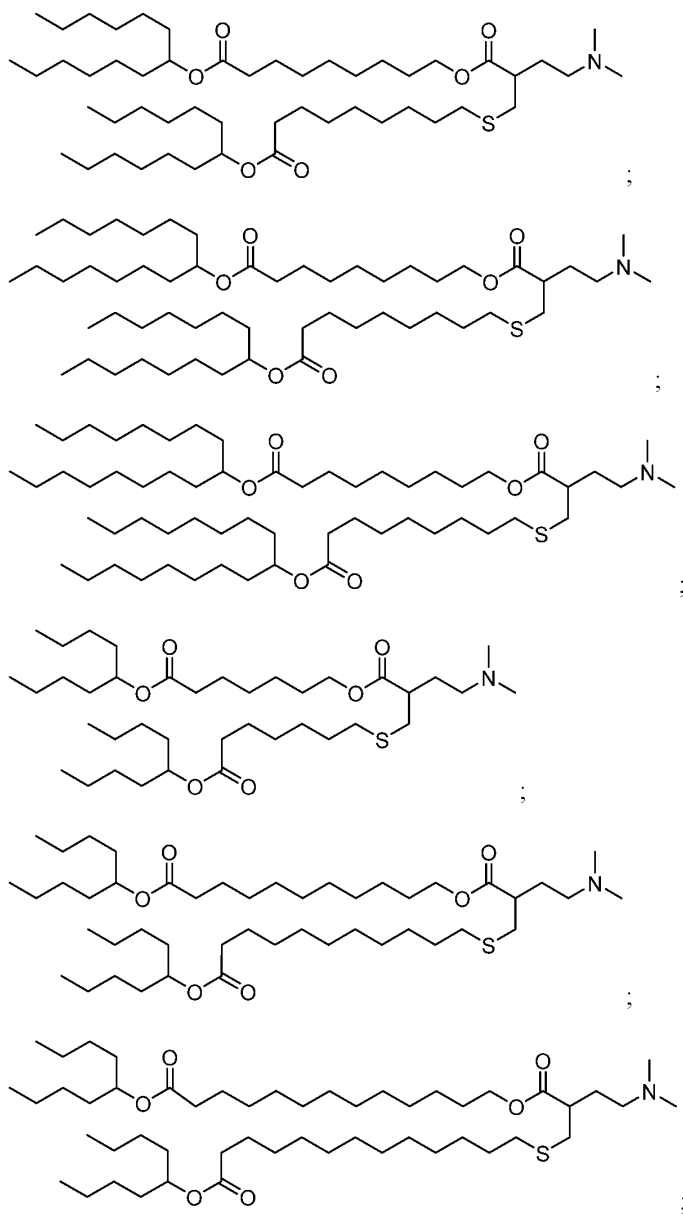
48



5

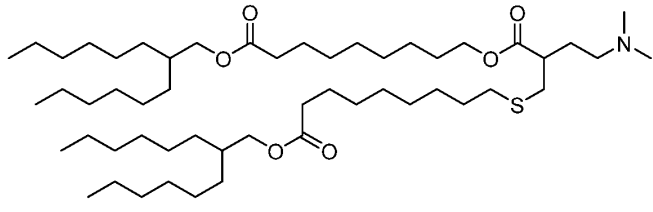
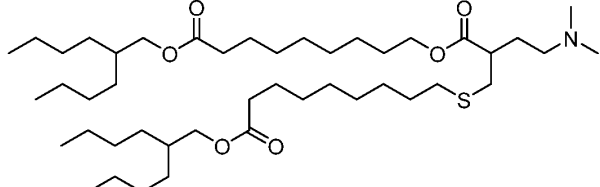
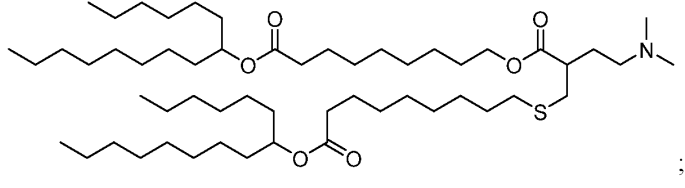
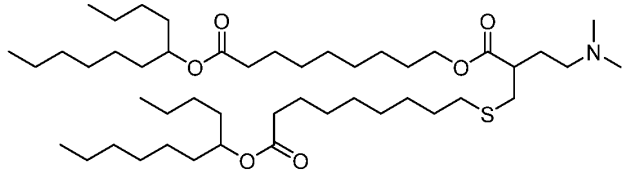
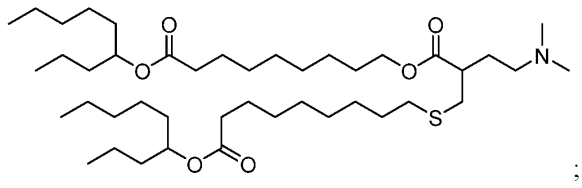
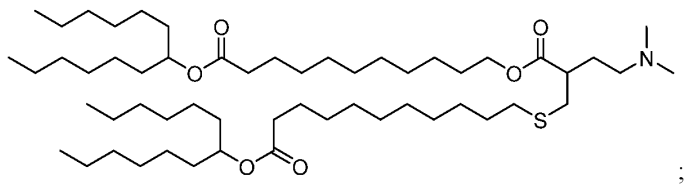


49



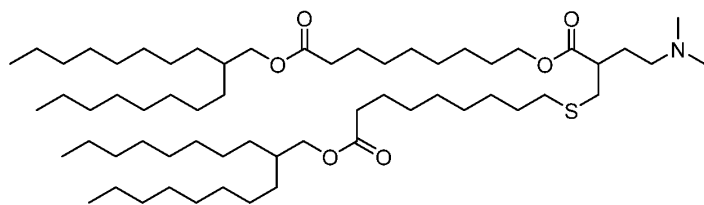
5

50

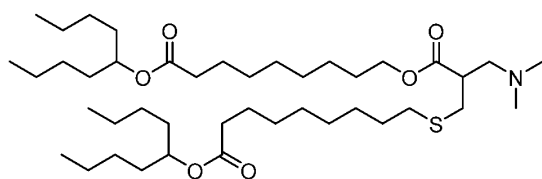


5

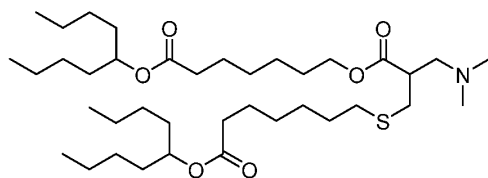
51



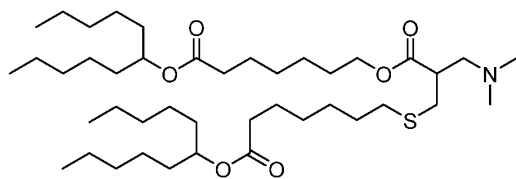
;



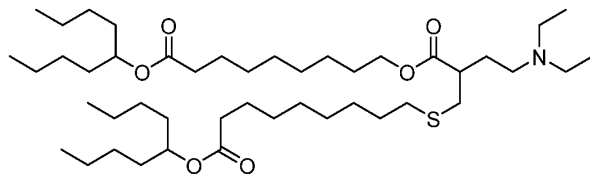
;



;

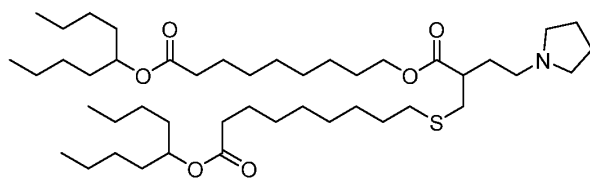


;



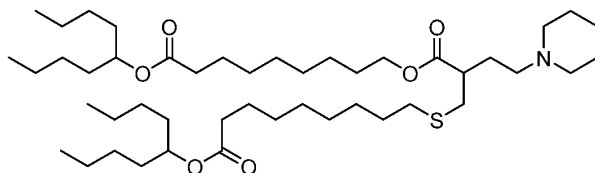
;

5



; and

52



Lipid Nanoparticles

[0121] The present disclosure provides for an LNP for delivery of a polynucleotide, such as an RNA, wherein the LNP comprises a compound of the present disclosure.

[0122] In embodiments, the LNPs have a mean diameter of from about 30 nm to about 160 nm, from about 40 nm to about 160 nm, from about 50 nm to about 160 nm, from about 60 nm to about 160 nm, from about 70 nm to about 160 nm, from about 50 nm to about 140 nm, from about 60 nm to about 130 nm, from about 70 nm to about 120 nm, from about 80 nm to about 120 nm, from about 90 nm to about 120 nm, from about 70 to about 110 nm, from about 80 nm to about 110 nm, or about 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, 150 nm, 155 nm or 160 nm. The diameter of the LNP may be measured by dynamic light scattering (DLS), transmission electron microscopy (TEM), scanning electron microscopy (SEM), or other methods such as are known in the art.

[0123] In some embodiments, the LNPs may be relatively homogenous. A polydispersity index may be used to indicate the homogeneity of the LNPs. A small, for example less than 0.3 or less than 0.2, polydispersity index generally indicates a narrow particle size distribution. A composition of the LNPs described herein 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 the LNP composition may be from about 0 to about 0.20 or 0.05 to 0.20.

[0124] The LNP may comprise more than one compound of Formula I or I-A to I-N or II-A to II-N, as appropriate. The inclusion of more than one such compound may, for example, be employed to achieve a desired pKa profile.

[0125] The LNP may comprise a compound of Formula I and an additional cationic and/or ionizable lipid, for example a cationic and/or ionizable lipid comprising a cyclic or non-cyclic amine. Such additional cationic and/or ionizable lipids may be selected from the non-limiting group consisting of:

- 3-(didodecylamino)-N1,N1,4-tridodecyl-1-piperazineethanamine (KL10),
- N1-[2-(didodecylamino)ethyl]-N1,N4,N4-tridodecyl-1,4-
- 10 piperazinediethanamine (KL22),
- 14,25-ditridecyl-15,18,21,24-tetraaza-octatriacontane (KL25),
- 1,2-dilinolexyloxy-N,N-dimethylaminopropane (DLin-DMA),
- 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA),
- (6Z,9Z,28Z,31Z)-heptatriacont-6,9,28,31-tetraene-19-yl 4-
- 15 (dimethylamino)butanoate (DLin-MC3-DMA),
- 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-KC2-
- DMA),
- 1,2-diolexyloxy-N,N-dimethylaminopropane (DODMA),
- 2-({8-[(3 β)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-[(9Z,12Z)-
- 20 octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA),
- (2R)-2-({8-[(3 β)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-
- [(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2R)),
- (2S)-2-({8-[(3 β)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-
- [(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2S)),
- 25 ((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate))
- and

8-[(2-hydroxyethyl)[6-oxo-6-(undecyloxy)hexyl]amino]-octanoic acid, 1-octylnonyl ester.

[0126] In embodiments, the LNP additionally comprises one or more of a PEG-lipid, a sterol structural lipid and/or a neutral lipid.

5 *PEGylated lipids*

[0127] In one embodiment, the present disclosure provides an LNP comprising a compound of the present disclosure and a PEGylated lipid.

[0128] It will be apparent to the skilled person that reference to a PEGylated lipid is a lipid that has been modified with polyethylene glycol. Exemplary PEGylated lipids include, but are not limited to, PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified dialkylamines, PEG-modified diacylglycerols, and PEG-modified dialkylglycerols. For embodiment, a PEG lipid includes PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, a PEG-DSPE lipid and combinations thereof.

15 *Neutral lipids*

[0129] In one embodiment, the present disclosure provides an LNP comprising a compound of the present disclosure and a neutral lipid.

[0130] Suitable neutral or zwitterionic lipids for use in the present disclosure will be apparent to the skilled person and include, in embodiments, 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-

phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), and sphingomyelin. The lipids can be saturated or unsaturated.

Structural lipids

[0131] In one embodiment, the present disclosure provides an LNP comprising a compound of the present disclosure and a structural lipid.

10 [0132] Exemplary structural lipids include, but are not limited to, cholesterol, fecosterol, sitosterol, campesterol, stigmasterol, brassicasterol, ergosterol, tomatidine, tomatine, ursolic acid and alpha-tocopherol.

[0133] In one embodiment, the structural lipid is a sterol. In embodiments, the structural lipid is cholesterol. In another embodiment, the structural lipid is campesterol.

15 [0134] In embodiments, the LNPs comprise an ionizable cationic lipid compound of the present disclosure; a neutral lipid; a sterol such as cholesterol; and a PEGylated lipid. The LNPs are formulated with a polynucleotide to be delivered to a subject.

Polynucleotides

20 [0135] The compounds of the present disclosure may form complexes with, and so be formulated into LNPs with, a range of polynucleotides including, but not limited to, mRNA, siRNA, antisense oligonucleotide, plasmid DNA, microRNA (miRNA), miRNA inhibitors (antagomirs/antimirs), messenger-RNA-interfering complementary RNA (micRNA), DNA, multivalent RNA, dicer substrate RNA, complementary DNA (cDNA), and the like. In this manner the LNPs and compositions may, in some
25 embodiments, be used to induce expression of a desired protein both *in vitro* and *in vivo* by contacting cells with an LNP comprising one or more novel compounds of the present disclosure, wherein the LNP encapsulates or is associated with a polynucleotide that is expressed to produce the desired protein, such as a mRNA or plasmid encoding the

desired protein. In alternative embodiments, the LNPs and compositions may be used to decrease the expression of target genes and proteins *in vitro* or *in vivo* by contacting cells with an LNP comprising one or more novel compounds of the present disclosure, wherein the LNP encapsulates or is associated with a polynucleotide that reduces target gene
5 expression, such as an antisense oligonucleotide or siRNA.

[0136] Therefore, in some embodiments, the polynucleotide is a mRNA encoding a polypeptide of interest, including any naturally or non-naturally occurring or otherwise modified polypeptide. A polypeptide encoded by an mRNA may be of any size and may have any secondary structure or activity. In some embodiments, a polypeptide encoded
10 by an mRNA may have a therapeutic effect when expressed in a cell.

[0137] In other embodiments, the polynucleotide is a siRNA capable of selectively knocking down or down regulating expression of a gene of interest. For example, an siRNA could be selected to silence a gene associated with a particular disease, disorder, or condition upon administration to a subject in need thereof of an LNP composition
15 comprising the siRNA. A siRNA may comprise a sequence that is complementary to an mRNA sequence that encodes a gene or protein of interest. In some embodiments, the siRNA may be an immunomodulatory siRNA.

[0138] In some embodiments, the polynucleotide is a shRNA or a vector or plasmid encoding the same. A shRNA may be produced inside a target cell upon delivery of an
20 appropriate construct to the nucleus. Constructs and mechanisms relating to shRNA are well-known in the relevant arts.

[0139] Polynucleotides useful for formulation with the LNPs incorporating an ionizable cationic compound of the present disclosure may include a first region of linked nucleosides encoding a polypeptide of interest (e.g., a coding region), a first flanking
25 region located at the 5'-terminus of the first region (e.g., a 5'-UTR), a second flanking region located at the 3'-terminus of the first region (e.g., a 3'-UTR), at least one 5'-cap region, and a 3'-stabilizing region. In some embodiments, a polynucleotide further includes a poly-A region or a Kozak sequence (e.g., in the 5'-UTR). In some cases, polynucleotides may contain one or more intronic sequences capable of being excised
30 from the polynucleotide. In some embodiments, a polynucleotide (e.g., an mRNA) may

include a 5' cap structure, a chain terminating nucleotide, a stem loop, a poly A sequence, and/or a polyadenylation signal. Any one of the regions of a polynucleotide may include one or more alternative components (e.g., an alternative nucleoside). For example, the 3'-stabilizing region may contain an alternative nucleoside such as an L-nucleoside, an
5 inverted thymidine, or a 2'-O-methyl nucleoside and/or the coding region, 5'-UTR, 3'-UTR, or cap region may include an alternative nucleoside such as a 5-substituted uridine (e.g., 5-methoxy uridine), a 1-substituted pseudouridine (e.g., 1-methyl-pseudouridine or 1-ethyl-pseudouridine), and/or a 5-substituted cytidine (e.g., 5-methyl-cytidine).

[0140] An exemplary polynucleotides useful for formulation with the LNPs
10 incorporating an ionizable cationic compound of the present disclosure include a first region of linked nucleosides encoding an antigenic polypeptide, a first flanking region located at the 5'-terminus of the first region (e.g., a 5'-UTR), a second flanking region located at the 3'-terminus of the first region (e.g., a 3'-UTR), at least one 5'-cap region, and a 3'-stabilizing region.

15 [0141] Polynucleotides suitable for use with the present LNPs may include one or more naturally occurring components, including any of the canonical nucleotides A (adenosine), G (guanosine), C (cytosine), U (uridine), or T (thymidine). In one embodiment, all or substantially all of the nucleotides comprising (a) the 5'-UTR, (b) the open reading frame (ORF), (c) the 3'-UTR, (d) the poly A tail, and any combination of
20 (a, b, c, or d above) comprise naturally occurring canonical nucleotides A (adenosine), G (guanosine), C (cytosine), U (uridine), or T (thymidine).

[0142] In some embodiments, polynucleotides may include one or more alternative components, as described herein, which impart useful properties including increased stability and/or the lack of a substantial induction of the innate immune response of a cell
25 into which the polynucleotide is introduced. For example, an alternative polynucleotide exhibits reduced degradation in a cell into which the polynucleotide is introduced, relative to a corresponding unaltered polynucleotide. These alternative species may enhance the efficiency of protein production, intracellular retention of the polynucleotides, and/or viability of contacted cells, as well as possess reduced
30 immunogenicity.

[0143] Polynucleotides may be naturally or non-naturally occurring. Polynucleotides may include one or more modified (e.g., altered or alternative) nucleobases, nucleosides, nucleotides, or combinations thereof. The polynucleotides may include any useful modification or alteration, such as to the nucleobase, the sugar,
5 or the internucleoside linkage (e.g., to a linking phosphate / to a phosphodiester linkage / to the phosphodiester backbone). In some embodiments, one or more alterations are present in each of the nucleobase, the sugar, and the internucleoside linkage.

[0144] Polynucleotides may or may not be uniformly altered along the entire length of the molecule. For example, one or more or all types of nucleotide (e.g., purine or
10 pyrimidine, or any one or more or all of A, G, U, C) may or may not be uniformly altered in a polynucleotide, or in a given predetermined sequence region thereof.

[0145] Different sugar alterations and/or internucleoside linkages (e.g., backbone structures) may exist at various positions in a polynucleotide. One of ordinary skill in the art will appreciate that the nucleotide analogs or other alteration(s) may be located at any
15 position(s) of a polynucleotide such that the function of the polynucleotide is not substantially decreased. An alteration may also be a 5'- or 3'-terminal alteration. In some embodiments, the polynucleotide includes an alteration at the 3'-terminus.

Nucleobase alternatives

[0146] The alternative nucleosides and nucleotides can include an alternative
20 nucleobase. A nucleobase of a polynucleotide is an organic base such as a purine or pyrimidine or a derivative thereof. A nucleobase may be a canonical base (e.g., adenine, guanine, uracil, thymine, and cytosine). These nucleobases can be altered or wholly replaced to provide polynucleotide molecules having enhanced properties, e.g., increased stability such as resistance to nucleases. Non-canonical or modified bases may include,
25 for example, one or more substitutions or modifications including but not limited to alkyl, aryl, halo, oxo, hydroxyl, alkyloxy, and/or thio substitutions; one or more fused or open rings; oxidation; and/or reduction.

[0147] Alternative nucleotide base pairing encompasses not only the standard adenine-thymine, adenine-uracil, or guanine-cytosine base pairs, but also base pairs
30 formed between nucleotides and/or alternative nucleotides including non-standard or

alternative bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures. One example of such non-standard base pairing is the base pairing between the alternative nucleotide inosine and adenine, cytosine, or uracil.

[0148] In some embodiments, the nucleobase is an alternative uracil. Exemplary nucleobases and nucleosides having an alternative uracil include pseudouridine (ψ), pyridin-4-one ribonucleoside, 5-aza-uracil, 6-aza-uracil, 2-thio-5-aza-uracil, 2-thio-uracil (s2U), 4-thio-uracil (s4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uracil (ho5U), 5-aminoallyl-uracil, 5-halo-uracil (e.g., 5-iodo-uracil or 5-bromo-uracil), 3-methyl-uracil (m3U), 5-methoxy-uracil (mo5U), uracil 5-oxyacetic acid (cmo5U), uracil 5-oxyacetic acid methyl ester (mcmo5U), 5-carboxymethyl-uracil (cm5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uracil (chm5U), 5-carboxyhydroxymethyl-uracil methyl ester (mchm5U), 5-methoxycarbonylmethyl-uracil (mcm5U), 5-methoxycarbonylmethyl-2-thio-uracil (mcm5s2U), 5-aminomethyl-2-thio-uracil (nm5s2U), 5-methylaminomethyl-uracil (mnm5U), 5-methylaminomethyl-2-thio-uracil (mnm5s2U), 5-methylaminomethyl-2-seleno-uracil (mnm5se2U), 5-carbamoylmethyl-uracil (ncm5U), 5-carboxymethylaminomethyl-uracil (cmnm5U), 5-carboxymethylaminomethyl-2-thio-uracil (cmnm5s2U), 5-propynyl-uracil, 1-propynyl-pseudouracil, 5-taurinomethyl-uracil (τ m5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uracil (τ m5s2U), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uracil (m5U, i.e., having the nucleobase deoxythymine), 1-methyl-pseudouridine (m ψ), 1-ethyl-pseudouridine (Et1 ψ), 5-methyl-2-thio-uracil (m5s2U), 1-methyl-4-thio-pseudouridine (m1s4 ψ), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine (m3 ψ), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouracil (D), dihydropseudouridine, 5,6-dihydrouracil, 5-methyl-dihydrouracil (m5D), 2-thio-dihydrouracil, 2-thio-dihydropseudouridine, 2-methoxy-uracil, 2-methoxy-4-thio-uracil, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, NI-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uracil (acp3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp3 ψ), 5-(isopentenylaminomethyl)uracil (inm5U), 5-(isopentenylaminomethyl)-2-thio-uracil

(inm5s2U), 5,2'-O-dimethyl-uridine (m5Um), 2-thio-2'-O-methyl-uridine (s2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine (cmnm5Um), 3,2'-O-dimethyl-uridine (m3Um), and 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm5Um), 1-thio-uracil, deoxythymidine, 5-(2-carbomethoxyvinyl)-uracil, 5-(carbamoylhydroxymethyl)-uracil, 5-carbamoylmethyl-2-thio-uracil, 5-carboxymethyl-2-thio-uracil, 5-cyanomethyl-uracil, 5-methoxy-2-thio-uracil, and 5-[3-(1-E-propenylamino)]uracil. In one example, the modified uracil is pseudouridine. In one example, the modified uracil is N1-methyl-pseudouridine.

- 10 [0149] In some embodiments, the nucleobase is an alternative cytosine. Exemplary nucleobases and nucleosides having an alternative cytosine include 5-aza-cytosine, 6-aza-cytosine, pseudoisocytidine, 3-methyl-cytosine (m3C), N4-acetyl-cytosine (ac4C), 5-formyl-cytosine (f5C), N4-methyl-cytosine (m4C), 5-methyl-cytosine (m5C), 5-halo-cytosine (e.g., 5-iodo-cytosine), 5-hydroxymethyl-cytosine (hm5C), 1-methyl-pseudoisocytidine, pyrrolo-cytosine, pyrrolo-pseudoisocytidine, 2-thio-cytosine (s2C), 2-thio-5-methyl-cytosine, 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-cytosine, 2-methoxy-5-methyl-cytosine, 4-methoxy-pseudoisocytidine, 4-methoxy-1-methyl-pseudoisocytidine, lysidine (k2C), 5,2'-O-dimethyl-cytidine (m5Cm), N4-acetyl-2'-O-methyl-cytidine (ac4Cm), N4,2'-O-dimethyl-cytidine (m4Cm), 5-formyl-2'-O-methyl-cytidine (f5Cm), N4,N4,2'-O-trimethyl-cytidine (m42Cm), 1-thio-cytosine, 5-hydroxy-cytosine, 5-(3-azidopropyl)-cytosine, and 5-(2-azidoethyl)-cytosine. In one example, the modified cytosine is 5-methyl-cytosine.

- 30 [0150] In some embodiments, the nucleobase is an alternative adenine. Exemplary nucleobases and nucleosides having an alternative adenine include 2-amino-purine, 2,6-diaminopurine, 2-amino-6-halo-purine (e.g., 2-amino-6-chloro-purine), 6-halo-purine (e.g., 6-chloro-purine), 2-amino-6-methyl-purine, 8-azido-adenine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-amino-purine, 7-deaza-8-aza-2-amino-purine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyl-adenine (m1A), 2-

methyl-adenine (m2A), N6-methyl-adenine (m6A), 2-methylthio-N6-methyl-adenine (ms2m6A), N6-isopentenyl-adenine (i6A), 2-methylthio-N6-isopentenyl-adenine (ms2i6A), N6-(cis-hydroxyisopentenyl)adenine (io6A), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenine (ms2io6A), N6-glycinylicarbamoyl-adenine (g6A), N6-threonylicarbamoyl-adenine (t6A), N6-methyl-N6-threonylicarbamoyl-adenine (m6t6A),
 5 2-methylthio-N6-threonylicarbamoyl-adenine (ms2g6A), N6,N6-dimethyl-adenine (m62A), N6-hydroxynorvalylcarbamoyl-adenine (hn6A), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenine (ms2hn6A), N6-acetyl-adenine (ac6A), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, N6,2'-O-dimethyl-adenosine
 10 (m6Am), N6,N6,2'-O-trimethyl-adenosine (m62Am), 1,2'-O-dimethyl-adenosine (m1Am), 2-amino-N6-methyl-purine, 1-thio-adenine, 8-azido-adenine, N6-(19-amino-pentaoxanadecyl)-adenine, 2,8-dimethyl-adenine, N6-formyl-adenine, and N6-hydroxymethyl-adenine.

[0151] In some embodiments, the nucleobase is an alternative guanine. Exemplary
 15 nucleobases and nucleosides having an alternative guanine include inosine (I), 1-methyl-inosine (mI), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine (o2yW), hydroxywybutosine (OHyW), undermodified hydroxywybutosine (OHyW*), 7-deaza-guanine, queuosine (Q), epoxyqueuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ),
 20 7-cyano-7-deaza-guanine (preQ0), 7-aminomethyl-7-deaza-guanine (preQ1), archaeosine (G+), 7-deaza-8-aza-guanine, 6-thio-guanine, 6-thio-7-deaza-guanine, 6-thio-7-deaza-8-aza-guanine, 7-methyl-guanine (m7G), 6-thio-7-methyl-guanine, 7-methyl-inosine, 6-methoxy-guanine, 1-methyl-guanine (m1G), N2-methyl-guanine (m2G), N2,N2-dimethyl-guanine (m22G), N2,7-dimethyl-guanine (m2,7G), N2, N2,7-
 25 dimethyl-guanine (m2,2,7G), 8-oxo-guanine, 7-methyl-8-oxo-guanine, 1-methyl-6-thio-guanine, N2-methyl-6-thio-guanine, N2,N2-dimethyl-6-thio-guanine, N2-methyl-2'-O-methyl-guanosine (m2Gm), N2,N2-dimethyl-2'-O-methyl-guanosine (m22Gm), 1-methyl-2'-O-methyl-guanosine (m1Gm), N2,7-dimethyl-2'-O-methyl-guanosine (m2,7Gm), 2'-O-methyl-inosine (Im), 1,2'-O-dimethyl-inosine (mllm), 1-thio-guanine,
 30 and O-6-methyl-guanine.

[0152] The alternative nucleobase of a nucleotide can be independently a purine, a pyrimidine, a purine or pyrimidine analog. For example, the nucleobase can be an alternative to adenine, cytosine, guanine, uracil, or hypoxanthine. In another embodiment, the nucleobase can also include, for example, naturally-occurring and synthetic derivatives of a base, including pyrazolo[3,4-d]pyrimidines, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo (e.g., 8-bromo), 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxy and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, deazaguanine, 7-deazaguanine, 3-deazaguanine, deazaadenine, 7-deazaadenine, 3-deazaadenine, pyrazolo[3,4-d]pyrimidine, imidazo[1,5-a]1,3,5 triazinones, 9-deazapurines, imidazo[4,5-d]pyrazines, thiazolo[4,5-d]pyrimidines, pyrazin-2-ones, 1,2,4-triazine, pyridazine; or 1,3,5 triazine.

[0153] Polynucleotides for formulation with LNPs comprising a compound of the present disclosure may be prepared according to any available technique known in the art. mRNA may be prepared by, for example, enzymatic synthesis which provides a process of template-directed synthesis of RNA molecules from an engineered DNA template comprised of an upstream bacteriophage promoter sequence linked to a downstream sequence encoding the gene of interest. Template DNA can be prepared for in vitro transcription from a number of sources with appropriate techniques which are well-known in the art including, but not limited to, plasmid DNA and polymerase chain reaction amplification (see Linpinsel, J.L and Conn, G.L., General protocols for preparation of plasmid DNA template and Bowman, J.C., Azizi, B., Lenz, T.K., Ray, P., and Williams, L.D. in RNA in vitro transcription and RNA purification by denaturing PAGE in Recombinant and in vitro RNA syntheses Methods v. 941 Conn G.L. (ed), New York, N.Y. Humana Press, 2012)

[0154] Transcription of the RNA occurs *in vitro* using the linearized DNA template in the presence of the corresponding RNA polymerase and adenosine, guanosine, uridine

and cytidine ribonucleoside triphosphates (rNTPs) under conditions that support polymerase activity while minimizing potential degradation of the resultant mRNA transcripts. *In vitro* transcription can be performed using a variety of commercially available kits including, but not limited to RiboMax Large Scale RNA Production System (Promega), MegaScript Transcription kits (Life Technologies) as well as with commercially available reagents including RNA polymerases and rNTPs. The methodology for *in vitro* transcription of mRNA is well-known in the art. (see, e.g. Losick, R., 1972, *In vitro* transcription, *Ann Rev Biochem* v.41 409-46; Kamakaka, R. T. and Kraus, W. L. 2001. *In Vitro Transcription*. *Current Protocols in Cell Biology*. 2:11.6:11.6.1-11.6.17; Beckert, B. And Masquida, B., (2010) *Synthesis of RNA by In Vitro Transcription in RNA in Methods in Molecular Biology* v. 703 (Neilson, H. Ed), New York, N.Y. Humana Press, 2010; Brunelle, J.L. and Green, R., 2013, Chapter Five - *In vitro* transcription from plasmid or PCR-amplified DNA, *Methods in Enzymology* v. 530, 101-114; all of which are incorporated herein by reference).

[0155] The desired *in vitro* transcribed mRNA is then purified from the undesired components of the transcription or associated reactions. Techniques for the isolation of the mRNA transcripts are well known in the art and include phenol/chloroform extraction or precipitation with either alcohol in the presence of monovalent cations or lithium chloride.

20

Lipid Nanoparticle Formation

[0156] LNPs comprising a compound of the present disclosure can be made using approaches which are well-known in the art of formulation. For example, suitable LNPs can be formed using mixing processes such as microfluidics, including herringbone micromixing, and T-junction mixing of two fluid streams, one of which contains the polynucleotide, typically in an aqueous solution, and the other of which has the various required lipid components, typically in ethanol.

[0157] The LNPs may then be prepared by combining a compound of Formula I, a phospholipid (such as DOPE or DSPC, which may be purchased from commercial sources including Avanti Polar Lipids, Alabaster, AL), a PEGylated lipid (such as 1,2-

30

dimyristoyl-sn-glycerol methoxypoly ethylene glycol, also known as PEG-DMG, which may be purchased from commercial sources including Avanti Polar Lipids, Alabaster, AL), and a structural lipid / sterol (such as cholesterol, which may be purchased from commercial sources including Sigma- Aldrich), at concentrations of, for example, about
5 50 mM in ethanol. Solutions should be refrigerated during storage at, for example, -20° C. The various lipids may be combined to yield the desired molar ratios and diluted with water and ethanol to a final desired lipid concentration of, for example, between about 5.5 mM and about 25 mM.

[0158] An LNP composition comprising a polynucleotide is prepared (as set out in
10 the examples) by combining the above lipid solution with a solution including the polynucleotide at, for example, a lipid component to polynucleotide wt:wt ratio from about 5:1 to about 50:1. The lipid solution may be rapidly injected using a NanoAssemblr microfluidic system at flow rates between about 3 ml/min and about 18 ml/min into the polynucleotide solution to produce a suspension with a water to ethanol ratio between
15 about 1:1 and about 4:1, or between about 2:1 and about 4:1.

[0159] For LNP compositions including a RNA, solutions of the RNA at concentrations of 1.0 mg/ml in deionized water may be diluted in 50 mM sodium citrate buffer at a pH between 3 and 6 to form a stock solution.

[0160] LNP compositions may be further processed, as is known in the art, in one
20 example by 10-fold dilution into 50 mM citrate buffer at pH 6 and subjected to tangential flow filtration (TFF) using a 300k molecular weight cut-off membrane (mPES) until concentrated to the original volume. Subsequently, the citrate buffer may be replaced with a buffer containing 20 mM Tris buffer at pH 7.5, 80 mM sodium chloride, and 3% sucrose using diafiltration with a 10-fold volume of the new buffer. The LNP solution
25 may be concentrated to a volume of between 5-10 mL, filtered using a 0.2 micron PES syringe filter, aliquoted into vials, and frozen at 1°C/min using a Corning® CoolCell® LX Cell Freezing Container until the samples reach -80°C. Samples may be stored at -80°C until required.

[0161] The method described above induces nano-precipitation and particle formation. Alternative processes including, but not limited to, T-junction and direct injection, may be used to achieve the same nano-precipitation.

[0162] In some embodiments, the lipid component of the LNP formulation comprises about 25 mol % to about 60 mol % compound of Formula I, about 2 mol % to about 25 mol % phospholipid (neutral lipid), about 18.5 mol % to about 60 mol % structural lipid (sterol), and about 0.2 mol % to about 10 mol % of PEGylated lipid, provided that the total mol % does not exceed 100%. In some embodiments, the lipid component of the LNP formulation comprises about 30 mol % to about 50 mol % compound of Formula I, about 5 mol % to about 20 mol % phospholipid, about 30 mol % to about 55 mol % structural lipid, and about 1 mol % to about 5 mol % of PEGylated lipid. In a particular embodiment, the lipid component includes about 40 mol % compound of the present disclosure, about 10 mol % phospholipid, about 48 mol % structural lipid, and about 2.0 mol % of PEG lipid. In some embodiments, the phospholipid may be DOPE or DSPC. In other embodiments, the PEG lipid may be PEG-DMG and/or the structural lipid may be cholesterol.

[0163] The efficiency of encapsulation of the polynucleotide within the LNPs comprising a compound of the present disclosure may be at least 50%, for example about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. In some embodiments, the encapsulation efficiency may be at least 80%. In certain embodiments, the encapsulation efficiency may be at least 90%.

Lipid Nanoparticle Compositions

[0164] The LNPs comprising a compound of the present disclosure and a polynucleotide can be formulated for administration via any accepted mode of administration of lipid particles including LNPs, liposomes, lipid vesicles and like lipid-based particles. The pharmaceutical compositions of the invention may be formulated into preparations in solid, semi-solid, liquid or gaseous forms, such as tablets, capsules, powders, granules, ointments, solutions, suspensions, suppositories, injections, inhalants, gels, microspheres, and aerosols. Typical routes of administering such pharmaceutical

LNP compositions include, without limitation, oral, topical, transdermal, inhalation, parenteral, sublingual, buccal, rectal, vaginal, and intranasal. The term parenteral as used herein includes subcutaneous injections, intravenous, intramuscular, intradermal, intrasternal injection or infusion techniques. The compositions administered to a subject
5 may be in the form of one or more dosage units, where for example, a tablet or injectable liquid volume may be a single dosage unit. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in this art; for example, see Remington: The Science and Practice of Pharmacy, 20th Edition (Philadelphia College of Pharmacy and Science, 2000).

- 10 [0165] Therefore, one embodiment of the present disclosure provides a composition (such as a pharmaceutical composition) comprising an LNP, which comprises a compound of the present disclosure, combined with a pharmaceutically acceptable carrier.

- [0166] In general terms, by “carrier” is meant a solid or liquid filler, binder, diluent,
15 encapsulating substance, emulsifier, wetting agent, solvent, suspending agent, coating or lubricant that may be safely administered to any subject, e.g., a human. Depending upon the particular route of administration, a variety of acceptable carriers, known in the art may be used, as for embodiment described in Remington's Pharmaceutical Sciences (Mack Publishing Co. N.J. USA, 1991).

- 20 [0167] An LNP is useful for parenteral, topical, oral, or local administration, intramuscular administration, aerosol administration, or transdermal administration, for prophylactic or for therapeutic treatment. In one embodiment, the LNP is administered parenterally, such as intramuscularly, subcutaneously or intravenously. In some embodiments, the LNP is administered intramuscularly.

- 25 [0168] Formulation of LNPs to be administered will vary according to the route of administration and formulation (e.g., solution, emulsion, capsule) selected. An appropriate pharmaceutical composition comprising an LNP to be administered can be prepared in a physiologically acceptable carrier. For solutions or emulsions, suitable carriers include, for embodiment, aqueous or alcoholic/aqueous solutions, emulsions or
30 suspensions, including saline and buffered media. Parenteral vehicles can include

sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. A variety of appropriate aqueous carriers are known to the skilled artisan, including water, buffered water, buffered saline, polyols (e.g., glycerol, propylene glycol, liquid polyethylene glycol), dextrose solution and glycine. Intravenous
5 vehicles can include various additives, preservatives, or fluid, nutrient or electrolyte replenishers (See, generally, Remington's Pharmaceutical Science, 16th Edition, Mack, Ed. 1980). The compositions can optionally contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents and toxicity adjusting agents, for embodiment, sodium
10 acetate, sodium chloride, potassium chloride, calcium chloride and sodium lactate. The LNPs can be stored in the liquid stage or can be lyophilized for storage and reconstituted in a suitable carrier prior to use according to art-known lyophilization and reconstitution techniques.

[0169] When the LNP composition is a vaccine composition then the carrier may
15 be water, typically pyrogen-free water; isotonic saline or buffered (aqueous) solutions, e.g. phosphate, citrate etc. buffered solutions. For injection of an LNP vaccine composition, water or preferably a buffer, more preferably an aqueous buffer, may be used, containing a sodium salt, preferably at least 50mM of a sodium salt, a calcium salt, preferably at least 0.01mM of a calcium salt, and optionally a potassium salt, such as at
20 least 3mM of a potassium salt. In an embodiment, the sodium, calcium and, optionally, potassium salts may be present as their chlorides, iodides, or bromides, or in the form of their hydroxides, carbonates, hydrogen carbonates, or sulfates, etc. Non-limiting examples of sodium salts include e.g. NaCl, NaI, NaBr, Na₂CO₃, NaHCO₃, Na₂SO₄, examples of the optional potassium salts include e.g. KCl, KI, KBr, K₂CO₃, KHCO₃,
25 K₂SO₄, and examples of calcium salts include e.g. CaCl₂, CaI₂, CaBr₂, CaCO₃, CaSO₄, Ca(OH)₂. Furthermore, organic anions of the aforementioned cations may be contained in the buffer. In certain embodiments, the buffer suitable for injection purposes, may contain salts selected from sodium chloride (NaCl), calcium chloride (CaCl₂) and optionally potassium chloride (KCl), wherein further anions may be present additional
30 to the chlorides. In embodiments, the salts in the injection buffer are present in a concentration of at least 50mM sodium chloride (NaCl), at least 3mM potassium chloride

(KCl) and at least 0.01mM calcium chloride (CaCl₂). The injection buffer may be hypertonic, isotonic or hypotonic with reference to the specific reference medium.

[0170] In some embodiments of a vaccine, one or more compatible solid or liquid fillers or diluents or encapsulating compounds may be employed which are suitable for administration to a person. Pharmaceutically acceptable carriers, fillers and diluents will have sufficiently high purity and sufficiently low toxicity to make them suitable for administration to a subject. Some examples of compounds which can be used as pharmaceutically acceptable carriers, fillers or constituents thereof are sugars, such as, for example, lactose, glucose, trehalose and sucrose; starches, such as, for example, corn starch or potato starch; dextrose; cellulose and its derivatives, such as, for example, sodium carboxymethylcellulose, ethylcellulose, cellulose acetate; powdered tragacanth; malt; gelatin; tallow; solid glidants, such as, for example, stearic acid, magnesium stearate; calcium sulfate; vegetable oils, such as, for example, groundnut oil, cottonseed oil, sesame oil, olive oil, corn oil and oil from theobroma; polyols, such as, for example, polypropylene glycol, glycerol, sorbitol, mannitol and polyethylene glycol; and alginic acid.

[0171] When the LNP composition is a vaccine composition it may further comprise one or more pharmaceutically acceptable adjuvants to enhance the immunostimulatory properties of the composition. The adjuvant may be any compound, which is suitable to support administration and delivery of the LNP composition and which may initiate or increase an immune response of the innate immune system, i.e. a non-specific immune response.

[0172] Such an adjuvant may be selected from any adjuvant known to a skilled person and suitable for the particular nature of the vaccine, i.e. for induction of a suitable immune response in a mammal. In embodiments, the adjuvant may be selected from the group consisting of: MF59[®] (squalene-water emulsion), TDM, MDP, muramyl dipeptide, pluronics, alum solution, aluminium hydroxide, ADJUMER[™] (polyphosphazene); aluminium phosphate gel; glucans from algae; algammulin; aluminium hydroxide gel (alum); highly protein-adsorbing aluminium hydroxide gel; low viscosity aluminium hydroxide gel; AF or SPT (emulsion of squalane (5%), Tween 80 (0.2%), Pluronic L121 (1.25%), phosphate-buffered saline, pH 7.4); AVRIDINE[™]

(propanediamine); BAY R1005TM ((N-(2-deoxy-2-L-leucylamino-b-D-glucopyranosyl)-N-octadecyl-dodecanoyl-amide hydroacetate); CALCITRIOLTM (1- α ,25-dihydroxy-vitamin D3); calcium phosphate gel; CAPTM (calcium phosphate nanoparticles); cholera holotoxin, cholera-toxin-A1-protein-A-D-fragment fusion protein, sub-unit B of the

5 cholera toxin; CRL 1005 (block copolymer P1205); cytokine-containing liposomes; DDA (dimethyldioctadecylammonium bromide); DHEA (dehydroepiandrosterone); DMPC (dimyristoylphosphatidylcholine); DMPG (dimyristoylphosphatidylglycerol); DOC/alum complex (deoxycholic acid sodium salt); Freund's complete adjuvant; Freund's incomplete adjuvant; gamma inulin; Gerbu adjuvant (mixture of: i) N-

10 acetylglucosaminyl-(1-4)-N-acetylmuramyl-L-alanyl-D-glutamine (GMDP), ii) dimethyldioctadecylammonium chloride (DDA), iii) zinc-L-proline salt complex (ZnPro-8); GM-CSF; GMDP (N-acetylglucosaminyl-(1-4)-N-acetylmuramyl-L-alanyl-D-isoglutamine); imiquimod (1-(2-methylpropyl)-1H-imidazo[4,5-c]quinoline-4-amine); ImmTherTM (N-acetylglucosaminyl-N-acetylmuramyl-L-Ala-D-isoGlu-L-Ala-

15 glycerol dipalmitate); DRVs (immunoliposomes prepared from dehydration-rehydration vesicles); interferon-gamma; interleukin-1 β ; interleukin-2; interleukin-7; interleukin-12; ISCOMSTM; ISCOPREP 7.0.3.TM; liposomes; LOXORIBINETM (7-allyl-8-oxoguanosine); LT oral adjuvant (E.coli labile enterotoxin-prototoxin); microspheres and microparticles of any composition;; MONTANIDE ISA 51TM (purified incomplete

20 Freund's adjuvant); MONTANIDE ISA 720TM (metabolisable oil adjuvant); MPLTM (3-Q-desacyl-4'-monophosphoryl lipid A); MTP-PE and MTP-PE liposomes ((N-acetyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1,2-dipalmitoyl-sn-glycero-3-(hydroxyphosphoryloxy))-ethylamide, monosodium salt); MURAMETIDETM (Nac-Mur-L-Ala-D-Gln-OCH₃); MURAPALMITINETM and D-MURAPALMITINETM (Nac-

25 Mur-L-Thr-D-isoGln-sn-glyceroldipalmitoyl); NAGO (neuraminidase-galactose oxidase); nanospheres or nanoparticles of any composition; NISVs (non-ionic surfactant vesicles); PLEURANTM (β -glucan); PLGA, PGA and PLA (homo- and co-polymers of lactic acid and glycolic acid; microspheres/nanospheres); PLURONIC L121TM; PMMA (polymethyl methacrylate); PODDSTM (proteinoid microspheres); polyethylene carbamate derivatives; poly-rA: poly-rU (polyadenylic acid-polyuridylic acid complex);

30 polysorbate 80 (Tween 80); protein cochleates (Avanti Polar Lipids, Inc., Alabaster, AL);

- STIMULON™ (QS-21); Quil-A (Quil-A saponin); S-28463 (4-amino-otec-dimethyl-2-ethoxymethyl-1H-imidazo[4,5 c]quinoline-1-ethanol); SAF-1™ ("Syntex adjuvant formulation"); Sendai proteoliposomes and Sendai-containing lipid matrices; Span-85 (sorbitan trioleate); Specol (emulsion of Marcol 52, Span 85 and Tween 85); squalene or
- 5 Robane® (2,6,10,15,19,23-hexamethyltetracosan and 2,6,10,15,19,23-hexamethyl-2,6,10,14,18,22-tetracosahexane); stearyltyrosine (octadecyltyrosine hydrochloride); Theramid® (N-acetylglucosaminyl-N-acetylmuramyl-L-Ala-D-isoGlu-L-Ala-dipalmitoxypropylamide); Theronyl-MDP (Termurtide™ or [thrl]-MDP; N-acetylmuramyl-L-threonyl-D-isoglutamine); Ty particles (Ty-VLPs or virus-like
- 10 particles); Walter-Reed liposomes (liposomes containing lipid A adsorbed on aluminium hydroxide), and lipopeptides, including Pam3Cys, in particular aluminium salts, such as Adju-phos, Alhydrogel, Rehydragel; emulsions, including CFA, SAF, IFA, MF59, Provax, TiterMax, Montanide, Vaxfectin; copolymers, including Optivax (CRL1005), L121, Poloxamer4010), etc.; liposomes, including Stealth, cochleates, including
- 15 BIORAL; plant derived adjuvants, including QS21, Quil A, Iscomatrix, ISCOM; adjuvants suitable for costimulation including Tomatine, biopolymers, including PLG, PMM, Inulin; microbe derived adjuvants, including Romurtide, DETOX, MPL, CWS, Mannose, CpG polynucleotide sequences, CpG7909, ligands of human TLR 1-10, ligands of murine TLR 1-13, ISS-1018, IC31, Imidazoquinolines, Ampligen, Ribi529,
- 20 IMOxine, IRIVs, VLPs, cholera toxin, heat-labile toxin, Pam3Cys, Flagellin, GPI anchor, LNFPIII/Lewis X, antimicrobial peptides, UC-1V150, RSV fusion protein, cdiGMP; and adjuvants suitable as antagonists including CGRP neuropeptide. In one preferred example the adjuvant may be the oil-in-water emulsion adjuvant MF59®, particularly if the vaccine is an influenza vaccine.
- 25 [0173] Upon formulation, compositions of the present disclosure will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically/prophylactically effective. The dosage ranges for the administration of the LNPs are those large enough to produce the desired effect. For embodiment, the composition comprises an effective amount of the encapsulated or associated RNA, e.g.,
- 30 the mRNA or self-replicating RNA. In one embodiment, the composition comprises a

therapeutically effective amount of the RNA. In another embodiment, the composition comprises a prophylactically effective amount of the RNA.

[0174] The dosage should not be so large as to cause adverse side effects. Generally, the dosage will vary with the age, condition, sex and extent of the disease in the patient and can be determined by one of skill in the art. The dosage can be adjusted by the individual physician in the event of any complication.

[0175] Preparation methods for the above compounds and compositions are described further herein and/or are known in the art.

10 Methods of Treatment and Producing a Polypeptide of Interest

[0176] Diseases, disorders, and/or conditions which may be a result of or related to aberrant protein or polypeptide may be treated by the present LNPs comprising a compound of the present disclosure and a polynucleotide and may include, but are not limited to, rare diseases, infectious diseases, cancer and proliferative diseases, genetic diseases, autoimmune diseases, diabetes, neurodegenerative diseases, cardio- and reno-vascular diseases, and metabolic diseases.

[0177] LNP compositions may be formulated in unit dosage form. The therapeutically effective or prophylactically effective dose for any particular patient will depend upon a variety of factors including the severity and identify of a disorder being treated; the specific composition employed; the age, body weight, general health, sex, and diet of the patient; the time of administration, route of administration, and rate of excretion of the specific pharmaceutical composition employed; the duration of the treatment; drugs used in combination or coincidental with the specific pharmaceutical composition employed; and like factors well known in the medical arts.

[0178] LNP compositions described herein may be used in combination with one or more other therapeutic, prophylactic, diagnostic, or imaging agents. They may be administered together in a single composition or administered separately in different compositions.

[0179] The present disclosure provides methods of producing a polypeptide of interest in a mammalian cell. Methods of producing polypeptides involve contacting a cell with an LNP composition, as described herein, including an mRNA encoding the polypeptide of interest. Upon contacting the cell with the LNP composition, the mRNA
5 may be taken up and translated in the cell to produce the polypeptide of interest.

[0180] The step of contacting an LNP composition including an mRNA with a cell may involve or cause transfection. A phospholipid including in the lipid component of the LNP composition may facilitate transfection and/or increase transfection efficiency, for example, by interacting and/or fusing with a cellular or intracellular membrane.
10 Transfection may allow for the translation of the mRNA within the cell.

[0181] In some embodiments, the LNP compositions described herein may be used therapeutically. For example, an mRNA included in the LNP composition may encode a therapeutic polypeptide (e.g., in a translatable region) and produce the therapeutic polypeptide upon contacting and/or entry (e.g., transfection) into a cell. In other
15 embodiments, an mRNA included in the LNP composition may encode a polypeptide that may improve or increase the immunity of a subject.

[0182] In embodiments, an mRNA included in an LNP composition may encode a recombinant polypeptide that may replace one or more polypeptides that may be substantially absent in a cell contacted with the LNP composition. The one or more
20 substantially absent polypeptides may be lacking due to a genetic mutation of the encoding gene or a regulatory pathway thereof. Alternatively, a recombinant polypeptide produced by translation of the mRNA may antagonize the activity of an endogenous protein present in, on the surface of, or secreted from the cell. An antagonistic recombinant polypeptide may be desirable to combat deleterious effects caused by
25 activities of the endogenous protein, such as altered activities or localization caused by mutation. In another alternative, a recombinant polypeptide produced by translation of the mRNA may indirectly or directly antagonize the activity of a biological moiety present in, on the surface of, or secreted from the cell. Antagonized biological moieties may include, but are not limited to, lipids (e.g., cholesterol), lipoproteins (e.g., low
30 density lipoprotein), nucleic acids, carbohydrates, and small molecule toxins. Recombinant polypeptides produced by translation of the mRNA may be engineered for

localization within the cell, such as within a specific compartment such as the nucleus, or may be engineered for secretion from the cell or for translocation to the plasma membrane of the cell.

[0183] In some embodiments, contacting a cell with an LNP composition including an mRNA may reduce the innate immune response of a cell to an exogenous polynucleotide. A cell may be contacted with a first LNP composition including a first amount of a first exogenous mRNA including a translatable region and the level of the innate immune response of the cell to the first exogenous mRNA may be determined. Subsequently, the cell may be contacted with a second LNP composition including a second amount of the first exogenous mRNA, the second amount being a lesser amount of the first exogenous mRNA compared to the first amount. Alternatively, the second composition may include a first amount of a second exogenous mRNA that is different from the first exogenous mRNA. The steps of contacting the cell with the first and second LNP compositions may be repeated one or more times. Additionally, efficiency of polypeptide production (e.g., translation) in the cell may be optionally determined, and the cell may be re-contacted with the first and/or second composition repeatedly until a target protein production efficiency is achieved.

[0184] In some embodiments, the present disclosure provides for the use of the LNPs comprising a compound of the present disclosure and a polynucleotide in the manufacture of a medicament for the treatment of a disease, disorder or condition. The disease, disorder or condition may be as described in any one or more embodiments herein.

[0185] The medicament may be for the prevention or treatment of a cancer, an infectious disease, an allergy, or an autoimmune disease. In embodiments, the medicament is a vaccine. The vaccine may be a tumor vaccine, an influenza vaccine, or a SARS-CoV-2 vaccine.

Use of LNPs in a Vaccine

[0186] In some embodiments, the LNP comprising a compound of the present disclosure and a polynucleotide may be a component of a vaccine. Vaccines include

compounds and preparations that are capable of providing immunity against one or more conditions related to infectious diseases and so may include mRNAs encoding infectious disease derived antigens and/or epitopes. Vaccines also include compounds and preparations that direct an immune response against cancer cells and can include mRNAs
5 encoding tumor cell derived antigens, epitopes, and/or neoepitopes. Compounds eliciting immune responses may include vaccines, corticosteroids (e.g., dexamethasone), and other species.

[0187] In embodiments, the vaccine may be an mRNA vaccine and so the LNP comprising a compound of the present disclosure encapsulates or is associated with an
10 mRNA molecule which comprises an mRNA sequence encoding an antigenic peptide or protein, or a fragment, variant or derivative thereof.

[0188] The antigenic peptides or proteins may be pathogenic antigens, tumour antigens, allergenic antigens or autoimmune self-antigens. Such pathogenic antigens may be those derived from pathogenic organisms, in particular bacterial, viral or
15 protozoological (multicellular) pathogenic organisms, which evoke an immunological reaction in a mammalian subject, such as a human. Pathogenic antigens may be surface antigens, for example proteins or fragments thereof, located at the surface of the virus or the bacterial or protozoological organism.

[0189] Pathogenic antigens of interest may include those derived from one or more
20 of: *Acinetobacter baumannii*, *Anaplasma* genus, *Anaplasma phagocytophilum*, *Ancylostoma braziliense*, *Ancylostoma duodenale*, *Area nobacterium haemolyticum*, *Ascaris lumbricoides*, *Aspergillus* genus, *Astroviridae*, *Babesia* genus, *Bacillus anthracis*, *Bacillus cereus*, *Bartonella henselae*, BK virus, *Blastocysts hominis*, *Blastomyces dermatitidis*, *Bordetella pertussis*, *Borrelia burgdorferi*, *Borrelia* genus,
25 *Borrelia* spp, *Brucella* genus, *Brugia malayi*, *Bunyaviridae* family, *Burkholderia cepacia* and other *Burkholderia* species, *Burkholderia mallei*, *Burkholderia pseudomallei*, *Caliciviridae* family, *Campylobacter* genus, *Candida albicans*, *Candida* spp, *Chlamydia trachomatis*, *Chlamydophila pneumoniae*, *Chlamydophila psittaci*, QD prion, *Clonorchis sinensis*, *Clostridium botulinum*, *Clostridium difficile*, *Clostridium perfringens*,
30 *Clostridium perfringens*, *Clostridium* spp, *Clostridium tetani*, *Coccidioides* spp, coronaviruses, *Corynebacterium diphtheriae*, *Coxiella burnetii*, Crimean-Congo hemorrhagic

- fever virus, *Cryptococcus neoformans*, *Cryptosporidium* genus, Cytomegalovirus (CMV), Dengue viruses (DEN-1, DEN-2, DEN-3 and DEN-4), *Dientamoeba fragilis*, Ebolavirus (EBOV), *Echinococcus* genus, *Ehrlichia chaffeensis*, *Ehrlichia ewingii*, *Ehrlichia* genus, *Entamoeba histolytica*, *Enterococcus* genus, Enterovirus genus,
- 5 Enteroviruses, mainly Coxsackie A virus and Enterovirus 71 (EV71), *Epidermophyton* spp, Epstein-Barr Virus (EBV), *Escherichia coli* 0157:H7, O111 and O104: H4, *Fasciola hepatica* and *Fasciola gigantica*, FFI prion, Filarioidea superfamily, Flaviviruses, *Francisella tularensis*, *Fusobacterium* genus, *Geotrichum candidum*, *Giardia intestinalis*, *Gnathostoma* spp, GSS prion, Guanarito virus, *Haemophilus ducreyi*, *Haemophilus*
- 10 *influenzae*, *Helicobacter pylori*, Henipavirus (Hendra virus Nipah virus), Hepatitis A Virus, Hepatitis B Virus (HBV), Hepatitis C Virus (HCV), Hepatitis D Virus, Hepatitis E Virus, Herpes simplex virus 1 and 2 (HSV-1 and HSV-2), *Histoplasma capsulatum*, HIV (Human immunodeficiency virus), *Hortaea werneckii*, Human bocavirus (HBoV), Human herpesvirus 6 (HHV-6) and Human herpesvirus 7 (HHV-7), Human
- 15 metapneumovirus (hMPV), Human papillomavirus (HPV), Human parainfluenza viruses (HPIV), Japanese encephalitis virus, JC virus, Junin virus, *Kingella kingae*, *Klebsiella granulomatis*, Kuru prion, Lassa virus, *Legionella pneumophila*, *Leishmania* genus, *Leptospira* genus, *Listeria monocytogenes*, Lymphocytic choriomeningitis virus (LCMV), Machupo virus, *Malassezia* spp, Marburg virus, Measles virus, *Metagonimus*
- 20 *yokagawai*, Microsporidia phylum, *Molluscum contagiosum* virus (MCV), Mumps virus, *Mycobacterium leprae* and *Mycobacterium lepromatosis*, *Mycobacterium tuberculosis*, *Mycobacterium ulcerans*, *Mycoplasma pneumoniae*, *Naegleria fowleri*, *Necator americanus*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Nocardia asteroides*, *Nocardia* spp, *Onchocerca volvulus*, *Orientia tsutsugamushi*, Orthomyxoviridae family
- 25 (Influenza), *Paracoccidioides brasiliensis*, *Paragonimus* spp, *Paragonimus westermani*, Parvovirus B19, *Pasteurella* genus, *Plasmodium* genus, *Pneumocystis jirovecii*, Poliovirus, Rabies virus, Respiratory syncytial virus (RSV), Rhinovirus, rhinoviruses, *Rickettsia akari*, *Rickettsia* genus, *Rickettsia prowazekii*, *Rickettsia rickettsii*, *Rickettsia typhi*, Rift Valley fever virus, Rotavirus, Rubella virus, Sabia virus, *Salmonella* genus,
- 30 *Sarcoptes scabiei*, SARS coronavirus, *Schistosoma* genus, *Shigella* genus, Sin Nombre virus, Hantavirus, *Sporothrix schenckii*, *Staphylococcus* genus, *Staphylococcus* genus,

- Streptococcus agalactiae, Streptococcus pneumoniae, Streptococcus pyogenes, Strongyloides stercoralis, Taenia genus, Taenia solium, Tick-borne encephalitis virus (TBEV), Toxocara canis or Toxocara cati, Toxoplasma gondii, Treponema pallidum, Trichinella spiralis, Trichomonas vaginalis, Trichophyton spp, Trichuris trichiura,
- 5 Trypanosoma brucei, Trypanosoma cruzi, Ureaplasma urealyticum, Varicella zoster virus (VZV), Varicella zoster virus (VZV), Variola major or Variola minor, vCJD prion, Venezuelan equine encephalitis virus, Vibrio cholerae, West Nile virus, Western equine encephalitis virus, Wuchereria bancrofti, Yellow fever virus, Yersinia enterocolitica, Yersinia pestis, and Yersinia pseudotuberculosis.
- 10 [0190] In certain embodiments, relevant antigens may be derived from the pathogens selected from: Severe Acute Respiratory Syndrome (SARS), Severe Acute Respiratory Syndrome Coronavirus and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-1 and SARS-CoV-2), Influenza virus, respiratory syncytial virus (RSV), Herpes simplex virus (HSV), human Papilloma virus (HPV), Human
- 15 immunodeficiency virus (HIV), Plasmodium, Staphylococcus aureus, Dengue virus, Chlamydia trachomatis, Cytomegalovirus (CMV), Hepatitis B virus (HBV), Mycobacterium tuberculosis, Rabies virus, and Yellow Fever Virus.
- [0191] In some embodiments, the relevant pathogenic antigen may be selected from: Outer membrane protein A OmpA, biofilm associated protein Bap, transport
- 20 protein MucK (Acinetobacter baumannii, Acinetobacter infections)); variable surface glycoprotein VSG, microtubule-associated protein MAPP15, trans-sialidase TSA (Trypanosoma brucei, African sleeping sickness (African trypanosomiasis)); HIV p24 antigen, HIV envelope proteins (Gp120, Gp41, Gp160), polyprotein GAG, negative factor protein Nef, trans-activator of transcription Tat (HIV (Human immunodeficiency virus),
- 25 AIDS (Acquired immunodeficiency syndrome)); galactose-inhibitable adherence protein GIAP, 29 kDa antigen Eh29, Gal/GalNAc lectin, protein CRT, 125 kDa immunodominant antigen, protein M17, adhesin ADH112, protein STIRP (Entamoeba histolytica, Amoebiasis); Major surface proteins 1-5 (MSP1a, MSP1b, MSP2, MSP3, MSP4, MSP5), type IV secretion system proteins (VirB2, VirB7, VirB11, VirD4)
- 30 (Anaplasma genus, Anaplasmosis); protective Antigen PA, edema factor EF, lethal factor LF, the S-layer homology proteins SLH (Bacillus anthracis, Anthrax);

- acranolysin, phospholipase D, collagen-binding protein CbpA (*Area nobacterium* haemolyticum, *Area nobacterium* haemolyticum infection); nucleocapsid protein NP, glycoprotein precursor GPC, glycoprotein GP1, glycoprotein GP2 (Junin virus, Argentine hemorrhagic fever); chitin-protein layer proteins, 14 kDa surface antigen
- 5 A14, major sperm protein MSP, MSP polymerization -organizing protein MPOP, MSP fiber protein 2 MFP2, MSP polymerization -activating kinase MPAK, ABA-1-like protein ALB, protein ABA-1, cuticulin CUT-1 (*Ascaris lumbricoides*, Ascariasis); 41 kDa allergen Asp v13, allergen Asp f3, major conidial surface protein rodlet A, protease Peplp, GPI-anchored protein Gellp, GPI-anchored protein Crflp (*Aspergillus* genus,
- 10 Aspergillosis); family VP26 protein, VP29 protein (Astroviridae, Astrovirus infection); Rhoptry-associated protein 1 RAP-1, merozoite surface antigens MSA-1, MSA-2 (a1, a2, b, c), 12D3, 11C5, 21B4, P29, variant erythrocyte surface antigen VESA1, Apical Membrane Antigen 1 AMA-1 (*Babesia* genus, Babesiosis); hemolysin, enterotoxin C, PXOI-51, glycolate oxidase, ABC-transporter, penicillin-binding protein, zinc
- 15 transporter family protein, pseudouridine synthase Rsu, plasmid replication protein RepX, oligoendopeptidase F, prophage membrane protein, protein HemK, flagellar antigen H, 28.5-kDa cell surface antigen (*Bacillus cereus*, *Bacillus cereus* infection); large T antigen LT, small T antigen, capsid protein VP1, capsid protein VP2 (BK virus, BK virus infection); 29 kDa-protein, caspase-3-like antigens, glycoproteins (*Blastocystis*
- 20 *hominis*, *Blastocystis hominis* infection); yeast surface adhesin WI-1 (*Blastomyces dermatitidis*, Blastomycosis); nucleoprotein N, polymerase L, matrix protein Z, glycoprotein GP (Machupo virus, Bolivian hemorrhagic fever); outer surface protein A OspA, outer surface protein OspB, outer surface protein OspC, decorin binding protein A DbpA, decorin binding protein B DbpB, flagellar filament 41 kDa core protein Fla,
- 25 basic membrane protein A precursor BmpA (Immunodominant antigen P39), outer surface 22 kDa lipoprotein precursor (antigen IPLA7), variable surface lipoprotein vlsE (*Borrelia* genus, *Borrelia* infection); Botulinum neurotoxins BoNT/A1, BoNT/A2, BoNT/A3, BoNT/B, BoNT/C, BoNT/D, BoNT/E, BoNT/F, BoNT/G, recombinant botulinum toxin F He domain FHc (*Clostridium botulinum*, Botulism (and Infant
- 30 botulism)); nucleocapsid, glycoprotein precursor (Sabia virus, Brazilian hemorrhagic fever); copper/Zinc superoxide dismutase SodC, bacterioferritin Bfr, 50S ribosomal

protein RplL, OmpA-like transmembrane domain-containing protein Omp31, immunogenic 39-kDa protein M5 P39, zinc ABC transporter periplasmic zinc-binding protein znuA, periplasmic immunogenic protein Bp26, 30S ribosomal protein S12 RpsL, glyceraldehyde-3-phosphate dehydrogenase Gap, 25 kDa outer-membrane immunogenic

5 protein precursor Omp25, invasion protein B lalB, trigger factor Tig, molecular chaperone DnaK, putative peptidyl-prolyl cis-trans isomerase SurA, lipoprotein Omp19, outer membrane protein MotY Omp16, conserved outer membrane protein D15, malate dehydrogenase Mdh, component of the Type-IV secretion system (T4SS) VirJ, lipoprotein of unknown function BAB1_0187 (*Brucella* genus, Brucellosis); members of

10 the ABC transporter family (LolC, OppA, and PotF), putative lipoprotein releasing system transmembrane protein LolC/E, flagellin FliC, *Burkholderia* intracellular motility A BimA, bacterial Elongation factor-Tu EF-Tu, 17 kDa OmpA-like protein, *boaA* coding protein, *boaB* coding protein (*Burkholderia cepacia* and other *Burkholderia* species, *Burkholderia* infection); mycolyl-transferase Ag85A, heat-shock protein Hsp65, protein

15 TB10.4, 19 kDa antigen, protein PstS3, heat-shock protein Hsp70 (*Mycobacterium ulcerans*, Buruli ulcer); norovirus major and minor viral capsid proteins VP1 and VP2, genome polyprotein, Sapovirus capsid protein VP1, protein Vp3, geome polyprotein (*Caliciviridae* family, *Calicivirus* infection (*Norovirus* and *Sapovirus*)); major outer membrane protein PorA, flagellin FlaA, surface antigen CjaA, fibronectin binding

20 protein CadF, aspartate/glutamate-binding ABC transporter protein Peb1A, protein FspA1, protein FspA2 (*Campylobacter* genus, *Campylobacteriosis*); glycolytic enzyme enolase, secreted aspartyl proteinases SAPI-10, glycoposphatidylinositol (GPI)-linked cell wall protein, protein Hyr1, complement receptor 3-related protein CR3-RP, adhesin Als3p, heat shock protein 90 kDa hsp90, cell surface hydrophobicity protein CSH

25 (usually *Candida albicans* and other *Candida* species, *Candidiasis*); 17-kDa antigen, protein P26, trimeric autotransporter adhesins TAAs, *Bartonella* adhesin A BadA, variably expressed outer-membrane proteins Vomps, protein Pap3, protein HbpA, envelope-associated protease HtrA, protein OMP89, protein GroEL, protein LalB, protein OMP43, dihydrolipoamide succinyltransferase SucB (*Bartonella henselae*, Cat-scratch disease); amastigote surface protein-2, amastigote-specific surface protein SSP4, cruzipain, trans-sialidase TS, trypomastigote surface glycoprotein TSA-1, complement

30

- regulatory protein CRP-10, protein G4, protein G2, paraxonemal rod protein PAR2, paraflagellar rod component Pari, mucin -Associated Surface Proteins MPSP (Trypanosoma cruzi, Chagas Disease (American trypanosomiasis)); envelope glycoproteins (gB, gC, gE, gH, gI, gK, gL), (Varicella zoster virus (VZV), Chickenpox);
- 5 major outer membrane protein MOMP, probable outer membrane protein PMPC, outer membrane complex protein B OmcB, heat shock proteins Hsp60 HSP10, protein IncA, proteins from the type III secretion system, ribonucleotide reductase small chain protein NrdB, plasmid protein Pgp3, chlamydial outer protein N CopN, antigen CT521, antigen CT425, antigen CT043, antigen TC0052, antigen TC0189, antigen TC0582, antigen
- 10 TC0660, antigen TC0726, antigen TC0816, antigen TC0828 (Chlamydia trachomatis, Chlamydia); low calcium response protein E LCrE, chlamydial outer protein N CopN, serine/threonine-protein kinase PknD, acyl-carrier-protein S-malonyltransferase FabD, single-stranded DNA-binding protein Ssb, major outer membrane protein MOMP, outer membrane protein 2 Omp2, polymorphic membrane protein family (Pmpl1, Pmp2, Pmp3,
- 15 Pmp4, Pmp5, Pmp6, Pmp7, Pmp8, Pmp9, PmplO, PmplI, Pmpl2, Pmpl3, Pmpl4, Pmpl5, Pmpl6, Pmpl7, Pmpl8, Pmpl9, Pmp20, Pmp21), (Chlamydophila pneumoniae, Chlamydophila pneumoniae infection); cholera toxin B CTB, toxin coregulated pilin A TcpA, toxin coregulated pilin TcpF, toxin co-regulated pilus biosynthesis protein F TcpF, cholera enterotoxin subunit A, cholera enterotoxin subunit B, Heat-stable
- 20 enterotoxin ST, mannose-sensitive hemagglutinin MSHA, outer membrane protein U Porin ompU, Poring B protein, polymorphic membrane protein-D (Vibrio cholerae, Cholera); propionyl-CoA carboxylase PCC, 14-3-3 protein, prohibitin, cysteine proteases, glutathione transferases, gelsolin, cathepsin L proteinase CatL, Tegumental Protein 20.8 kDa TP20.8, tegumental protein 31.8 kDa TP31.8, lysophosphatidic acid
- 25 phosphatase LPAP, (Clonorchis sinensis, Clonorchiasis); surface layer proteins SLPs, glutamate dehydrogenase antigen GDH, toxin A, toxin B, cysteine protease Cwp84, cysteine protease Cwpl3, cysteine protease Cwpl9, Cell Wall Protein CwpV, flagellar protein FliC, flagellar protein FliD (Clostridium difficile, Clostridium difficile infection);
- 30 rhinoviruses: capsid proteins VP1, VP2, VP3, VP4; coronaviruses: spike proteins S, envelope proteins E, membrane proteins M, nucleocapsid proteins N (usually rhinoviruses and coronaviruses, Common cold (Acute viral rhinopharyngitis; Acute

- coryza)); prion protein Prp (CJD prion, Creutzfeldt -Jakob disease (CJD)); envelope protein Gc, envelope protein Gn, nucleocapsid proteins (Crimean-Congo hemorrhagic fever virus, Crimean-Congo hemorrhagic fever (CCHF)); virulence-associated DEAD-box RNA helicase VAD1, galactoxylomannan-protein GalXM, glucuronoxylomannan
- 5 GXM, mannoprotein MP (*Cryptococcus neoformans*, Cryptococcosis); acidic ribosomal protein P2 CpP2, mucin antigens Mucl, Muc2, Muc3 Muc4, Muc5, Muc6, Muc7, surface adherence protein CP20, surface adherence protein CP23, surface protein CP12, surface protein CP21, surface protein CP40, surface protein CP60, surface protein CP15, surface-associated glycopeptides gp40, surface-associated glycopeptides gpl5, oocyst wall
- 10 protein AB, profilin PRF, apyrase (*Cryptosporidium* genus, Cryptosporidiosis); fatty acid and retinol binding protein-1 FAR-1, tissue inhibitor of metalloproteinase TIMP (TMP), cysteine proteinase ACEY-1, cysteine proteinase ACCP-1, surface antigen Ac-16, secreted protein 2 ASP-2, metalloprotease 1 MTP-1, aspartyl protease inhibitor API-1, surface-associated antigen SAA-1, adult-specific secreted factor Xa serine protease
- 15 inhibitor anticoagulant AP, cathepsin D-like aspartic protease ARR-1 (usually *Ancylostoma braziliense*; multiple other parasites, Cutaneous larva migrans (CLM)); cathepsin L-like proteases, 53/25-kDa antigen, 8kDa family members, cysticercus protein with a marginal trypsin-like activity TsAg5, oncosphere protein TSOL18, oncosphere protein TSOL45-1A, lactate dehydrogenase A LDHA, lactate dehydrogenase
- 20 B LDHB (*Taenia solium*, Cysticercosis); pp65 antigen, membrane protein ppl5, capsid-proximal tegument protein ppl50, protein M45, DNA polymerase UL54, helicase UL105, glycoprotein gM, glycoprotein gN, glcoprotein H, glycoprotein B gB, protein UL83, protein UL94, protein UL99 (*Cytomegalovirus* (CMV), Cytomegalovirus infection); capsid protein C, premembrane protein prM, membrane protein M, envelope
- 25 protein E (domain I, domain II, domain II), protein NS1, protein NS2A, protein NS2B, protein NS3, protein NS4A, protein 2K, protein NS4B, protein NS5 (Dengue viruses (DEN-1, DEN-2, DEN-3 and DEN-4)-Flaviviruses, Dengue fever); 39 kDa protein (*Dientamoeba fragilis*, Dientamoebiasis); diphtheria toxin precursor Tox, diphtheria toxin DT, pilin-specific sortase SrtA, shaft pilin protein SpaA, tip pilin protein SpaC, minor
- 30 pilin protein SpaB, surface-associated protein DIP1281 (*Corynebacterium diphtheriae*, Diphtheria); glycoprotein GP, nucleoprotein NP, minor matrix protein VP24, major

- matrix protein VP40, transcription activator VP30, polymerase cofactor VP35, RNA polymerase L (Ebola virus (EBOV), Ebola hemorrhagic fever); prion protein (vQD prion, Variant Creutzfeldt-Jakob disease (vCJD, nvCJD)); UvrABC system protein B, protein FlpI, protein Flp2, protein Flp3, protein TadA, hemoglobin receptor HgbA, outer
- 5 membrane protein TdhA, protein CpsRA, regulator CpxR, protein SapA, 18 kDa antigen, outer membrane protein NcaA, protein LspA, protein LspA1, protein LspA2, protein LspB, outer membrane component DsrA, lectin DltA, lipoprotein Hip, major outer membrane protein OMP, outer membrane protein OmpA2 (*Haemophilus ducreyi*, Chancroid); aspartyl protease 1 Pepl, phospholipase B PLB, alpha-mannosidase 1
- 10 AMN1, glucanotransferase GEL1, urease URE, peroxisomal matrix protein Pmpl, proline-rich antigen Pra, humal T-cell reative protein TcrP (*Coccidioides immitis* and *Coccidioides posadasii*, Coccidioidomycosis); allergen Tri r 2, heat shock protein 60 Hsp60, fungal actin Act, antigen Tri r2, antigen Tri r4, antigen Tri tl, protein IV, glycerol-3-phosphate dehydrogenase Gpdl, osmosensor HwSholA, osmosensor HwSholB,
- 15 histidine kinase HwHhk7B, allergen Mala s 1, allergen Mala s 11, thioredoxin Trx Mala s 13, allergen Mala f, allergen Mala s (usually *Trichophyton* spp, *Epidermophyton* spp., *Malassezia* spp., *Hortaea werneckii*, Dermatophytosis); protein EG95, protein EG10, protein EG18, protein EgA31, protein EM18, antigen EPC1, antigen B, antigen 5, protein P29, protein 14-3-3, 8-kDa protein, myophilin, heat shock protein 20 HSP20,
- 20 glycoprotein GP-89, fatty acid binding protein FAPB (*Echinococcus* genus, Echinococcosis); major surface protein 2 MSP2, major surface protein 4 MSP4, MSP variant SGV1, MSP variant SGV2, outer membrane protein OMP, outer membrane protein 19 OMP-19, major antigenic protein MAPI, major antigenic protein MAP1-2, major antigenic protein MAP1B, major antigenic protein MAP1-3, Erum2510 coding
- 25 protein, protein GroEL, protein GroES, 30-kDa major outer membrane proteins, GE 100-kDa protein, GE 130-kDa protein, GE 160-kDa protein (*Ehrlichia* genus, Ehrlichiosis); secreted antigen SagA, sagA-like proteins SalA and SalB, collagen adhesin Scm, surface proteins FmsI (EbpA(fm), Fms5 (EbpB(fm), Fms9 (EpbC(fm) and FmsIO, protein EbpC(fm), 96 kDa immunoprotective glycoprotein Gl (*Enterococcus* genus,
- 30 *Enterococcus* infection); genome polyprotein, polymerase 3D, viral capsid protein VP1, viral capsid protein VP2, viral capsid protein VP3, viral capsid protein VP4, protease 2A,

protease 3C (Enterovirus genus, Enterovirus infection); outer membrane proteins OM, 60 kDa outer membrane protein, cell surface antigen OmpA, cell surface antigen OmpB (sca5), 134 kDa outer membrane protein, 31 kDa outer membrane protein, 29.5 kDa outer membrane protein, cell surface protein SCA4, cell surface protein Adr1 (RP827), cell surface protein Adr2 (RP828), cell surface protein SCA1, Invasion protein invA, cell division protein fts, secretion proteins sec Ofamily, virulence proteins virB, tlyA, tlyC, parvulin-like protein Pip, preprotein translocase SecA, 120-kDa surface protein antigen SPA, 138 kD complex antigen, major 100-kD protein (protein I), intracytoplasmic protein D, protective surface protein antigen SPA (*Rickettsia prowazekii*, Epidemic typhus); Epstein-Barr nuclear antigens (EBNA-1, EBNA-2, EBNA-3A, EBNA-3B, EBNA-3C, EBNA-leader protein (EBNA-LP)), latent membrane proteins (LMP-1, LMP-2A, LMP-2B), early antigen EBV-EA, membrane antigen EBV-MA, viral capsid antigen EBV-VCA, alkaline nuclease EBV-AN, glycoprotein H, glycoprotein gp350, glycoprotein gp110, glycoprotein gp42, glycoprotein gHgL, glycoprotein gB (Epstein-Barr Virus (EBV), Epstein-Barr Virus Infectious Mononucleosis); capsid protein VP2, capsid protein VP1, major protein NS1 (Parvovirus B19, Erythema infectiosum (Fifth disease)); pp65 antigen, glycoprotein 105, major capsid protein, envelope glycoprotein H, protein U51 (Human herpesvirus 6 (HHV-6) and Human herpesvirus 7 (HHV-7), Exanthem subitum); thioredoxin-glutathione reductase TGR, cathepsins LI and L2, Kunitz-type protein KTM, leucine aminopeptidase LAP, cysteine proteinase Fas2, saposin-like protein-2 SAP-2, thioredoxin peroxidases TPx, Prx-1, Prx-2, cathepsin I cysteine proteinase CL3, protease cathepsin L CL1, phosphoglycerate kinase PGK, 27-kDa secretory protein, 60 kDa protein HSP35alpha, glutathione transferase GST, 28.5 kDa tegumental antigen 28.5 kDa TA, cathepsin B3 protease CatB3, Type I cystatin stefin-1, cathepsin L5, cathepsin Llg and cathepsin B, fatty acid binding protein FABP, leucine aminopeptidases LAP (*Fasciola hepatica* and *Fasciola gigantica*, Fasciolosis); prion protein (FFI prion, Fatal familial insomnia (FFI)); venom allergen homolog-like protein VAL-1, abundant larval transcript ALT-1, abundant larval transcript ALT- 2, thioredoxin peroxidase TPX, vespid allergen homologue VAH, thioredoxin peroxidase 2 TPX-2, antigenic protein SXP (peptides N, N1, N2, and N3), activation associated protein-1 ASP-1, Thioredoxin TRX, transglutaminase BmTGA, glutathione-S-

transferases GST, myosin, vespid allergen homologue VAH, 175 kDa collagenase, glyceraldehyde-3-phosphate dehydrogenase GAPDH, cuticular collagen Col-4, secreted larval acidic proteins SLAPs, chitinase CHI-1, maltose binding protein MBP, glycolytic enzyme fructose-1,6-bisphosphate aldolase Fba, tropomyosin TMY-1, nematode specific

5 gene product OvB20, onchocystatin CPI-2, Cox-2 (Filarioidea superfamily, Filariasis); phospholipase C PLC, heat-labile enterotoxin B, Iota toxin component Ib, protein CPE1281, pyruvate ferredoxin oxidoreductase, elongation factor G EF-G, perfringolysin

10 0 Pfo, glyceraldehyde-3-phosphate dehydrogenase GapC, Fructose-bisphosphate aldolase Alf2, Clostridium perfringens enterotoxin CPE, alpha toxin AT, alpha toxoid

15 ATd, epsilon-toxoid ETd, protein HP, large cytotoxin TpeL, endo-beta-N-acetylglucosaminidase Naglu, phosphoglyceromutase Pgm (Clostridium perfringens, Food poisoning by Clostridium perfringens); leukotoxin IktA, adhesion FadA, outer membrane protein RadD, high-molecular weight arginine-binding protein (Fusobacterium genus, Fusobacterium infection); phospholipase C PLC, heat-labile

20 enterotoxin B, Iota toxin component Ib, protein CPE1281, pyruvate ferredoxin oxidoreductase, elongation factor G EF-G, perfringolysin 0 Pfo, glyceraldehyde-3-phosphate dehydrogenase GapC, fructose-bisphosphate aldolase Alf2, Clostridium perfringens enterotoxin CPE, alpha toxin AT, alpha toxoid ATd, epsilon-toxoid ETd, protein HP, large cytotoxin TpeL, endo-beta-N-acetylglucosaminidase Naglu,

25 phosphoglyceromutase Pgm (usually Clostridium perfringens; other Clostridium species, Gas gangrene (Clostridial myonecrosis)); lipase A, lipase B, peroxidase Decl (Geotrichum candidum, Geotrichosis); prion protein (GSS prion, Gerstmann-Straussler-Scheinker syndrome (GSS)); cyst wall proteins CWP1, CWP2, CWP3, variant surface protein VSP, VSP1, VSP2, VSP3, VSP4, VSP5, VSP6, 56 kDa antigen, pyruvate

30 ferredoxin oxidoreductase PFOR, alcohol dehydrogenase E ADHE, alpha-giardin, alpha8-giardin, alpha-giardin, beta-giardin, cysteine proteases, glutathione-S-transferase GST, arginine deiminase ADI, fructose-1,6-bisphosphate aldolase FBA, Giardia trophozoite antigens GTA (GTA1, GTA2), ornithine carboxyl transferase OCT, striated fiber-assemblin-like protein SALP, uridine phosphoryl-like protein UPL, alpha-tubulin, beta-tubulin (Giardia intestinalis, Giardiasis); members of the ABC transporter family (LolC, OppA, and PotF), putative lipoprotein releasing system transmembrane

protein LolC/E, flagellin FliC, Burkholderia intracellular motility A BimA, bacterial Elongation factor-Tu EF-Tu, 17 kDa OmpA-like protein, boaA coding protein (Burkholderia mallei, Glanders); cyclophilin CyP, 24 kDa third-stage larvae protein GS24, excretion -secretion products ESPs (40, 80, 120 and 208 kDa) (Gnathostoma spinigerum and Gnathostoma hispidum, Gnathostomiasis); pilin proteins, minor pilin-associated subunit pilC, major pilin subunit and variants pilE, pilS, phase variation protein porA, Porin B PorB, protein TraD, Neisserial outer membrane antigen H.8, 70kDa antigen, major outer membrane protein PI, outer membrane proteins PIA and PIB, W antigen, surface protein A NspA, transferrin binding protein TbpA, transferrin binding protein TbpB, PBP2, mtrR coding protein, ponA coding protein, membrane permease FbpBC, FbpABC protein system, LbpAB proteins, outer membrane protein Opa, outer membrane transporter FetA, iron -repressed regulator MpeR (Neisseria gonorrhoeae, Gonorrhea); outer membrane protein A OmpA, outer membrane protein C OmpC, outer membrane protein K17 OmpK17 (Klebsiella granulomatis, Granuloma inguinale (Donovanosis)); fibronectin-binding protein Sfb, fibronectin/fibrinogen-binding protein FBP54, fibronectin-binding protein FbaA, M protein type 1 Emml, M protein type 6 Emm6, immunoglobulin-binding protein 35 Sib35, Surface protein R28 Spr28, superoxide dismutase SOD, C5a peptidase ScpA, antigen I/II Agl/II, adhesin AspA, G-related alpha2-macroglobulin-binding protein GRAB, surface fibrillar protein M5 (Streptococcus pyogenes, Group A streptococcal infection); C protein β antigen, arginine deiminase proteins, adhesin BibA, 105 kDa protein BPS, surface antigens c, surface antigens R, surface antigens X, trypsin-resistant protein R1, trypsin-resistant protein R3, trypsin-resistant protein R4, surface immunogenic protein Sip, surface protein Rib, Leucine-rich repeats protein LrrG, serine-rich repeat protein Srr-2, C protein alpha-antigen Bca, Beta antigen Bag, surface antigen Epsilon, alpha-like protein ALP1, alpha-like protein ALP5 surface antigen delta, alpha-like protein ALP2, alphas-like protein ALP3, alpha-like protein ALP4, Cbeta protein Bac (Streptococcus agalactiae, Group B streptococcal infection); transferrin-binding protein 2 Tbp2, phosphatase P4, outer membrane protein P6, peptidoglycan-associated lipoprotein Pal, protein D, protein E, adherence and penetration protein Hap, outer membrane protein 26 Omp26, outer membrane protein P5 (Fimbrin), outer membrane protein D15, outer membrane protein

OmpP2, 5'-nucleotidase NucA, outer membrane protein PI, outer membrane protein P2, outer membrane lipoprotein Pep, Lipoprotein E, outer membrane protein P4, fuculokinase FucK, [Cu,Zn]-superoxide dismutase SodC, protease HtrA, protein 0145, alpha-galactosylceramide (Haemophilus influenzae, Haemophilus influenzae infection);

5 polymerase 3D, viral capsid protein VP1, viral capsid protein VP2, viral capsid protein VP3, viral capsid protein VP4, protease 2A, protease 3C (Enteroviruses, mainly Coxsackie A virus and Enterovirus 71 (EV71), Hand, foot and mouth disease (HFMD)); RNA polymerase L, protein L, glycoprotein Gn, glycoprotein Gc, nucleocapsid protein S, envelope glycoprotein Gl, nucleoprotein NP, protein N, polyprotein M (Sin Nombre

10 virus, Hantavirus, Hantavirus Pulmonary Syndrome (HPS)); heat shock protein HspA, heat shock protein HspB, citrate synthase GltA, protein UreB, heat shock protein Hsp60, neutrophil-activating protein NAP, catalase KatA, vacuolating cytotoxin VacA, urease alpha UreA, urease beta UreB, protein Cpn10, protein groES, heat shock protein Hsp10, protein MopB, cytotoxicity-associated 10 kDa protein CAG, 36 kDa antigen, beta-

15 lactamase HcpA, Beta-lactamase HcpB (Helicobacter pylori, Helicobacter pylori infection); integral membrane proteins, aggregation-prone proteins, O-antigen, toxin-antigens Stx2B, toxin-antigen Stx1B, adhesion-antigen fragment Int28, protein EspA, protein EspB, Intimin, protein Tir, protein IntC300, protein Eae (Escherichia coli 0157:H7, O111 and O104:H4, Hemolytic-uremic syndrome (HUS)); RNA polymerase L,

20 protein L, glycoprotein Gn, glycoprotein Gc, nucleocapsid protein S, envelope glycoprotein Gl, nucleoprotein NP, protein N, polyprotein M (Bunyaviridae family, Hemorrhagic fever with renal syndrome (HFRS)); glycoprotein G, matrix protein M, nucleoprotein N, fusion protein F, polymerase L, protein W, proteinC, phosphoprotein p, non-structural protein V (Henipavirus (Hendra virus Nipah virus), Henipavirus

25 infections); polyprotein, glycoprotein Gp2, hepatitis A surface antigen HBsAg, protein 2A, virus protein VP1, virus protein VP2, virus protein VP3, virus protein VP4, protein P1B, protein P2A, protein P3AB, protein P3D (Hepatitis A Virus, Hepatitis A); hepatitis B surface antigen HBsAg, Hepatitis B core antigen HbcAg, polymerase, protein Hbx, preS2 middle surface protein, surface protein L, large S protein, virus protein VP1, virus

30 protein VP2, virus protein VP3, virus protein VP4 (Hepatitis B Virus (HBV), Hepatitis B); envelope glycoprotein E1 gp32 gp35, envelope glycoprotein E2 NS1 gp68 gp70,

capsid protein C, core protein Core, polyprotein, virus protein VP1, virus protein VP2,
 virus protein VP3, virus protein VP4, antigen G, protein NS3, protein NS5A, (Hepatitis
 C Virus, Hepatitis C); virus protein VP1, virus protein VP2, virus protein VP3, virus
 protein VP4, large hepatitis delta antigen, small hepatitis delta antigen (Hepatitis D
 5 Virus, Hepatitis D); virus protein VP1, virus protein VP2, virus protein VP3, virus protein
 VP4, capsid protein E2 (Hepatitis E Virus, Hepatitis E); glycoprotein L UL1, uracil-DNA
 glycosylase UL2, protein UL3, protein UL4, DNA replication protein UL5, portal protein
 UL6, virion maturation protein UL7, DNA helicase UL8, replication origin-binding
 protein UL9, glycoprotein M UL10, protein UL11, alkaline exonuclease UL12, serine-
 10 threonine protein kinase UL13, tegument protein UL14, terminase UL15, tegument
 protein UL16, protein UL17, capsid protein VP23 UL18, major capsid protein VP5
 UL19, membrane protein UL20, tegument protein UL21, Glycoprotein H (UL22),
 Thymidine Kinase UL23, protein UL24, protein UL25, capsid protein P40 (UL26, VP24,
 VP22A), glycoprotein B (UL27), ICP18.5 protein (UL28), major DNA-binding protein
 15 ICP8 (UL29), DNA polymerase UL30, nuclear matrix protein UL31, envelope
 glycoprotein UL32, protein UL33, inner nuclear membrane protein UL34, capsid protein
 VP26 (UL35), large tegument protein UL36, capsid assembly protein UL37, VP19C
 protein (UL38), ribonucleotide reductase (Large subunit) UL39, ribonucleotide
 reductase (Small subunit) UL40, tegument protein/virion host shutoff VHS protein
 20 (UL41), DNA polymerase processivity factor UL42, membrane protein UL43,
 glycoprotein C (UL44), membrane protein UL45, tegument proteins VP11/12 (UL46),
 tegument protein VP13/14 (UL47), virion maturation protein VP16 (UL48, Alpha-TIF),
 envelope protein UL49, dUTP diphosphatase UL50, tegument protein UL51, DNA
 helicase/primase complex protein UL52, glycoprotein K (UL53), transcriptional
 25 regulation protein IE63 (ICP27, UL54), protein UL55, protein UL56, viral replication
 protein ICP22 (IE68, US1), protein US2, serine/threonine-protein kinase US3,
 glycoprotein G (US4), glycoprotein J (US5), glycoprotein D (US6), glycoprotein I
 (US7), glycoprotein E (US8), tegument protein US9, capsid/tegument protein US10,
 Vmw21 protein (US11), ICP47 protein (IE12, US12), major transcriptional activator
 30 ICP4 (IE175, RSI), E3 ubiquitin ligase ICPO (IE110), latency-related protein 1 LRP1,
 latency-related protein 2 LRP2, neurovirulence factor RL1 (ICP34.5), latency-associated

transcript LAT (Herpes simplex virus 1 and 2 (HSV-1 and HSV-2), Herpes simplex); heat shock protein Hsp60, cell surface protein H1C, dipeptidyl peptidase type IV DppIV, M antigen, 70 kDa protein, 17 kDa histone-like protein (*Histoplasma capsulatum*, Histoplasmosis); fatty acid and retinol binding protein-1 FAR-1, tissue inhibitor of

5 metalloproteinase TIMP (TMP), cysteine proteinase ACEY-1, cysteine proteinase ACCP-1, surface antigen Ac-16, secreted protein 2 ASP-2, metalloprotease 1 MTP-1, aspartyl protease inhibitor API-1, surface-associated antigen SAA-1, surface-associated antigen SAA-2, adult-specific secreted factor Xa, serine protease inhibitor anticoagulant AP, cathepsin D-like aspartic protease ARR-1, glutathione S-transferase GST, aspartic

10 protease APR-1, acetylcholinesterase AChE (*Ancylostoma duodenale* and *Necator americanus*, Hookworm infection); protein NSI, protein NP1, protein VP1, protein VP2, protein VP3 (Human bocavirus (HBoV), Human bocavirus infection); major surface protein 2 MSP2, major surface protein 4 MSP4, MSP variant SGV1, MSP variant SGV2, outer membrane protein OMP, outer membrane protein 19 OMP-19, major antigenic

15 protein MAPI, major antigenic protein MAP1-2, major antigenic protein MAP1B, major antigenic protein MAP1-3, Erum2510 coding protein, protein GroEL, protein GroES, 30-kDa major outer membrane proteins, GE 100-kDa protein, GE 130-kDa protein, GE 160-kDa protein (*Ehrlichia ewingii*, Human ewingii ehrlichiosis); major surface proteins 1-5 (MSPIa, MSPIb, MSP2, MSP3, MSP4, MSP5), type IV secretion system proteins

20 VirB2, VirB7, VirB11, VirD4 (*Anaplasma phagocytophilum*, Human granulocytic anaplasmosis (HGA)); protein NSI, small hydrophobic protein NS2, SH protein, fusion protein F, glycoprotein G, matrix protein M, matrix protein M2-1, matrix protein M2-2, phosphoprotein P, nucleoprotein N, polymerase L (Human metapneumovirus (hMPV), Human metapneumovirus infection); major surface protein 2 MSP2, major surface

25 protein 4 MSP4, MSP variant SGV1, MSP variant SGV2, outer membrane protein OMP, outer membrane protein 19 OMP-19, major antigenic protein MAPI, major antigenic protein MAP1-2, major antigenic protein MAP1B, major antigenic protein MAP1-3, Erum2510 coding protein, protein GroEL, protein GroES, 30-kDa major outer membrane proteins, GE 100-kDa protein, GE 130-kDa protein, GE 160-kDa protein

30 (*Ehrlichia chaffeensis*, Human monocytic ehrlichiosis); replication protein E1, regulatory protein E2, protein E3, protein E4, protein E5, protein E6, protein E7, protein E8, major

- capsid protein LI, minor capsid protein L2 (Human papillomavirus (HPV), Human papillomavirus (HPV) infection); fusion protein F, hemagglutinin-neuramidase HN, glycoprotein G, matrix protein M, phosphoprotein P, nucleoprotein N, polymerase L (Human parainfluenza viruses (HPIV), Human parainfluenza virus infection);
- 5 Hemagglutinin (HA), Neuraminidase (NA), Nucleoprotein (NP), M1 protein, M2 protein, NS1 protein, NS2 protein (NEP protein: nuclear export protein), PA protein, PB1 protein (polymerase basic 1 protein), PB1-F2 protein and PB2 protein (Orthomyxoviridae family, Influenza virus (flu)); genome polyprotein, protein E, protein M, capsid protein C (Japanese encephalitis virus, Japanese encephalitis); RTX toxin, type IV pili, major
 - 10 pilus subunit PilA, regulatory transcription factors PilS and PilR, protein sigma54, outer membrane proteins (Kingella kingae, Kingella kingae infection); prion protein (Kuru prion, Kuru); nucleoprotein N, polymerase L, matrix protein Z, glycoprotein GP (Lassa virus, Lassa fever); peptidoglycan-associated lipoprotein PAL, 60 kDa chaperonin Cpn60 (groEL, HspB), type IV pilin PilE, outer membrane protein MIP, major outer
 - 15 membrane protein MompS, zinc metalloproteinase MSP (Legionella pneumophila, Legionellosis (Legionnaires' disease, Pontiac fever)); P4 nuclease, protein WD, ribonucleotide reductase M2, surface membrane glycoprotein Pg46, cysteine proteinase CP, glucose-regulated protein 78 GRP-78, stage-specific S antigen-like protein A2, ATPase Fl, beta-tubulin, heat shock protein 70 Hsp70, KMP-11, glycoprotein GP63,
 - 20 protein BT1, nucleoside hydrolase NH, cell surface protein Bl, ribosomal protein Pl-like protein PI, sterol 24-c-methy transferase SMT, LACK protein, histone HI, SPB1 protein, thiol specific antioxidant TSA, protein antigen STI1, signal peptidase SP, histone H2B, surface antigen PSA-2, cysteine proteinase b Cpb (Leishmania genus, Leishmaniasis); major membrane protein I, serine-rich antigen- 45 kDa, 10 kDa caperonin GroES, HSP
 - 25 kDa antigen, amino-oxononanoate synthase AONS, protein recombinase A RecA, Acetyl-/propionyl-coenzyme A carboxylase alpha, alanine racemase, 60 kDa chaperonin 2, ESAT-6-like protein EcxB (L-ESAT-6), protein Lsr2, protein ML0276, Heparin-binding hemagglutinin HBHA, heat-shock protein 65 Hsp65, mycPI or ML0041 coding protein, htrA2 or ML0176 coding protein, htrA4 or ML2659 coding protein, gcp or
 - 30 ML0379 coding protein, clpC or ML0235 coding protein (Mycobacterium leprae and Mycobacterium lepromatosis, Leprosy); outer membrane protein LipL32, membrane

protein LIC10258, membrane protein LP30, membrane protein LIC12238, Ompa-like protein Lsa66, surface protein LigA, surface protein LigB, major outer membrane protein OmpL1, outer membrane protein LipL41, protein LigAni, surface protein LcpA, adhesion protein LipL53, outer membrane protein UpL32, surface protein Lsa63, flagellin FlaBl, 5 membran lipoprotein LipL21, membrane protein pL40, leptospiral surface adhesin Lsa27, outer membrane protein OmpL36, outer membrane protein OmpL37, outer membrane protein OmpL47, outer membrane protein OmpL54, acyltransferase LpxA (*Leptospira* genus, *Leptospirosis*); listeriolysin O precursor Hly (LLO), invasion-associated protein lap (P60), Listeriolysin regulatory protein PrfA, Zinc 10 metalloproteinase Mpl, Phosphatidylinositol- specific phospholipase C PLC (PlcA, PlcB), O-acetyltransferase Oat, ABC-transporter permease Im.G_1771, adhesion protein LAP, LAP receptor Hsp60, adhesin LapB, haemolysin listeriolysin OLLO, protein ActA, Internalin A InIA, protein InIB (*Listeria monocytogenes*, *Listeriosis*); outer surface protein A OspA, outer surface protein OspB, outer surface protein OspC, decorin binding 15 protein A DbpA, decorin binding protein B DbpB, flagellar filament 41 kDa core protein Fla, basic membrane protein A BmpA (Immunodominant antigen P39), outer surface 22 kDa lipoprotein precursor (antigen IPLA7), variable surface lipoprotein vlsE (usually *Borrelia burgdorferi* and other *Borrelia* species, Lyme disease (Lyme borreliosis)); venom allergen homolog-like protein VAL-1, abundant larval transcript ALT-1, 20 abundant larval transcript ALT- 2, thioredoxin peroxidase TPX, vespid allergen homologue VAH, thioredoxin peroxidase 2 TPX-2, antigenic protein SXP (peptides N, N1, N2, and N3), activation associated protein-1 ASP-1, thioredoxin TRX, transglutaminase BmTGA, glutathione-S-transferases GST, myosin, vespid allergen homologue VAH, 175 kDa collagenase, glyceraldehyde-3-phosphate dehydrogenase 25 GAPDH, cuticular collagen Col-4, Secreted Larval Acidic Proteins SLAPs, chitinase CHI-1, maltose binding protein MBP, glycolytic enzyme fructose-1,6-bisphosphate aldolase Fba, tropomyosin TMY-1, nematode specific gene product OvB20, onchocystatin CPI-2, protein Cox-2 (*Wuchereria bancrofti* and *Brugia malayi*, Lymphatic filariasis (Elephantiasis)); glycoprotein GP, matrix protein Z, polymerase L, 30 nucleoprotein N (Lymphocytic choriomeningitis virus (LCMV), Lymphocytic choriomeningitis); thrombospondin-related anonymous protein TRAP, SSP2 Sporozoite

surface protein 2, apical membrane antigen 1 AMA1, rhoptry membrane antigen RMA1, acidic basic repeat antigen ABRA, cell-traversal protein PF, protein Pvs25, merozoite surface protein 1 MSP-1, merozoite surface protein 2 MSP-2, ring-infected erythrocyte surface antigen RESA, liver stage antigen 3 LSA-3, protein Eba-175, serine repeat antigen 5 SERA-5, circumsporozoite protein CS, merozoite surface protein 3 MSP3, merozoite surface protein 8 MSP8, enolase PF10, hepatocyte erythrocyte protein 17 kDa HEP17, erythrocyte membrane protein 1 EMP1, protein Kbeta merozoite surface protein 4/5 MSP 4/5, heat shock protein Hsp90, glutamate-rich protein GLURP, merozoite surface protein 4 MSP-4, protein STARP, circumsporozoite protein-related antigen precursor CRA (*Plasmodium* genus, Malaria); nucleoprotein N, membrane-associated protein VP24, minor nucleoprotein VP30, polymerase cofactor VP35, polymerase L, matrix protein VP40, envelope glycoprotein GP (Marburg virus, Marburg hemorrhagic fever (MHF)); protein C, matrix protein M, phosphoprotein P, non-structural protein V, hemagglutinin glycoprotein H, polymerase L, nucleoprotein N, fusion protein F (Measles virus, Measles); members of the ABC transporter family (LolC, OppA, and PotF), putative lipoprotein releasing system transmembrane protein LolC/E, flagellin FliC, *Burkholderia* intracellular motility A BimA, bacterial Elongation factor-Tu EF-Tu, 17 kDa OmpA-like protein, boaA coding protein, boaB coding protein (*Burkholderia pseudomallei*, Melioidosis (Whitmore's disease)); pilin proteins, minor pilin-associated subunit pilC, major pilin subunit and variants pilE, pilS, phase variation protein porA, Porin B PorB, protein TraD, *Neisseria* outer membrane antigen H.8, 70kDa antigen, major outer membrane protein PI, outer membrane proteins PIA and PIB, W antigen, surface protein A NspA, transferrin binding protein TbpA, transferrin binding protein TbpB, PBP2, mtrR coding protein, ponA coding protein, membrane permease FbpBC, FbpABC protein system, LbpAB proteins, outer membrane protein Opa, outer membrane transporter FetA, iron-repressed regulator MpeR, factor H-binding protein fHbp, adhesin NadA, protein NhbA, repressor FarR (*Neisseria meningitidis*, Meningococcal disease); 66 kDa protein, 22 kDa protein (usually *Metagonimus yokagawai*, Metagonimiasis); polar tube proteins (34, 75, and 170 kDa in *Glugea*, 35, 55 and 150kDa in *Encephalitozoon*), kinesin-related protein, RNA polymerase II largest subunit, similar to integral membrane protein YIPA, a nti-silencing protein 1, heat shock transcription

factor HSF, protein kinase, thymidine kinase, NOP-2 like nucleolar protein (Microsporidia phylum, Microsporidiosis); CASP8 and FADD-like apoptosis regulator, Glutathione peroxidase GPX1, RNA helicase NPH-II NPH2, Poly(A) polymerase catalytic subunit PAPL, Major envelope protein P43K, early transcription factor 70 kDa
5 subunit VETFS, early transcription factor 82 kDa subunit VETFL, metalloendopeptidase GI-type, nucleoside triphosphatase INPH1, replication protein A28-like MC134L, RNA polymease 7 kDa subunit RP07 (Mollusum contagiosum virus (MCV), Mollusum contagiosum (MC)); matrix protein M, phosphoprotein P/V, small hydrophobic protein SH, nucleoprotein N, protein V, fusion glycoprotein F, hemagglutinin-neuraminidase
10 HN, RNA polymerase L (Mumps virus, Mumps); Outer membrane proteins OM, cell surface antigen OmpA, cell surface antigen OmpB (sca5), cell surface protein SCA4, cell surface protein SCA1, intracytoplasmic protein D, crystalline surface layer protein SLP, protective surface protein antigen SPA (Rickettsia typhi, Murine typhus (Endemic typhus)); adhesin PI, adhesion P30, protein pll6, protein P40, cytoskeletal protein
15 HMW1, cytoskeletal protein HMW2, cytoskeletal protein HMW3, MPN152 coding protein, MPN426 coding protein, MPN456 coding protein, MPN-500coding protein (Mycoplasma pneumoniae, Mycoplasma pneumonia); NocA, Iron dependent regulatory protein, VapA, VapD, VapF, VapG, caseinolytic protease, filament tip-associated 43-kDa protein, protein P24, protein P61, 15-kDa protein, 56-kDa protein (usually Nocardia
20 asteroides and other Nocardia species, Nocardiosis); venom allergen homolog-like protein VAL-1, abundant larval transcript ALT-1, abundant larval transcript ALT- 2, thioredoxin peroxidase TPX, vespid allergen homologue VAH, thiordoxin peroxidase 2 TPX-2, antigenic protein SXP (peptides N, N1, N2, and N3), activation associated protein-1 ASP-1, Thioredoxin TRX, transglutaminase BmTGA, glutathione-S-
25 transferases GST, myosin, vespid allergen homologue VAH, 175 kDa collagenase, glyceraldehyde-3-phosphate dehydrogenase GAPDH, cuticular collagen Col-4, Secreted Larval Acidic Proteins SLAPs, chitinase CHI-1, maltose binding protein MBP, glycolytic enzyme fructose-1,6-bisphosphate aldolase Fba, tropomyosin TMY-1, nematode specific gene product OvB20, onchocystatin CPI-2, Cox-2 (Onchocerca
30 volvulus, Onchocerciasis (River blindness)); 43 kDa secreted glycoprotein, glycoprotein gpO, glycoprotein gp75, antigen Pb27, antigen Pb40, heat shock protein Hsp65, heat

shock protein Hsp70, heat shock protein Hsp90, protein P10, triosephosphate isomerase TPI, N-acetyl-glucosamine-binding lectin Paracoccin, 28 kDa protein Pb28 (Paracoccidioides brasiliensis, Paracoccidioidomycosis (South American blastomycosis)); 28-kDa cruzipain-like cystein protease Pw28CCP (usually

5 Paragonimus westermani and other Paragonimus species, Paragonimiasis); outer membrane protein OmpH, outer membrane protein Omp28, protein PM1539, protein PM0355, protein PM1417, repair protein MutL, protein BcbC, protein PM0305, formate dehydrogenase-N, protein PM0698, protein PM1422, DNA gyrase, lipoprotein PlpE, adhesive protein Cp39, heme acquisition system receptor HasR, 39 kDa capsular protein,

10 iron-regulated OMP IROMP, outer membrane protein OmpA87, fimbrial protein Ptf, fimbrial subunit protein PtfA, transferrin binding protein Tbpl, esterase enzyme MesA, Pasteurella multocida toxin PMT, adhesive protein Cp39 (Pasteurella genus, Pasteurellosis); "filamentous hemagglutinin FhaB, adenylate cyclase CyaA, pertussis toxin subunit 4 precursor PtxD, pertactin precursor Pm, toxin subunit 1 PtxA, protein

15 Cpn60, protein brkA, pertussis toxin subunit 2 precursor PtxB, pertussis toxin subunit 3 precursor PtxC, pertussis toxin subunit 5 precursor PtxE, pertactin Pm, protein Fim2, protein Fim3; " (Bordetella pertussis, Pertussis (Whooping cough)); "Fl capsule antigen, virulence-associated V antigen, secreted effector protein LcrV, V antigen, outer membrane protease Pla, secreted effector protein YopD, putative secreted protein-

20 tyrosine phosphatase YopH, needle complex major subunit YscF, protein kinase YopO, putative autotransporter protein YapF, inner membrane ABC-transporter YbtQ (Irp7), putative sugar binding protein YPO0612, heat shock protein 90 HtpG, putative sulfatase protein YdeN, outer-membrane lipoprotein carrier protein LolA, secretion chaperone YerA, putative lipoprotein YPO0420, hemolysin activator protein HpmB,

25 pesticin/yersiniabactin outer membrane receptor Psn, secreted effector protein YopE, secreted effector protein YopF, secreted effector protein YopK, outer membrane protein YopN, outer membrane protein YopM, Coagulase/fibrinolysin precursor Pla; " (Yersinia pestis, Plague); protein PhpA, surface adhesin PsaA, pneumolysin Ply, ATP-dependent protease CIP, lipote-protein ligase LplA, cell wall surface anchored protein

30 psrP, sortase SrtA, glutamyl-tRNA synthetase GltX, choline binding protein A CbpA, pneumococcal surface protein A PspA, pneumococcal surface protein C PspC, 6-

phosphogluconate dehydrogenase Gnd, iron-binding protein PiaA, Murein hydrolase
 LytB, proteon LytC, protease Al (*Streptococcus pneumoniae*, Pneumococcal infection);
 major surface protein B, kexin-like protease KEX1, protein A12, 55 kDa antigen P55,
 major surface glycoprotein Msg (*Pneumocystis jirovecii*, *Pneumocystis pneumonia*
 5 (PCP)); genome polypeptide, polymerase 3D, viral capsid protein VP1, viral capsid
 protein VP2, viral capsid protein VP3, viral capsid protein VP4, protease 2A, protease
 3C (*Poliovirus*, *Poliomyelitis*); protein Nf1, extendin-3, secretory lipase, cathepsin B-like
 protease, cysteine protease, cathepsin, peroxiredoxin, protein CrylAc (usually *Naegleria*
fowleri, Primary amoebic meningoencephalitis (PAM)); agnoprotein, large T antigen,
 10 small T antigen, major capsid protein VP1, minor capsid protein Vp2 (JC virus,
Progressive multifocal leukoencephalopathy); low calcium response protein E LCrE,
 chlamydial outer protein N CopN, serine/threonine-protein kinase PknD, acyl-carrier-
 protein S-malonyltransferase FabD, single-stranded DNA-binding protein Ssb, major
 outer membrane protein MOMP, outer membrane protein 2 Omp2, polymorphic
 15 membrane protein family (Pmp1, Pmp2, Pmp3, Pmp4, Pmp5, Pmp6, Pmp7, Pmp8, Pmp9,
 Pmp10, Pmp11, Pmp12, Pmp13, Pmp14, Pmp15, Pmp16, Pmp17, Pmp18, Pmp19, Pmp20,
 Pmp21) (*Chlamydia psittaci*, *Psittacosis*); outer membrane protein PI, heat shock
 protein B HspB, peptide ABC transporter, GTP-binding protein, protein IcmB,
 ribonuclease R, phosphatase SixA, protein DsbD, outer membrane protein TolC, DNA-
 20 binding protein PhoB, ATPase DotB, heat shock protein B HspB, membrane protein
 Coml, 28 kDa protein, DNA-3-methyladenine glycosidase I, outer membrane protein
 OmpH, outer membrane protein AdaA, glycine cleavage system T-protein (*Coxiella*
burnetii, Q fever); nucleoprotein N, large structural protein L, phosphoprotein P, matrix
 protein M, glycoprotein G (*Rabies virus*, *Rabies*); fusionprotein F, nucleoprotein N,
 25 matrix protein M, matrix protein M2-1, matrix protein M2-2, phosphoprotein P, small
 hydrophobic protein SH, major surface glycoprotein G, polymerase L, non-structural
 protein 1 NS1, non-structural protein 2 NS2 (*Respiratory syncytial virus* (RSV),
Respiratory syncytial virus infection); genome polypeptide, polymerase 3D, viral capsid
 protein VP1, viral capsid protein VP2, viral capsid protein VP3, viral capsid protein VP4,
 30 protease 2A, protease 3C (*Rhinovirus*, *Rhinovirus infection*); outer membrane proteins
 OM, cell surface antigen OmpA, cell surface antigen OmpB (sca5), cell surface protein

SCA4, cell surface protein SCA1, protein PS120, intracytoplasmic protein D, protective surface protein antigen SPA (*Rickettsia* genus, Rickettsial infection); outer membrane proteins OM, cell surface antigen OmpA, cell surface antigen OmpB (sca5), cell surface protein SCA4, cell surface protein SCA1, intracytoplasmic protein D (*Rickettsia akari*,
5 *Rickettsialpox*); envelope glycoprotein GP, polymerase L, nucleoprotein N, non-structural protein NSS (Rift Valley fever virus, Rift Valley fever (RVF)); outer membrane proteins OM, cell surface antigen OmpA, cell surface antigen OmpB (sca5), cell surface protein SCA4, cell surface protein SCA1, intracytoplasmic protein D (*Rickettsia rickettsii*, Rocky mountain spotted fever (RMSF)); non -structural protein 6
10 NS6, non -structural protein 2 NS2, intermediate capsid protein VP6, inner capsid protein VP2, non-structural protein 3 NS3, RNA-directed RNA polymerase L, protein VP3, non-structural protein 1 NS1, non -structural protein 5 NS5, outer capsid glycoprotein VP7, nonstructural glycoprotein 4 NS4, outer capsid protein VP4; (Rotavirus, Rotavirus infection); polyprotein P200, glycoprotein EI, glycoprotein E2, protein NS2, capsid
15 protein C (Rubella virus, Rubella); chaperonin GroEL (MopA), inositol phosphate phosphatase SopB, heat shock protein HsIU, chaperone protein DnaJ, protein TviB, protein IroN, flagellin FliC, invasion protein SipC, glycoprotein gp43, outer membrane protein LamB, outer membrane protein PagC, outer membrane protein TolC, outer membrane protein NmpC, outer membrane protein FadL, transport protein SadA,
20 transferase WgaP, effector proteins SifA, SteC, SseL, SseJ and SseF (*Salmonella* genus, Salmonellosis); "protein 14, non -structural protein NS7b, non -structural protein NS8a, protein 9b, protein 3a, nucleoprotein N, non-structural protein NS3b, non -structural protein NS6, protein 7a, non-structural protein NS8b, membrane protein M, envelope small membrane protein EsM, replicase polyprotein Ia, spike glycoprotein S, replicase
25 polyprotein Iab; SARS coronavirus, SARS (Severe Acute Respiratory Syndrome)); serin protease, Atypical *Sarcoptes* Antigen 1 ASA1, glutathione S-transferases GST, cysteine protease, serine protease, apolipoprotein (*Sarcoptes scabiei*, Scabies); glutathione S-transferases GST, paramyosin, hemoglobinase SM32, major egg antigen, 14 kDa fatty acid-binding protein Sml4, major larval surface antigen P37, 22,6 kDa tegumental
30 antigen, calpain CANP, triphosphate isomerase Tim, surface protein 9B, outer capsid protein VP2, 23 kDa integral membrane protein Sm23, Cu/Zn -superoxide dismutase,

glycoprotein Gp, myosin (*Schistosoma* genus, Schistosomiasis (Bilharziosis)); 60 kDa chaperonin, 56 kDa type-specific antigen, pyruvate phosphate dikinase, 4-hydroxybenzoate octaprenyltransferase (*Orientia tsutsugamushi*, Scrub typhus); dehydrogenase GuaB, invasion protein Spa32, invasin IpaA, invasin IpaB, invasin IpaC, 5 invasin IpaD, invasin IpaH, invasin IpaJ (*Shigella* genus, Shigellosis (Bacillary dysentery)); protein P53, virion protein US10 homolog, transcriptional regulator IE63, transcriptional transactivator IE62, protease P33, alpha trans-inducing factor 74 kDa protein, deoxyuridine 5'-triphosphate nucleotidohydrolase, transcriptional transactivator IE4, membrane protein UL43 homolog, nuclear phosphoprotein UL3 homolog, nuclear 10 protein UL4 homolog, replication origin-binding protein, membrane protein 2, phosphoprotein 32, protein 57, DNA polymerase processivity factor, portal protein 54, DNA primase, tegument protein UL14 homolog, tegument protein UL21 homolog, tegument protein UL55 homolog, tripartite terminase subunit UL33 homolog, tripartite terminase subunit UL15 homolog, capsid-binding protein 44, virion-packaging protein 15 43 (*Varicella zoster virus* (VZV), Shingles (Herpes zoster)); truncated 3-beta hydroxy-5-ene steroid dehydrogenase homolog, virion membrane protein A13, protein A19, protein A31, truncated protein A35 homolog, protein A37.5 homolog, protein A47, protein A49, protein A51, semaphorin-like protein A43, serine proteinase inhibitor 1, serine proteinase inhibitor 2, serine proteinase inhibitor 3, protein A6, protein B15, 20 protein C1, protein C5, protein C6, protein F7, protein F8, protein F9, protein F11, protein F14, protein F15, protein F16 (*Variola major* or *Variola minor*, Smallpox (*Variola*)); adhesin/glycoprotein gp70, proteases (*Sporothrix schenckii*, Sporotrichosis); heme-iron binding protein IsdB, collagen adhesin Cna, clumping factor A ClfA, protein MecA, fibronectin-binding protein A FnbA, enterotoxin type A EntA, enterotoxin type B EntB, 25 enterotoxin type C EntC1, enterotoxin type C EntC2, enterotoxin type D EntD, enterotoxin type E EntE, Toxic shock syndrome toxin-1 TSST-1, Staphylokinase, Penicillin binding protein 2a PBP2a (MecA), secretory antigen SssA (*Staphylococcus* genus, Staphylococcal food poisoning); heme-iron binding protein IsdB, collagen adhesin Cna, clumping factor A ClfA, protein MecA, fibronectin-binding protein A 30 FnbA, enterotoxin type A EntA, enterotoxin type B EntB, enterotoxin type C EntC1, enterotoxin type C EntC2, enterotoxin type D EntD, enterotoxin type E EntE, Toxic

shock syndrome toxin-1 TSST-1, Staphylokinase, Penicillin binding protein 2a PBP2a (MecA), secretory antigen SssA (Staphylococcus genus e.g. aureus, Staphylococcal infection); antigen Ss-IR, antigen NIE, strongylastacin, Na⁺-K⁺ ATPase Sseat-6, tropomyosin SsTmy-1, protein LEC-5, 41 kDa antigen P5, 41-kDa larval protein, 31-kDa

5 larval protein, 28-kDa larval protein (Strongyloides stercoralis, Strongyloidiasis); glycerophosphodiester phosphodiesterase GlpQ (Gpd), outer membrane protein TmpB, protein Tp92, antigen TpFl, repeat protein Tpr, repeat protein F TprF, repeat protein G TprG, repeat protein I TprI, repeat protein J TprJ, repeat protein K TprK, treponemal membrane protein A TmpA, lipoprotein, 15 kDa Tpp15, 47 kDa membrane antigen,

10 miniferritin TpFl, adhesin Tp0751, lipoprotein TP0136, protein TpN17, protein TpN47, outer membrane protein TP0136, outer membrane protein TP0155, outer membrane protein TP0326, outer membrane protein TP0483, outer membrane protein TP0956 (Treponema pallidum, Syphilis); Cathepsin L-like proteases, 53/25-kDa antigen, 8kDa family members, cysticercus protein with a marginal trypsin-like activity TsAg5,

15 oncosphere protein TSOL18, oncosphere protein TSOL45-1A, lactate dehydrogenase A LDHA, lactate dehydrogenase B LDHB (Taenia genus, Taeniasis); tetanus toxin TetX, tetanus toxin C TTC, 140 kDa S layer protein, flavoprotein beta-subunit CT3, phospholipase (lecithinase), phosphocarrier protein HPr (Clostridium tetani, Tetanus (Lockjaw)); genome polyprotein, protein E, protein M, capsid protein C (Tick-borne

20 encephalitis virus (TBEV), Tick-borne encephalitis); 58-kDa antigen, 68-kDa antigens, Toxocara larvae excretory-secretory antigen TES, 32-kDa glycoprotein, glycoprotein TES-70, glycoprotein GP31, excretory-secretory antigen TcES-57, perienteric fluid antigen Pe, soluble extract antigens Ex, excretory/secretory larval antigens ES, antigen TES-120, polyprotein allergen TBA-1, cathepsin L-like cysteine protease c-cpl-1, 26-

25 kDa protein (Toxocara canis or Toxocara cati, Toxocariasis (Ocular Larva Migrans (OLM) and Visceral Larva Migrans (VLM))); microneme proteins (MIC1, MIC2, MIC3, MIC4, MIC5, MIC6, MIC7, MIC8), rhoptry protein Rop2, rhoptry proteins (Rop1, Rop2, Rop3, Rop4, Rop5, Rop6, Rop7, Rop16, Rjop17), protein SRL, surface antigen P22, major antigen p24, major surface antigen p30, dense granule proteins (GRA1, GRA2, GRA3,

30 GRA4, GRA5, GRA6, GRA7, GRA8, GRA9, GRA10), 28 kDa antigen, surface antigen SAG1, SAG2 related antigen, nucleoside-triphosphatase 1, nucleoside-triphosphatase 2,

protein Stt3, HesB-like domain-containing protein, rhomboid-like protease 5, toxomepsin 1 (*Toxoplasma gondii*, Toxoplasmosis); 43 kDa secreted glycoprotein, 53 kDa secreted glycoprotein, paramyosin, antigen Ts21, antigen Ts87, antigen p46000, TSL-1 antigens, caveolin-1 CAV-1, 49 kDa newborn larva antigen, prosaposin

5 homologue, serine protease, serine proteinase inhibitor, 45 -kDa glycoprotein Gp45 (*Trichinella spiralis*, Trichinellosis); Myb-like transcriptional factors (Myb1, Myb2, Myb3), adhesion protein AP23, adhesion protein AP33, adhesin protein AP33-3, adhesins AP51, adhesin AP65, adhesion protein AP65-1, alpha-actinin, kinesin-associated protein, teneurin, 62 kDa proteinase, subtilisin-like serine protease SUB1,

10 cysteine proteinase gene 3 CP3, alpha-enolase Enol, cysteine proteinase CP30, heat shock proteins (Hsp70, Hsp60) , immunogenic protein P270, (*Trichomonas vaginalis*, Trichomoniasis); beta-tubulin, 47-kDa protein, secretory leucocyte-like proteinase-1 SLP-1, 50-kDa protein TT50, 17 kDa antigen, 43/47 kDa protein (*Trichuris trichiura*, Trichuriasis (Whipworm infection)); protein ESAT-6 (EsxA), 10 kDa filtrate antigen

15 EsxB, secreted antigen 85-B FBPB, fibronectin-binding protein A FbpA (Ag85A), serine protease PepA, PPE family protein PPE18, fibronectin-binding protein D FbpD, immunogenic protein MPT64, secreted protein MPT51, catalase-peroxidase-peroxynitritase T KATG, periplasmic phosphate-binding lipoprotein PSTS3 (PBP-3, Phos-1), iron-regulated heparin binding hemagglutinin Hbha, PPE family protein PPE14,

20 PPE family protein PPE68, protein Mtb72F, protein Apa, immunogenic protein MPT63, periplasmic phosphate-binding lipoprotein PSTS1 (PBP-1), molecular chaperone DnaK, cell surface lipoprotein Mpt83, lipoprotein P23, phosphate transport system permease protein pstA, 14 kDa antigen, fibronectin-binding protein C FbpCl, Alanine dehydrogenase TB43, Glutamine synthetase 1, ESX-1 protein, protein CFP10, TB10.4

25 protein, protein MPT83, protein MTB12, protein MTB8, Rpf-like proteins, protein MTB32, protein MTB39, crystallin, heat -shock protein HSP65, protein PST-S (usually *Mycobacterium tuberculosis*, Tuberculosis); outer membrane protein FobA, outer membrane protein FobB, intracellular growth locus IglCl, intracellular growth locus IglC2, aminotransferase Wbtl, chaperonin GroEL, 17 kDa major membrane protein

30 TUL4, lipoprotein LpnA, chitinase family 18 protein, isocitrate dehydrogenase, Nif3 family protein, type IV pili glycosylation protein, outer membrane protein tolC, FAD

binding family protein, type IV pilin multimeric outer membrane protein, two component sensor protein KdpD, chaperone protein DnaK, protein TolQ (*Francisella tularensis*, Tularemia); "MB antigen, urease, protein GyrA, protein GyrB, protein ParC, protein ParE, lipid associated membrane proteins LAMP, thymidine kinase TK, phospholipase

5 PL-A1, phospholipase PL-A2, phospholipase PL-C, surface-expressed 96-kDa antigen; (*Ureaplasma urealyticum*, *Ureaplasma urealyticum* infection); non-structural polyprotein, structural polyprotein, capsid protein CP, protein El, protein E2, protein E3, protease PI, protease P2, protease P3 (Venezuelan equine encephalitis virus, Venezuelan equine encephalitis); glycoprotein GP, matrix protein Z, polymerase L, nucleoprotein N

10 (Guanarito virus, Venezuelan hemorrhagic fever); polyprotein, protein E, protein M, capsid protein C, protease NS3, protein NS1, protein NS2A, protein AS2B, protein NS4A, protein NS4B, protein NS5 (West Nile virus, West Nile Fever); capsid protein CP, protein El, protein E2, protein E3, protease P2 (Western equine encephalitis virus, Western equine encephalitis); genome polyprotein, protein E, protein M, capsid protein

15 C, protease NS3, protein NS1, protein NS2A, protein AS2B, protein NS4A, protein NS4B, protein NS5 (Yellow fever virus, Yellow fever); putative Yop targeting protein YobB, effector protein YopD, effector protein YopE, protein YopH, effector protein YopJ, protein translocation protein YopK, effector protein YopT, protein YpkA, flagellar biosyntheses protein FlhA, peptidase M48, potassium efflux system KefA,

20 transcriptional regulator RovA, adhesin Ifp, translocator protein LcrV, protein PcrV, invasin Inv, outer membrane protein OmpF-like porin, adhesin YadA, protein kinase C, phospholipase CI, protein PsaA, mannosyltransferase-like protein WbyK, protein YscU, antigen YPMa (*Yersinia pseudotuberculosis*, *Yersinia pseudotuberculosis* infection); effector protein YopB, 60 kDa chaperonin, protein WbcP, tyrosine protein phosphatase

25 YopH, protein YopQ, enterotoxin, Galactoside permease, reductase NrdE, protein YasN, Invasin Inv, adhesin YadA, outer membrane porin F OmpF, protein UspA1, protein EibA, protein Hia, cell surface protein Ail, chaperone SycD, protein LcrD, protein LcrG, protein LcrV, protein SycE, protein YopE, regulator protein TyeA, protein YopM, protein YopN, protein YopO, protein YopT, protein YopD, protease ClpP,

30 protein MyfA, protein FilA, and protein PsaA (*Yersinia enterocolitica*, Yersiniosis).

[0192] In embodiments wherein the infectious disease is influenza, the mRNA molecule may have a coding region encoding at least one antigenic peptide or protein derived from hemagglutinin (HA), neuraminidase (NA), nucleoprotein (NP), matrix protein 1 (M1), matrix protein 2 (M2), non-structural protein 1 (NS1), non-structural
5 protein 2 (NS2), nuclear export protein (NEP), polymerase acidic protein (PA), polymerase basic protein PB1, PB1-F2, or polymerase basic protein 2 (PB2) of an influenza virus or a fragment or variant thereof.

[0193] In certain embodiments, the coding region encodes at least one antigenic peptide or protein derived from hemagglutinin (HA) and/or neuraminidase (NA) of an
10 influenza virus or a fragment or variant thereof. The HA and/or NA may, independently, be derived from an influenza A virus or an influenza B virus or a fragment of either.

[0194] In embodiments wherein the infectious disease is influenza, the mRNA molecule may have a coding region encoding at least one antigenic peptide or protein derived from Spike (S) protein.

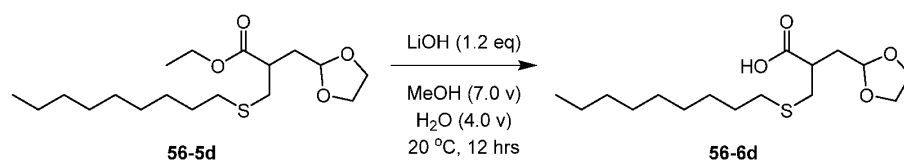
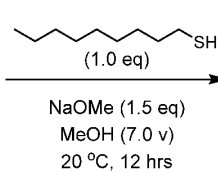
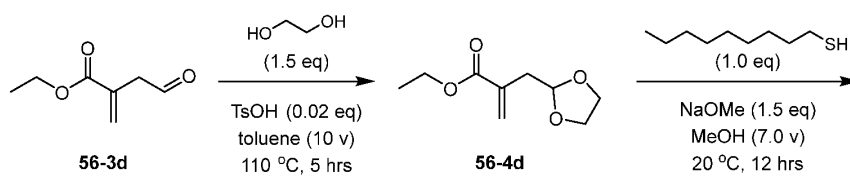
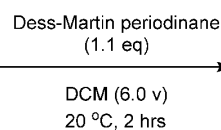
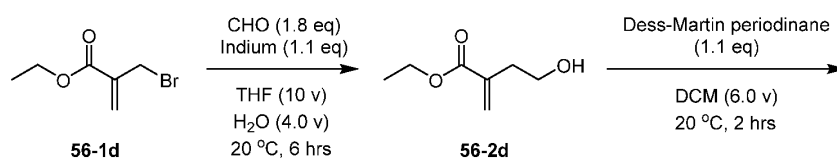
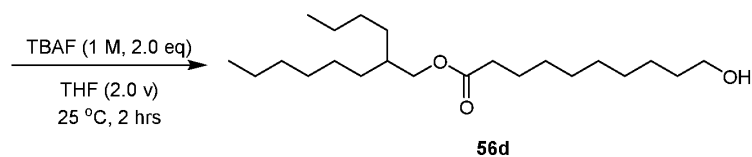
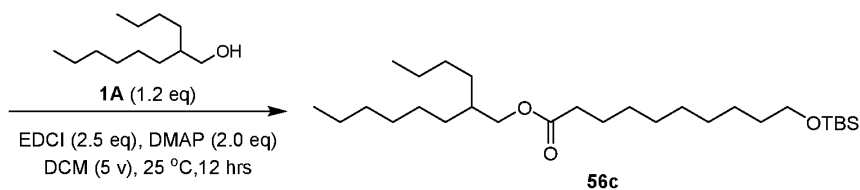
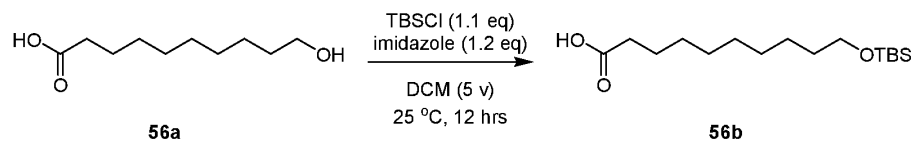
15

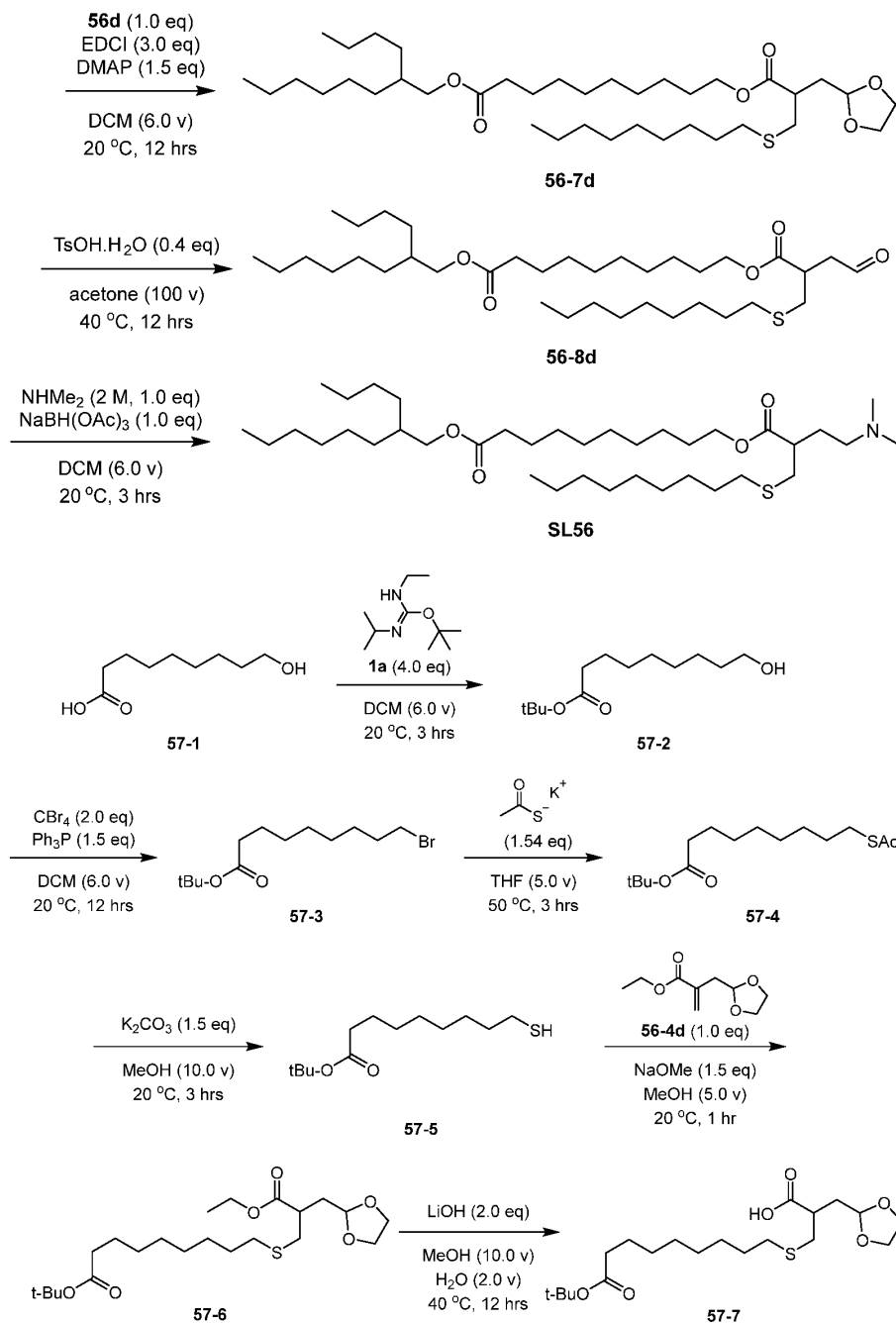
EXAMPLES

The following synthetic procedures describe approaches by which compounds of the present disclosure are made. Simple modifications of conditions or the nature of any particular substrate may be changed to achieve any compound within the scope of
20 Formula I and/or Formula II including straightforward synthesis modification to prepare the reversed tail ester (L^2) analog. Variations in the synthetic pathway are demonstrated to allow for variation in the end compound structure.

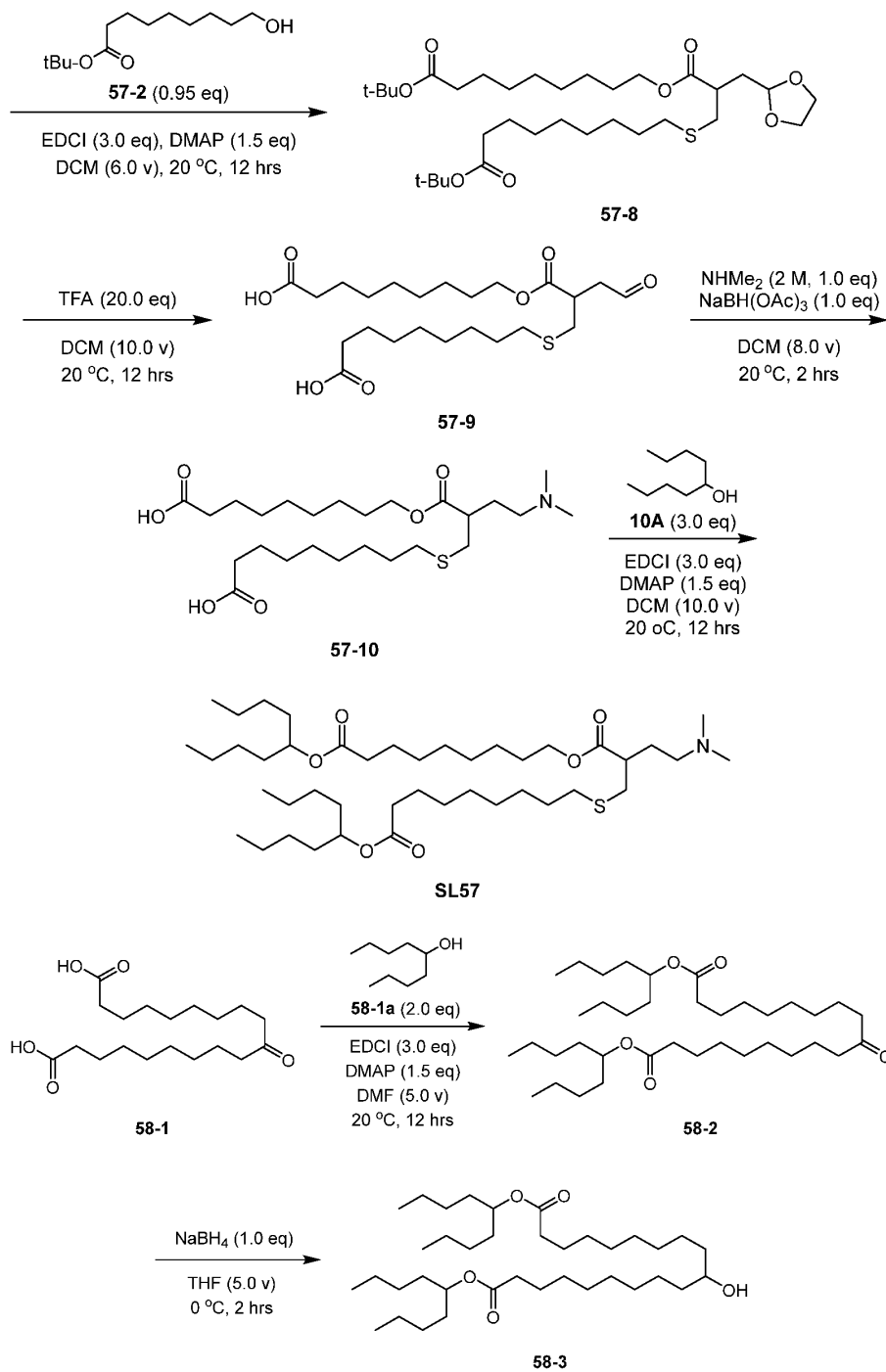
Example 1a: Synthetic Approach to Compounds SL56-SL61

100

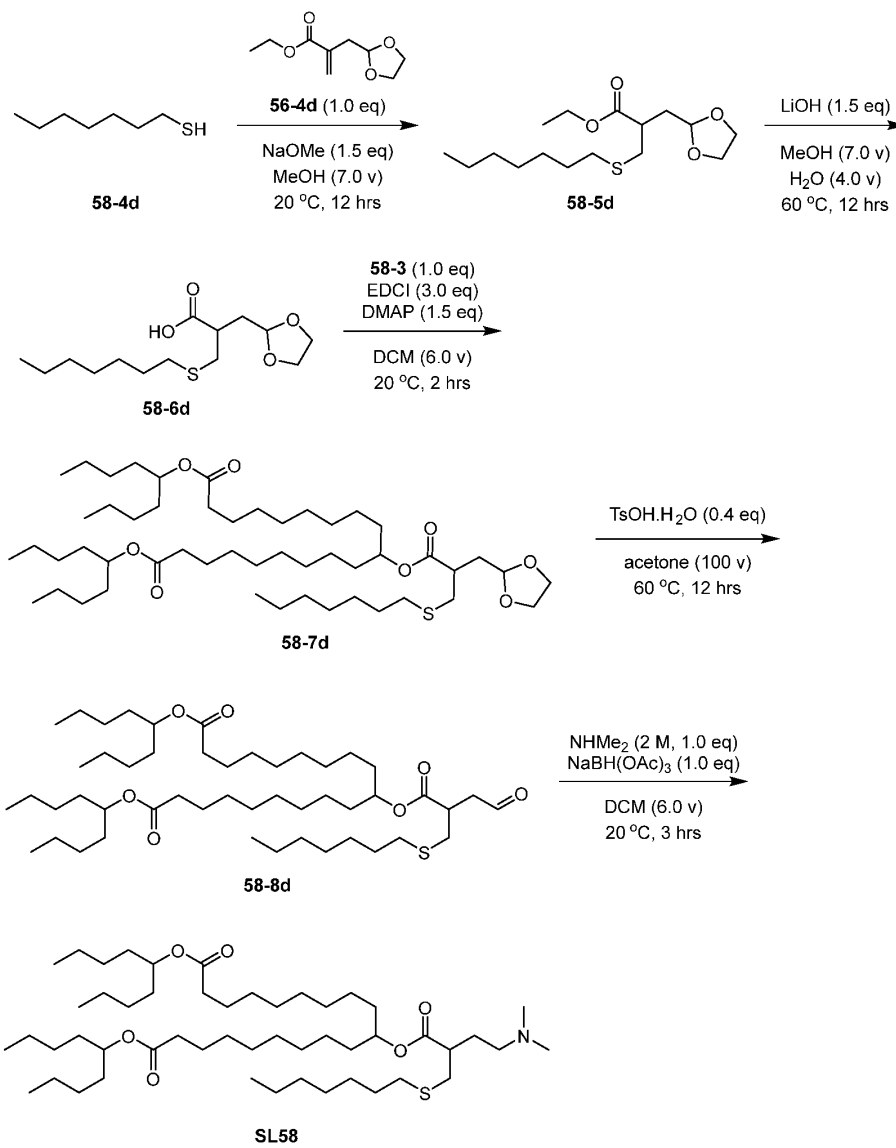




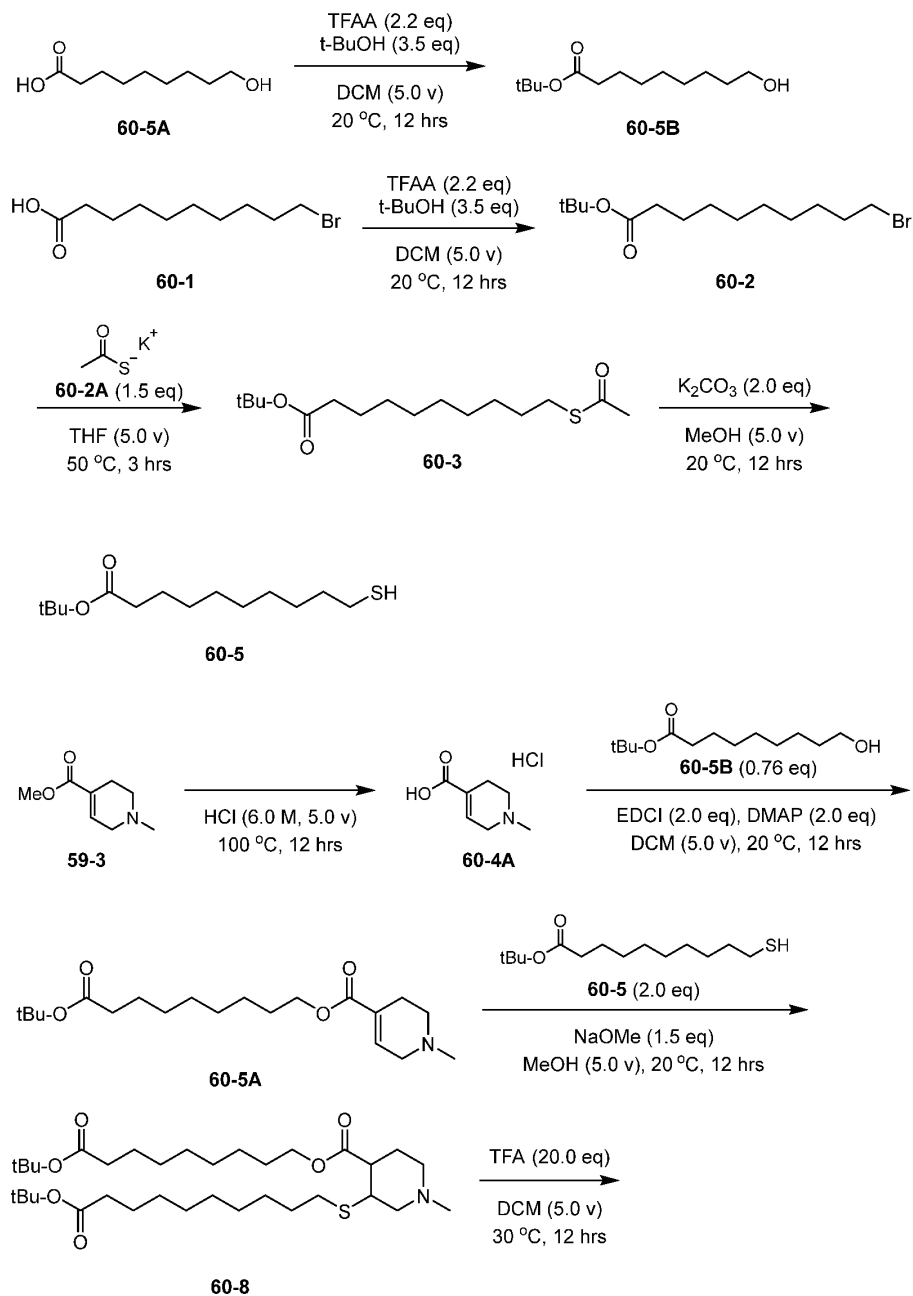
102



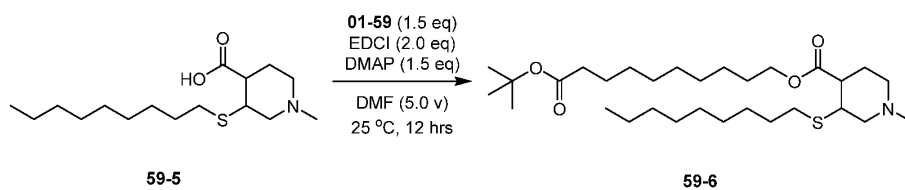
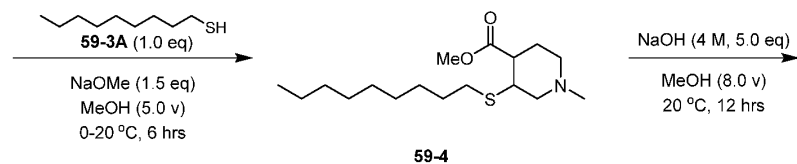
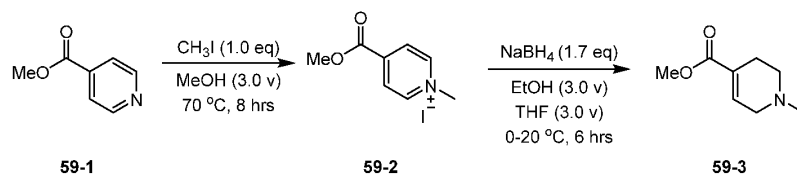
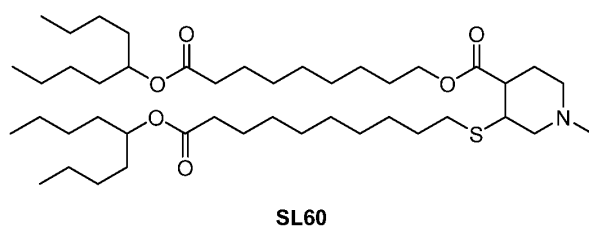
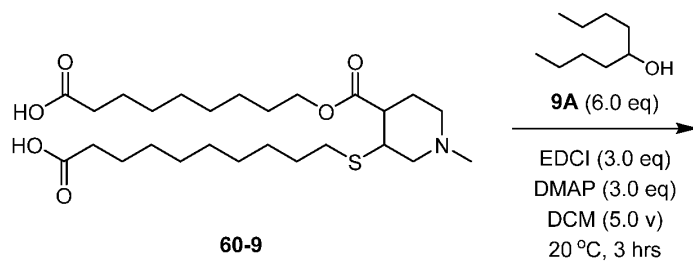
103



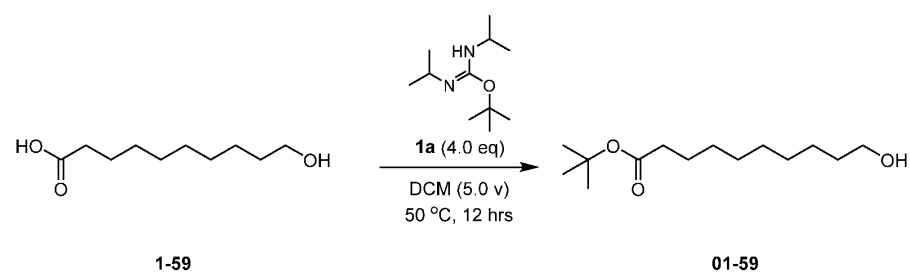
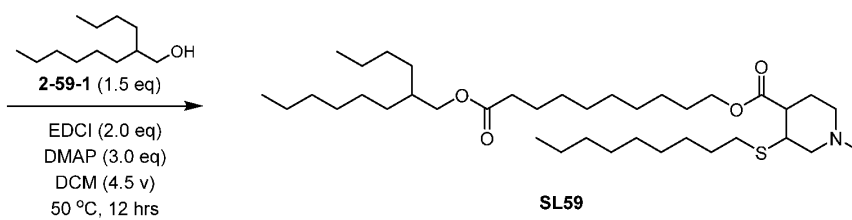
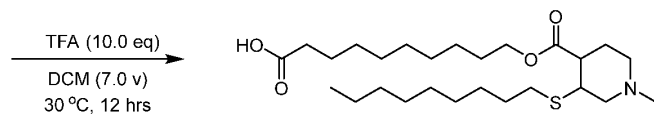
104



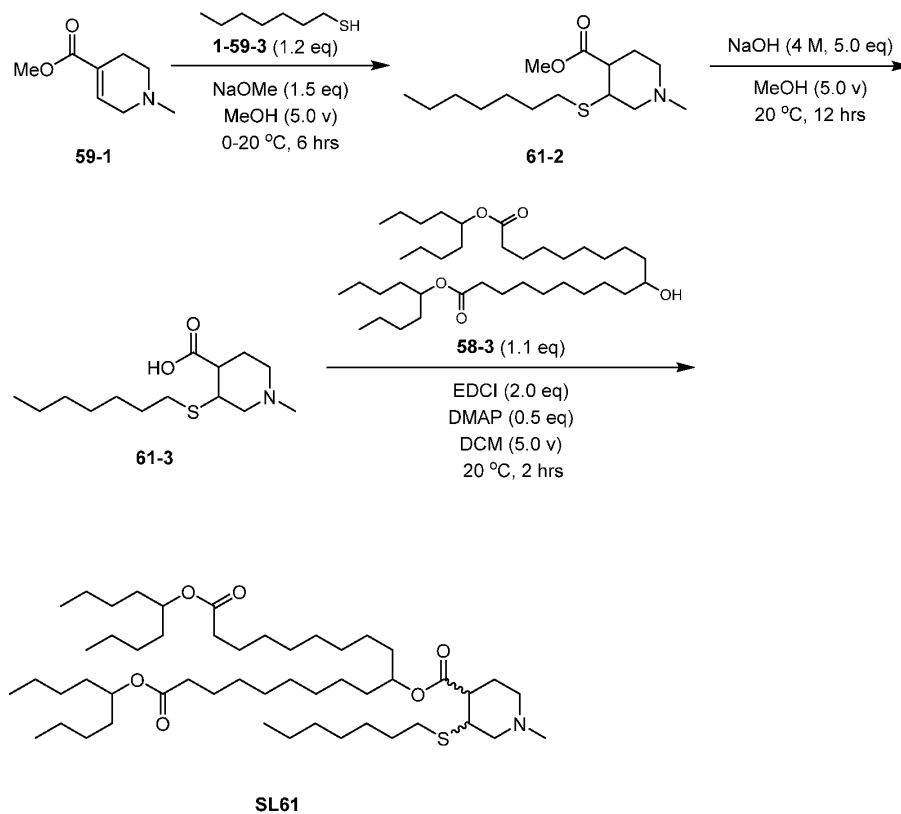
105



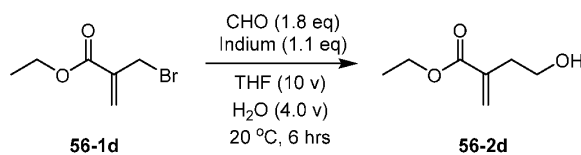
106



107



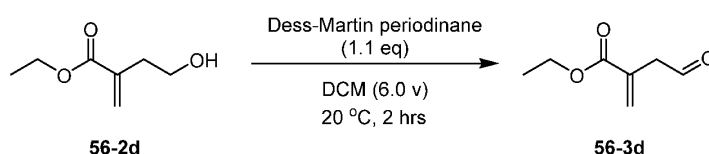
Preparation of compound 56-2d



- 5 To a solution of compound **56-1d** (75.0 g, 389 mmol, 1.0 *eq*) in H₂O (405 mL) and THF (700 mL) was added formaldehyde (56.8 g, 699 mmol, 52.1 mL, 37% purity, 1.8 *eq*) and Indium (49.1 g, 427 mmol, 6.72 mL, 1.1 *eq*). The mixture was stirred at 20 °C for 6 hrs. TLC (petroleum ether/ethyl acetate = 10/1, R_f of compound **56-1d** was 0.69, R_f of compound **56-2d** was 0.30) showed the solution was finished. The reaction
- 10 was quenched by H₂O (1500 mL), and the resulting solution was extracted with EtOAc

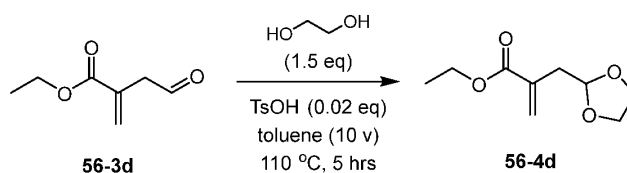
(1000 mL). The combined organic layers were washed with brine (500 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give compound **56-2d** (40.0 g, 277 mmol, 71.4% yield) as yellow oil. ¹H NMR: ET73307-106-P1A1 (400 MHz, CDCl₃) δ 6.25 (d, J = 1.2 Hz, 1H), 5.67 (d, J = 1.2 Hz, 1H), 4.20-4.26 (m, 2H), 3.77 (t, J = 6.4 Hz, 2H), 2.57-2.61 (m, 2H), 1.32 (t, J = 7.2 Hz, 3H).

Preparation of compound 56-3d



To a solution of compound **56-2d** (40.0 g, 278 mmol, 1.0 eq) in DCM (210 mL) was added Dess-Martin periodinane (129 g, 305 mmol, 94.6 mL, 1.1 eq). The mixture was stirred at 20 °C for 2 hrs. TLC (petroleum ether/ethyl acetate = 5/1, R_f of compound **56-2d** was 0.49, R_f of compound **56-3d** was 0.53) showed the solution was finished. The reaction was quenched by H₂O (200 mL), and the resulting solution was extracted with EtOAc (100 mL). The combined organic layers were washed with brine (50 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 500/1 to 20/1) to give compound **56-3d** (39.0 g, 274 mmol, 98.9% yield) as colorless oil. ¹H NMR: ET73307-110-P1A (400 MHz, CDCl₃) δ 9.72 (s, 1H), 6.42 (s, 1H), 5.73 (s, 1H), 4.24 (q, J = 7.2 Hz, 3H), 3.42 (s, 2H), 1.31 (t, J = 7.2 Hz, 3H).

Preparation of compound 56-4d

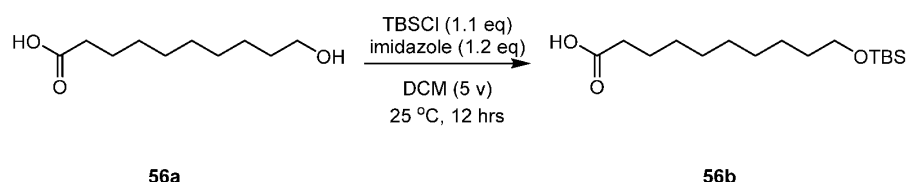


To a solution of compound **56-3d** (39.0 g, 274 mmol, 1.0 eq) in toluene (270 mL) was added 4-methylbenzenesulfonic acid (945 mg, 5.49 mmol, 0.02 eq) and ethylene glycol

(25.5 g, 412 mmol, 23 mL, 1.5 eq). The mixture was stirred at 110 °C for 5 hrs. TLC (petroleum ether/ethyl acetate = 5/1, R_f of compound **56-3d** was 0.47, R_f of compound **56-4d** was 0.52) showed the solution was finished. The reaction was quenched by H₂O (500 mL), and the resulting solution was extracted with EtOAc (300 mL). The combined organic layers were washed with brine (200 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give compound **56-4d** (37.0 g, 199 mmol, 72.4% yield) as colorless oil. ¹H NMR: ET73307-111-P1A (400 MHz, CDCl₃) δ 6.29 (s, 1H), 5.73 (s, 1H), 5.07 (t, J = 4.8 Hz, 1H), 4.22 (q, J = 7.2 Hz, 2H), 3.97-4.00 (m, 2H), 3.84-3.88 (m, 2H), 2.69 (d, J = 4.8 Hz, 2H), 1.31 (t, J = 7.2 Hz, 3H).

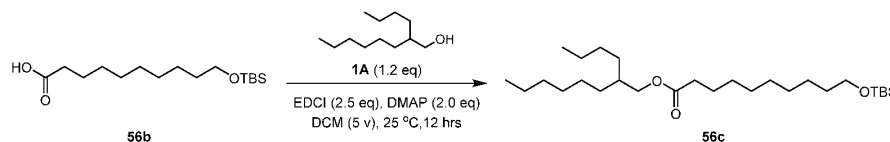
10

Preparation of compound 56b



To a solution of compound **56a** (30.0 g, 159 mmol, 1.0 eq) in DCM (150 mL) was added TBSCl (26.4 g, 175 mmol, 21.6 mL, 1.1 eq) and imidazole (13.0 g, 191 mmol, 1.2 eq). The solution was stirred at 25°C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 10/1, R_f of compound **56a** was 0.00, R_f of compound **56b** was 0.47) showed the starting material was consumed completely. The reaction was poured into water (20 mL). The organic phase was collected and the aqueous phase was extracted with DCM (20 mL x 2). The combined organic phase was washed with water (10 mL x 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 500/0 to 100/1) to give compound **56b** (26.1 g, 86.4 mmol, 54.2% yield) as colorless oil. ¹H NMR: ET74125-18-P1A (400 MHz, CDCl₃) δ 3.57-3.62 (t, J = 6.8 Hz, 2H), 2.33-2.38 (t, J = 7.2 Hz, 2H), 1.61-1.68 (m, 2H), 1.48-1.54 (m, 2H), 1.26-1.37 (m, 10H), 0.89-0.91 (m, 9H), 0.03-0.07 (m, 6H).

Preparation of compound 56c

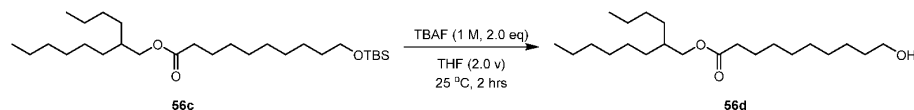


To a solution of compound **56b** (26.0 g, 85.9 mmol, 1.0 *eq*) in DCM (150 mL) was added EDCI (41.2 g, 215 mmol, 2.5 *eq*), DMAP (21.0 g, 172 mmol, 2.0 *eq*) and compound **1A** (19.2 g, 103 mmol, 1.2 *eq*). The solution was stirred at 25 °C for 12 hrs.

- 5 TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **56b** was 0.30, R_f of compound **56c** was 0.60) showed the starting material was consumed completely. The reaction solution was poured into NH_4Cl (aq., Sat., 20 mL). The solution was extracted with DCM (20 mL) three times. The organic layer was dried over sodium sulfate and concentrated under reduced pressure to give a residue. The residue was
- 10 purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 300/1 to 20/1) to give compound **56c** (37.1 g, 78.8 mmol, 91.7% yield) as colorless oil. ^1H NMR: ET74125-21-P1A (400 MHz, CDCl_3) δ 3.91-3.93 (m, 2H), 3.52-3.57 (t, $J = 6.8$ Hz, 2H), 2.23-2.28 (t, $J = 7.2$ Hz, 2H), 1.51-1.61 (m, 3H), 1.41-1.48 (m, 2H), 1.39-1.41 (s, 1H), 1.18-1.28 (m, 25H), 0.82-0.87 (m, 15H), 0.00-0.03 (m, 6H).

15

Preparation of compound **56d**

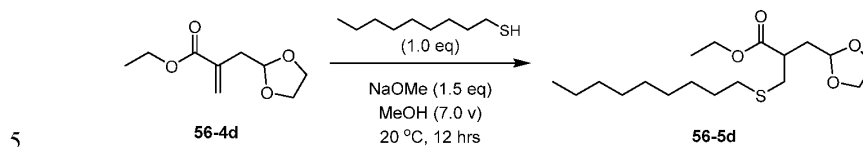


To a solution of compound **56c** (37.1 g, 78.8 mmol, 1.0 *eq*) in THF (80 mL) was added TBAF (1 M, 158 mL, 2.0 *eq*). The solution was stirred at 25 °C for 2 hrs.

- 20 TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **56c** was 0.60, R_f of compound **56d** was 0.30) showed the starting material was consumed completely. The reaction solution was poured into HCl (1 M, 100 mL). The solution was extracted with EtOAc (100 mL) three times. The organic was concentrated under vacuum to give a compound **56d** (20.0 g, 56.1 mmol, 71.2% yield) was obtained as colorless oil.
- 25 Compound **56d** (20.0 g, 56.1 mmol, 71.2% yield) was used next step without purification. ^1H NMR: ET74125-23-P1A (400 MHz, CDCl_3) δ 3.96-3.99 (m, 2H), 3.62-

3.68 (m, 2H), 2.28-2.33 (t, $J = 7.6$ Hz, 2H), 1.56-1.67 (m, 5H), 1.25-1.34 (m, 26H), 0.87-0.92 (m, 6H).

Preparation of compound 56-5d

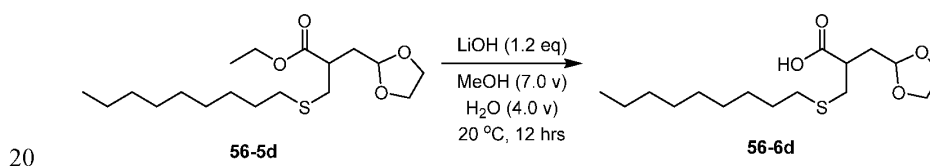


To a solution of compound **56-4d** (7.00 g, 37.6 mmol, 1.0 eq) in MeOH (70 mL) was added NaOMe (3.05 g, 56.4 mmol, 1.5 eq) and nonane-1-thiol (6.03 g, 37.6 mmol, 1.0 eq). The mixture was stirred at 20 °C for 12 hrs. TLC (petroleum ether/ethyl acetate = 5/1, R_f of compound **56-4d** was 0.43, R_f of compound **56-5d** was 0.50) showed the solution was finished. The reaction was quenched by H₂O (200 mL), and the resulting solution was extracted with EtOAc (100 mL). The combined organic layers were washed with brine (80 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give compound **56-5d** (11.0 g, 31.7 mmol, 84.4% yield) was obtained as colorless oil. ¹H NMR: ET73307-112-P1A (400 MHz, CDCl₃) δ 4.94-4.96 (m, 1H), 4.14-4.21 (m, 1H), 3.92-3.99 (m, 2H), 3.81-3.87 (m, 2H), 3.72 (s, 1H), 2.76-2.84 (m, 2H), 2.64-2.71 (m, 1H), 2.51 (t, $J = 7.2$ Hz, 2H), 2.10-2.15 (m, 1H), 1.95-2.00 (m, 1H), 1.53-1.61 (m, 2H), 1.25-1.38 (m, 14H), 0.88 (t, $J = 6.4$ Hz, 3H).

10

15

Preparation of compound 56-6d

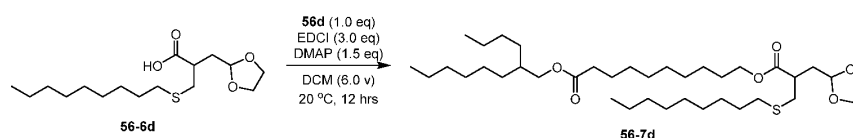


To a solution of compound **56-5d** (11.0 g, 31.7 mmol, 1.0 eq) in MeOH (60 mL) was added LiOH.H₂O (2.00 g, 47.6 mmol, 1.5 eq). The mixture was stirred at 60 °C for 12 hrs. TLC (petroleum ether/ethyl acetate = 5/1, R_f of compound **56-5d** was 0.56, R_f of compound **56-6d** was 0.14) showed the solution was finished. The reaction solution was poured into HCl (2 M, 100 mL). The aqueous layer was extracted with Ethyl acetate (80

25

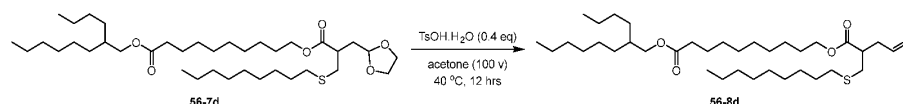
mL). The combined organic layer was washed with brine (100 mL), dried over Na₂SO₄, filtered and concentrated under vacuum to give compound **56-6d** (8.00 g, 25.1 mmol, 79.1% yield) as colorless oil. ¹H NMR: ET73307-115-P1A (400 MHz, CDCl₃) δ 5.01 (t, *J* = 4.0 Hz, 1H), 3.94-4.02 (m, 2H), 3.84-3.90 (m, 2H), 2.82-2.88 (m, 2H), 2.67-2.73 (m, 1H), 2.54 (t, *J* = 7.2 Hz, 2H), 2.14-2.21 (m, 1H), 2.02-2.08 (m, 1H), 1.54-1.62 (m, 2H), 1.25-1.39 (m, 12H), 0.89 (t, *J* = 6.4 Hz, 3H).

Preparation of compound **56-7d**



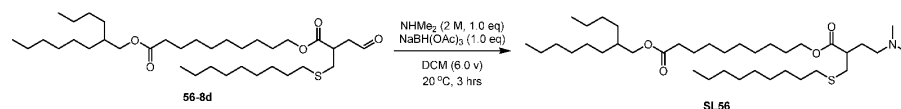
- 10 To a solution of compound **56-6d** (7.00 g, 22.0 mmol, 1.0 *eq*) in DCM (42.0 mL) were added compound **56d** (7.84 g, 22.0 mmol, 1.0 *eq*), EDCI (12.6 g, 65.9 mmol, 3.0 *eq*) and DMAP (4.03 g, 33.0 mmol, 1.5 *eq*). The mixture was stirred at 20 °C for 12 hrs. TLC (petroleum ether/ethyl acetate = 5/1, *R_f* of compound **56-6d** was 0.37, *R_f* of compound **56-7d** was 0.73) showed the solution was finished. The reaction was
- 15 quenched by H₂O (100 mL), and the resulting solution was extracted with EtOAc (80 mL). The combined organic layers were washed with brine (50 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 500/1 to 20/1) to give compound **56-7d** (8.00 g, 12.2 mmol, 55.4% yield) as colorless oil. ¹H
- 20 NMR: ET73307-118-P1A (400 MHz, CDCl₃) δ 4.95 (t, *J* = 4.0 Hz, 1H), 4.10 (t, *J* = 6.8 Hz, 2H), 3.92-3.98 (m, 4H), 3.84-3.87 (m, 2H), 2.76-2.81 (m, 2H), 2.67-2.71 (m, 1H), 2.51 (t, *J* = 7.2 Hz, 2H), 2.30 (t, *J* = 7.6 Hz, 2H), 2.10-2.17 (m, 1H), 1.94-1.99 (m, 1H), 1.53-1.67 (m, 8H), 1.28-1.39 (m, 38H), 0.87-0.92 (m, 9H).

25 Preparation of compound **56-8d**



To a solution of compound **56-7d** (7.00 g, 10.7 mmol, 1.0 *eq*) in ACETONE (700 mL) was added TsOH.H₂O (811 mg, 4.26 mmol, 0.4 *eq*). The mixture was stirred at 40 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **56-7d** was 0.54, R_f of compound **56-8d** was 0.59) showed the solution was finished. The reaction was quenched by H₂O (1000 mL), and the resulting solution was extracted with EtOAc (500 mL). The combined organic layers were washed with brine (300 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give compound **56-8d** (7 g, crude) as yellow oil.

10 Preparation of compound **SL56**

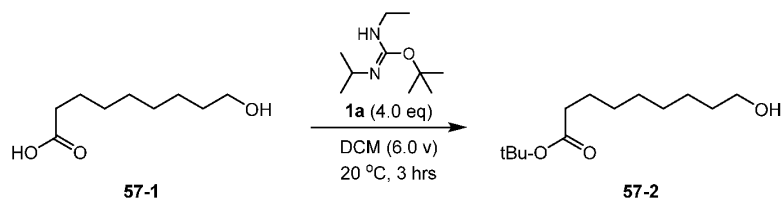


To a solution of compound **56-8d** (7.00 g, 11.4 mmol, 1.0 *eq*) in DCM (70.0 mL) was added NaBH(OAc)₃ (2.42 g, 11.4 mmol, 1.0 *eq*) and NHMe₂ (2 M, 5.71 mL, 1.0 *eq*). The mixture was stirred at 20 °C for 3 hrs. TLC (Dichloromethane/Methanol = 10/1, R_f of compound **56-8d** was 0.80, R_f of compound **SL56** was 0.40) showed the solution was finished. The reaction was quenched by H₂O (200 mL), and the resulting solution was extracted with EtOAc (200 mL). The combined organic layers were washed with brine (100 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, DCM/MeOH = 500/1 to 20/1). **SL 56** (1.20 g, 1.87 mmol, 16.4% yield) was obtained as a light-yellow oil. LCMS: ET73307-127-P1A (M+H⁺): 642.5. ¹H NMR: ET73307-127-P1A1 (400 MHz, CDCl₃) δ 4.10 (t, *J* = 6.8 Hz, 2H), 3.98 (d, *J* = 5.6 Hz, 12H), 2.76-2.82 (m, 1H), 2.62-2.67 (m, 2H), 2.51 (t, *J* = 7.2 Hz, 2H), 2.24-2.32 (m, 4H), 2.21 (s, 6H), 1.76-7.84 (m, 2H), 1.53-1.67 (m, 7H), 1.27-1.36 (m, 38H), 0.87-0.90 (m, 9H).

25

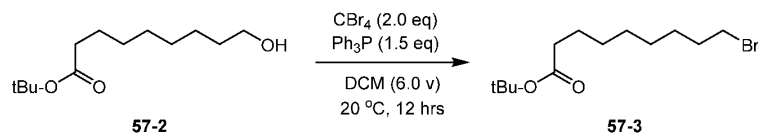
Preparation of compound **57-2**

114



To a solution of compound **57-1** (25.0 g, 144 mmol, 1.0 *eq*) in DCM (150 mL) was added compound **1a** (115 g, 574 mmol, 4.0 *eq*). The solution was stirred at 20 °C for 3 hrs. TLC (Dichloromethane/Methanol = 8/1, R_f of compound **57-1** was 0.41, R_f of compound **57-2** was 0.54) showed the reaction was completed. The reaction mixture was filtered and the filter cake was washed with 100 mL of DCM. The combined filtrates were concentrated to dryness to give yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 50/1 to 5/1) to give colorless oil. Compound **57-2** (25.0 g, 109 mmol, 75.6% yield) was obtained as yellow oil. ¹H NMR: ET73270-114-P1A (400 MHz, CDCl₃) δ 3.64 (t, J = 6.6 Hz, 2H), 2.20 (t, J = 6.6 Hz, 2H), 1.55-1.59 (m, 4H), 1.45 (s, 9H), 1.18-1.32 (m, 8H).

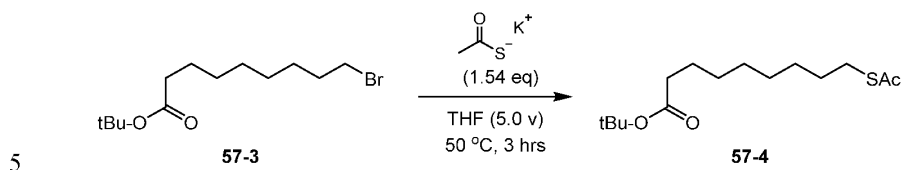
Preparation of compound 57-3



To a solution of compound **57-2** (17.0 g, 73.8 mmol, 1.0 *eq*) in DCM (102 mL) was added CBr₄ (48.9 g, 148 mmol, 2.0 *eq*) and Ph₃P (29.0 g, 111 mmol, 1.5 *eq*). The solution was stirred at 20 °C for 12 hrs. TLC (Dichloromethane/Methanol = 8/1, R_f of compound **57-2** was 0.57, R_f of compound **57-3** was 0.90) showed the reaction was completed. The reaction was poured into aqueous solution of H₂O (aq., 150 mL). The resulting solution was extracted with DCM (100 mL x 2). Dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 1/0 to 50/1) to give yellow oil. Compound **57-3** (20.0 g, 68.2 mmol, 92.4% yield) was obtained as yellow oil. ¹H NMR: ET73270-119-P1A (400 MHz, CDCl₃) δ 3.44 (t, J =

6.8 Hz, 2H), 2.23 (t, $J = 7.6$ Hz, 2H), 1.85-1.90 (m, 2H), 1.61 (t, $J = 7.0$ Hz, 2H), 1.48 (s, 11H), 1.34 (s, 6H).

Preparation of compound 57-4

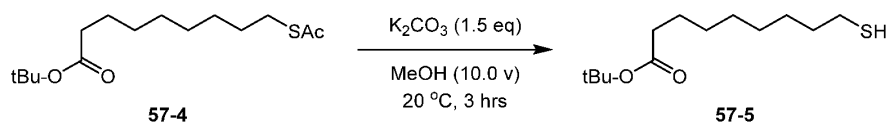


To a solution of compound **57-3** (20.0 g, 68.2 mmol, 1.0 eq) in THF (100 mL) was added $\text{C}_2\text{H}_5\text{OSK}$ (12.0 g, 105 mmol, 1.54 eq). The suspension was stirred at 50 °C for 3 hrs. TLC (Petroleum ether/Ethyl acetate = 8/1, R_f of compound **57-3** was 0.61, R_f of compound **57-4** was 0.67) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 100 mL). The resulting solution was extracted with EtOAc (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. Compound **57-4** (19 g, 65.87 mmol, 96.58% yield) was obtained as yellow oil. ^1H NMR: ET73270-120-P1A (400 MHz, CDCl_3) δ 2.86 (t, $J = 7.4$ Hz, 2H), 2.33 (s, 3H), 2.20 (t, $J = 7.4$ Hz, 2H), 1.55-1.59 (m, 4H), 1.45 (s, 9H), 1.30-1.43 (m, 8H).

10

15

Preparation of compound 57-5



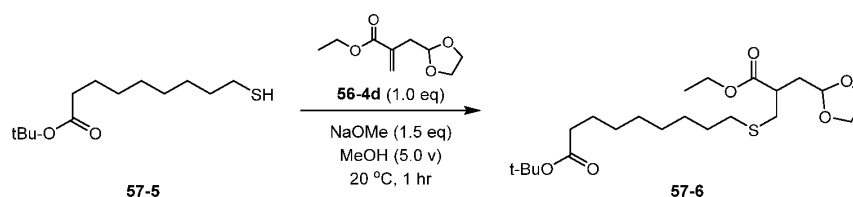
To a solution of compound **57-4** (19.0 g, 65.9 mmol, 1.0 eq) in MeOH (200 mL) was added K_2CO_3 (13.7 g, 98.8 mmol, 1.5 eq). The suspension was stirred at 20 °C for 3 hrs. TLC (Petroleum ether/Ethyl acetate = 8/1, R_f of compound **57-4** was 0.59, R_f of compound **57-5** was 0.68) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 100 mL). The resulting solution was extracted with EtOAc (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product

20

25

as yellow oil. Compound **57-5** (16.0 g, 64.9 mmol, 98.6% yield) was obtained as yellow oil. ^1H NMR: ET73270-120-P1A (400 MHz, CDCl_3) δ 2.50-2.55 (m, 2H), 2.20 (t, J = 7.6 Hz, 2H), 1.57-1.61 (m, 4H), 1.45 (s, 9H), 1.33-1.35 (m, 2H), 1.26-1.31 (m, 6H).

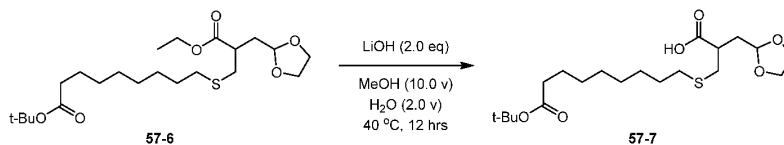
5 Preparation of compound 57-6



To a solution of compound **56-5** (11.5 g, 61.7 mmol, 0.95 *eq*) in MeOH (80 mL) was added NaOMe (5.26 g, 97.4 mmol, 1.5 *eq*) and compound **57-4d** (16.00 g, 64.93 mmol, 1.00 *eq*). The solution was stirred at 20 °C for 1 hr. TLC (Petroleum ether/Ethyl acetate = 1/1, R_f of compound **57-5** was 0.26, R_f of compound **57-6** was 0.32) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 100 mL). The resulting solution was extracted with EtOAc (100 mL x 2). Dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 80/1 to 10/1) to give yellow oil. Compound **57-6** (25.0 g, 57.8 mmol, 89.0% yield) was obtained as yellow oil. ^1H NMR: ET73270-124-P1A (400 MHz, CDCl_3) δ 4.87-4.88 (m, H), 3.87-3.89 (m, 2H), 3.76-3.77 (m, 2H), 2.71-2.73 (m, 2H), 2.61 (m, 1H), 2.42 (t, J = 7.2 Hz, 2H), 2.13 (t, J = 7.6 Hz, 2H), 2.10-2.11, (m, 1H), 1.95-1.97 (m, 1H), 1.49-1.52 (m, 4H), 1.37 (s, 9H), 1.21-1.22 (m, 2H), 1.17-1.21 (m, 9H).

20

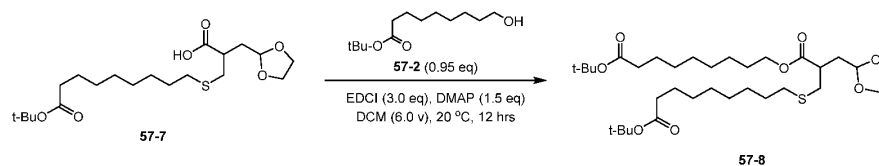
Preparation of compound 57-7



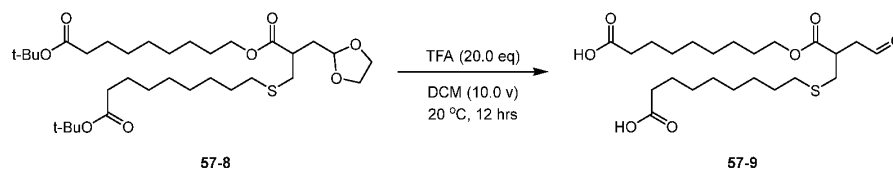
To a solution of compound **57-6** (20.0 g, 46.2 mmol, 1.0 *eq*) in MeOH (200 mL) and H_2O (40 mL) was added $\text{LiOH}\cdot\text{H}_2\text{O}$ (3.88 g, 92.5 mmol, 2.0 *eq*). The

suspension was stirred at 40 °C for 12 hrs. TLC (Dichloromethane/Methanol = 8/1, R_f of compound **57-6** was 0.72, R_f of compound **57-7** was 0.52) showed the reaction was completed. The reaction was poured into aqueous solution of HCl (aq., 1M, 200 mL). The resulting solution was extracted with EtOAc (200 mL x 2). Dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 80/1 to 10/1) to give yellow oil. Compound **57-7** (7.30 g, 18.0 mmol, 39.0% yield) was obtained as yellow oil. ¹H NMR: ET73270-126-P1A (400 MHz, CDCl₃) δ 5.31 (t, *J* = 7.2 Hz, 1H), 3.98-4.00 (m, 2H), 3.87-3.88 (m, 2H), 2.84-2.86 (m, 2H), 2.70-2.71 (m, 1H), 2.52 (t, *J* = 6.8 Hz, 2H), 2.21 (t, *J* = 7.6 Hz, 2H), 1.56-1.60 (m, 4H), 1.45 (s, 9H), 1.29-1.30 (m, 8H).

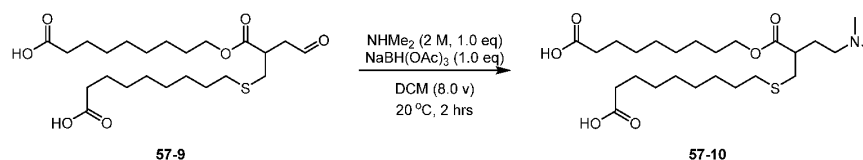
Preparation of compound 57-8



15 To a solution of compound **57-7** (7.30 g, 18.0 mmol, 1.0 *eq*) in DCM (43.8 mL) was added EDCI (10.4 g, 54.1 mmol, 3.0 *eq*), DMAP (3.31 g, 27.1 mmol, 1.5 *eq*) and compound **57-2** (3.95 g, 17.1 mmol, 0.95 *eq*). The solution was stirred at 20 °C for 12 hrs. TLC (Dichloromethane/Methanol, *R_f* of compound **57-7** was 0.34, *R_f* of compound **57-8** was 0.91) showed the reaction was completed. The reaction was poured into
20 H₂O (100 mL). The resulting solution was extracted with DCM (100 mL x 2). Dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 1/0 to 20/1) to give yellow oil. Compound **57-8** (3.80 g, 9.73 mmol, 53.9% yield) was obtained as yellow oil. ¹H NMR: ET73270-37-P1A (400 MHz, CDCl₃)
25 δ 4.94 (t, *J* = 4.2 Hz, 1H), 4.08-4.14 (m, 2H), 3.94-3.96 (m, 2H), 3.83-3.84 (m, 2H), 2.78-2.80 (m, 2H), 2.68-2.68 (m, 1H), 2.51 (t, *J* = 7.6 Hz, 2H), 2.20 (t, *J* = 7.6 Hz, 4H), 2.10-2.19 (m, 1H), 2.09-2.10 (m, 1H), 1.58-1.60 (m, 8H), 1.44 (s, 18H), 1.26-1.31 (m, 16H).

Preparation of compound 57-9

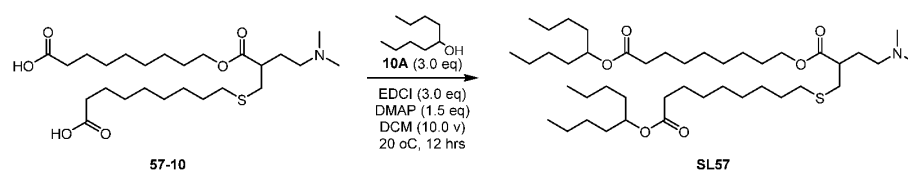
To a solution of compound **57-8** (6.00 g, 9.73 mmol, 1.0 eq) in DCM (60 mL) was added TFA (22.2 g, 195 mmol, 14.5 mL, 20.0 eq). The solution was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **57-8** was 0.42, R_f of compound **57-9** was 0.09) showed the reaction was completed. Concentrated the reaction under vacuum to give the crude product as yellow oil. Compound **57-9** (5.00 g, crude) was obtained as yellow oil. ^1H NMR: ET73270-133-P1A (400 MHz, CDCl_3) δ 9.80 (s, 1H), 4.11 (t, $J = 6.6$ Hz, 2H), 3.13-3.14 (m, 1H), 2.87-3.95 (m, 3H), 2.67-2.68 (m, 1H), 2.50 (t, $J = 7.4$ Hz, 2H), 2.36 (t, $J = 7.2$ Hz, 4H), 1.55-1.66 (m, 6H), 1.33 (s, 18H).

Preparation of compound 57-10

To a solution of compound **57-9** (5.00 g, 10.85 mmol, 1.0 eq) in DCM (40 mL) was added NHMe_2 (2M, 5.43 mL, 1.0 eq). The solution was stirred at 20 °C for 1 hr. Then added NaBH(OAc)_3 (2.30 g, 10.9 mmol, 1.0 eq). The solution was stirred at 20 °C for 1 hr. TLC (Dichloromethane/Methanol = 8/1, R_f of compound **57-9** was 0.34, R_f of compound **57-10** was 0.02) showed the reaction was completed. The reaction was poured into aqueous solution of NaHCO_3 (aq. 50 mL). The resulting solution was extracted with DCM (100 mL x 2). Then acidified the aqueous phase with 2M HCl to pH=3. The aqueous phase was extracted with EtOAc (100 mL x 2). Dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. Compound **57-10** (4.30 g, 8.78 mmol, 80.9% yield) was obtained as colorless oil.

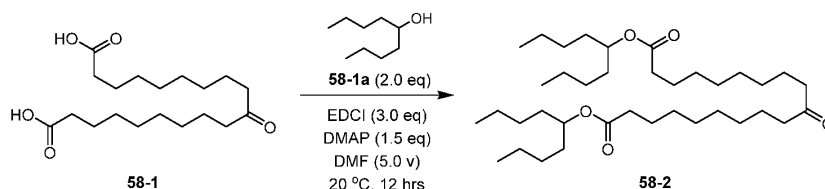
^1H NMR: ET73270-136-P1A (400 MHz, CDCl_3) δ 4.10-4.15 (m, 2H), 3.08-3.24 (m, 1H), 2.98-3.18 (m, 1H), 2.85 (s, 6H), 2.70-2.73 (m, 2H), 2.52 (t, $J = 7.2$ Hz, 2H), 2.35 (t, $J = 7.2$ Hz, 4H), 2.11-2.20 (m, 2H), 1.55-1.65 (m, 8H), 1.26-1.34 (m, 16H).

5 Preparation of **SL57**



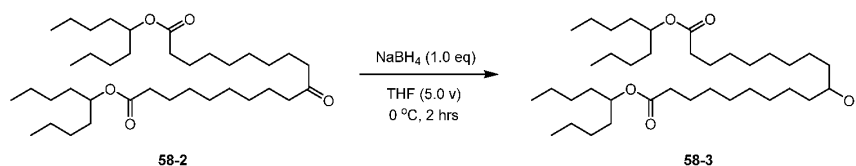
To a solution of compound **57-10** (4.30 g, 8.78 mmol, 1.0 eq) in DCM (43 mL) was added EDCI (5.05 g, 26.3 mmol, 3.0 eq), DMAP (1.61 g, 13.17 mmol, 1.5 eq) and compound **10A** (3.80 g, 26.3 mmol, 3.0 eq). The solution was stirred at 20 °C for 12 hrs. LCMS showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 100 mL). The resulting solution was extracted with DCM (100 mL x 2). Dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO_2 , Dichloromethane/Methanol = 200/1 to 50/1) to give light yellow oil. **SL57** (1.00 g, 1.34 mmol, 15.3% yield, 99.7% purity) was obtained as a light-yellow oil. LCMS: ET73270-137-P1A, RT = 0.791, $\text{M}+\text{H}^+ = 742.59$. ^1H NMR: ET73270-137-P1A (400 MHz, CDCl_3) δ 4.88 (t, $J = 6.4$ Hz, 2H), 4.10 (t, $J = 6.6$ Hz, 2H), 2.77-2.78 (m, 1H), 2.63-2.65 (m, 2H), 2.49 (t, $J = 7.6$ Hz, 2H), 2.29 (t, $J = 7.4$ Hz, 6H), 2.20 (s, 6H), 1.76-1.81 (m, 2H), 1.53-1.62 (m, 8H), 1.51-1.53 (m, 8H), 1.27-1.35 (m, 32H), 0.89 (t, $J = 6.8$ Hz, 12H).

Preparation of compound **58-2**



To a solution of compound **58-1** (18.0 g, 52.6 mmol, 1.0 *eq*) in DCM (200 mL) was added compound **58-1a** (15.2 g, 105 mmol, 2.0 *eq*), EDCI (30.2 g, 158 mmol, 3.0 *eq*) and DMAP (9.63 g, 78.8 mmol, 1.5 *eq*). The mixture was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, *R_f* of compound **58-1** was 0.31, *R_f* of compound **58-2** was 0.61) showed the solution was finished. The reaction was quenched by H₂O (500 mL), and the resulting solution was extracted with EtOAc (300 mL). The combined organic layers were washed with brine (200 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 500/1 to 20/1). Compound **58-2** (15.0 g, 25.2 mmol, 47.9% yield) was obtained as yellow oil.

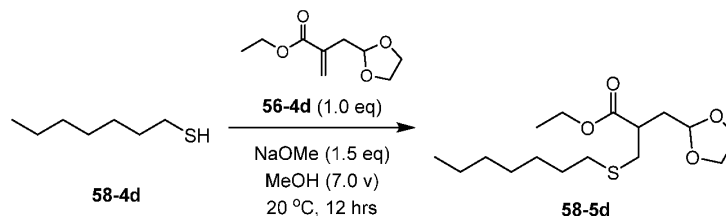
Preparation of compound 58-3



To a solution of compound **58-2** (15.0 g, 25.2 mmol, 1.0 *eq*) in THF (80 mL) was added NaBH₄ (954 mg, 25.2 mmol, 1.0 *eq*). The mixture was stirred at 0 °C for 2 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, *R_f* of compound **58-2** was 0.58, *R_f* of compound **58-3** was 0.53) showed the solution was finished. The reaction was quenched by H₂O (200 mL), and the resulting solution was extracted with EtOAc (100 mL). The combined organic layers were washed with brine (50 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 200/1 to 20/1). Compound **58-3** (13.0 g, 21.8 mmol, 86.4% yield) was obtained as colorless oil. ¹H NMR: ET73307-117-P1A (400 MHz, CDCl₃) δ 4.85-4.91 (m, 2H), 3.56-3.59 (m, 1H), 2.29 (t, *J* = 7.6 Hz, 4H), 1.61-1.64 (m, 4H), 1.49-1.55 (m, 8H), 1.41-1.45 (m, 4H), 1.24-1.36 (m, 36H), 0.89 (t, *J* = 6.8 Hz, 12H).

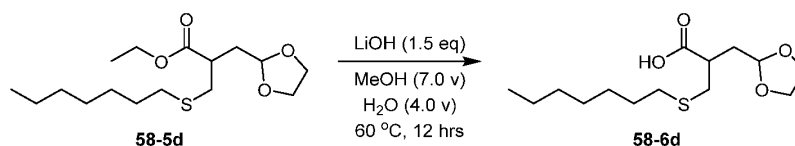
Preparation of compound 58-5d

121



To a solution of compound **56-4d** (7.00 g, 37.6 mmol, 1.0 *eq*) in MeOH (70 mL) was added NaOMe (3.05 g, 56.4 mmol, 1.5 *eq*) and compound **58-4d** (4.97 g, 37.6 mmol, 5.89 mL, 1.0 *eq*). The mixture was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **56-4d** was 0.43, R_f of compound **58-5d** was 0.50) showed the solution was finished. The reaction was quenched by H₂O (300 mL), and the resulting solution was extracted with EtOAc (200 mL). The combined organic layers were washed with brine (100 mL), dried over Na₂SO₄ and concentrated under reduced pressure to give product. Compound **58-5d** (10.0 g, 31.4 mmol, 83.5% yield) was obtained as colorless oil. ¹H NMR: ET73307-114-P1A (400 MHz, CDCl₃) δ 4.94-4.96 (m, 1H), 3.92-3.99 (m, 2H), 3.81-3.87 (m, 2H), 3.71 (s, 2H), 2.77-2.84 (m, 2H), 2.66-2.71 (m, 1H), 2.51 (t, J = 7.6 Hz, 2H), 2.10-2.17 (m, 1H), 1.95-2.01 (m, 1H), 1.53-1.61 (m, 2H), 1.25-1.38 (m, 10H), 0.89 (t, J = 6.8 Hz, 3H).

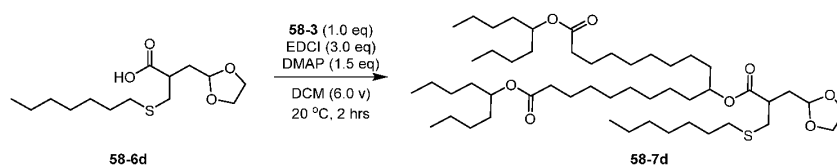
15 Preparation of compound **58-6d**



To a solution of compound **58-5d** (10.0 g, 31.4 mmol, 1.0 *eq*) in MeOH (60 mL) was added LiOH.H₂O (1.98 g, 47.1 mmol, 1.5 *eq*). The mixture was stirred at 60 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 5/1, R_f of compound **58-5d** was 0.56, R_f of compound **58-6d** was 0.14) showed the solution was finished. The reaction solution was poured into HCl (2 M, 100 mL). The aqueous layer was extracted with Ethyl acetate (80 mL). The combined organic layer was washed with brine (100 mL), dried over Na₂SO₄, filtered and concentrated under vacuum to give product. Compound **58-6d** (9.00 g, 31.0 mmol, 98.7% yield) was obtained as colorless oil. ¹H NMR: ET73307-116-P1A (400

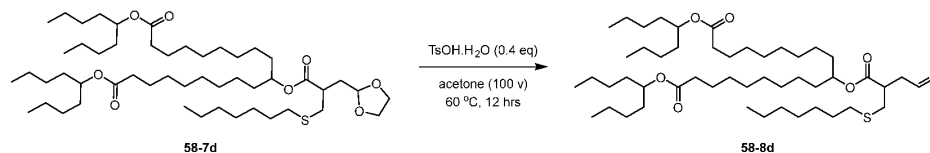
MHz, CDCl₃) δ 5.01 (t, J = 4.0 Hz, 1H), 3.96-4.02 (m, 2H), 3.84-3.90 (m, 2H), 2.81-2.88 (m, 2H), 2.67-2.73 (m, 1H), 2.54 (t, J = 7.2 Hz, 2H), 2.14-2.21 (m, 1H), 2.01-2.07 (m, 1H), 1.54-1.62 (m, 2H), 1.25-1.39 (m, 10H), 0.89 (t, J = 6.4 Hz, 3H).

5 Preparation of compound 58-7d



To a solution of compound **58-6d** (4.00 g, 13.7 mmol, 1.0 *eq*) and compound **58-3** (8.22 g, 13.7 mmol, 1.0 *eq*) in DCM (24 mL) was added EDCI (7.92 g, 41.3 mmol, 3.0 *eq*) and DMAP (2.52 g, 20.6 mmol, 1.5 *eq*). The solution was stirred at 20 °C for 2 hrs. TLC (petroleum ether/ethyl acetate = 5/1, *R_f* of compound **58-6d** was 0.24, *R_f* of compound **58-7d** was 0.38) showed the starting material was consumed completely. The reaction was poured into water (100 mL). The resulting solution was extracted with EtOAc (50 mL x 3). The combined organic phase was washed with brine (50 mL) and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 100/1 to 0/1) to give compound **58-7d** (5.00 g, 5.75 mmol, 41.7% yield) as yellow oil. ¹H NMR: ET73400-23-P1A (400 MHz, CDCl₃) δ 4.90-4.93 (m, 1H), 4.84-4.89 (m, 2H), 3.93-3.97 (m, 2H), 3.82-3.84 (m, 2H), 2.75-2.81 (m, 2H), 2.65-2.69 (m, 1H), 2.51 (t, *J* = 7.2 Hz, 2H), 2.28 (t, *J* = 7.2 Hz, 4H), 2.65-2.69 (m, 1H), 2.09-2.16 (m, 1H), 1.92-1.98 (m, 1H), 1.49-1.63 (m, 19H), 1.20-1.37(m, 44H), 0.89 (t, *J* = 6.8 Hz, 15H).

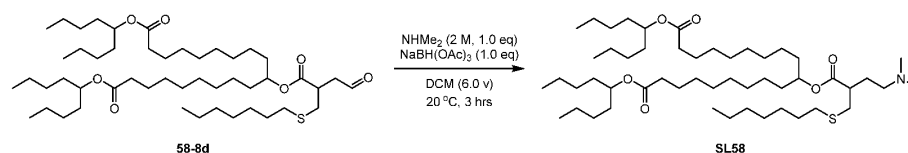
Preparation of compound 58-8d



To a solution of compound **58-7d** (4.00 g, 4.60 mmol, 1.0 eq) in ACETONE (400 mL) was added TsOH.H₂O (350 mg, 1.84 mmol, 0.4 eq). The solution was stirred at 60

°C for 12 hrs. TLC (petroleum ether/ethyl acetate = 5/1, R_f of compound **58-7d** was 0.38, R_f of compound **58-8d** was 0.43) indicated 60% of Reactant 1 was remained. The suspension was concentrated under reduced pressure to give a residue. The residue was poured into water (100 mL). The resulting solution was extracted with EtOAc (50 mL x 3). The combined organic phase was washed with brine (50 mL) and concentrated under vacuum to give the crude product as yellow oil. Compound **58-8d** (4.00 g, crude) was obtained as yellow oil which the residue was used next step and without purification.

10 Preparation of SL58

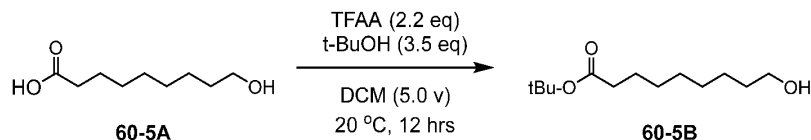


To a solution of compound **58-8d** (4.00 g, 4.85 mmol, 1.0 *eq*) in DCM (30 mL) was added NaBH(OAc)₃ (1.03 g, 4.85 mmol, 1.0 *eq*) and N-methylmethanamine (2 M, 2.42 mL, 1.0 *eq*). The solution was stirred at 20 °C for 3 hrs. TLC (dichloromethane/methanol = 10/1, R_f of compound **58-8d** was 0.34, R_f of **SL58** was 0.43) showed the starting material was consumed completely. The reaction was poured into water (50 mL). The resulting solution was extracted with EtOAc (30 mL x 3). The combined organic phase was washed with water (30 mL x 3) and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 100/1 to 0/1) to give **SL58** (1.00 g, 1.17 mmol, 24.1% yield) as a light-yellow oil. ¹H NMR: ET73400-32-P1A (400 MHz, CDCl₃) δ 4.84-4.90 (m, 3H), 2.75-2.81 (m, 1H), 2.60-2.65 (m, 2H), 2.51 (t, *J* = 6.8 Hz, 2H), 2.25-2.30 (m, 6H), 2.20 (s, 6H), 1.75-1.84 (m, 2H), 1.49-1.65 (m, 18H), 1.20-1.39(m, 44H), 0.89 (t, *J* = 7.2 Hz, 15H).

25

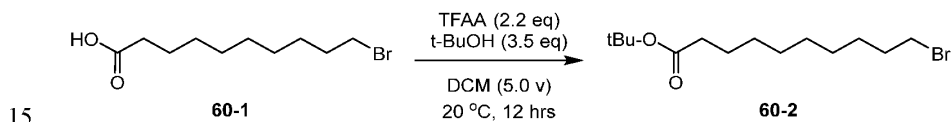
Preparation of compound 60-5B

124

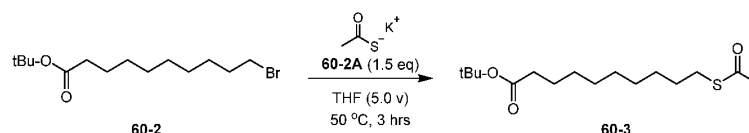


To a solution of compound **60-5A** (25.0 g, 144 mmol, 1.0 *eq*) in DCM (125 mL) was added TFAA (66.3 g, 316 mmol, 43.9 mL, 2.2 *eq*) and t-BuOH (37.2 g, 502 mmol, 48.0 mL, 3.5 *eq*). The suspension was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 10/1, R_f of compound **60-5A** was 0.29, R_f of compound **60-5B** was 0.80) showed the reaction was completed. The reaction was poured into aqueous solution of H₂O (aq. 100 mL). The resulting solution was extracted with DCM (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 1/0 to 10/1) to give colorless oil. Compound **60-5B** (28.0 g, 122 mmol, 84.7% yield) was obtained as colorless oil.

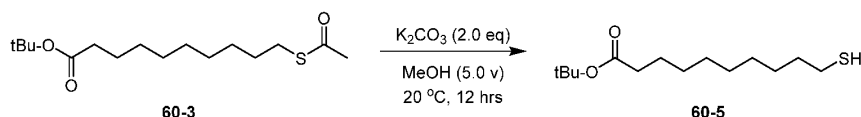
Preparation of compound **60-2**



To a solution of compound **60-1** (40.0 g, 159 mmol, 1.0 *eq*) in DCM (200 mL) was added TFAA (73.6 g, 350 mmol, 48.7 mL, 2.2 *eq*) and t-BuOH (41.3 g, 557 mmol, 53.3 mL, 3.5 *eq*). The solution was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 10/1, R_f of compound **60-1** was 0.29, R_f of compound **60-2** was 0.80) showed the reaction was completed. The reaction was poured into aqueous solution of H₂O (aq. 100 mL). The resulting solution was extracted with DCM (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 1/0 to 10/1) to give colorless oil. Compound **60-2** (45.0 g, 146 mmol, 91.9% yield) was obtained as yellow oil.

Preparation of compound 60-3

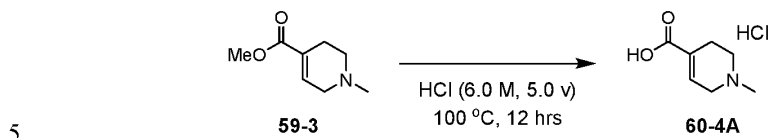
To a solution of compound **60-2** (40.0 g, 130 mmol, 1.0 *eq*) in THF (200 mL) was added $\text{C}_2\text{H}_3\text{OSK}$ (22.3 g, 195 mmol, 1.5 *eq*). The suspension was stirred at 50 °C for 3 hrs. TLC (Petroleum ether/Ethyl acetate = 10/1, R_f of compound **60-2** was 0.21, R_f of compound **60-3** was 0.69) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 200 mL). The resulting solution was extracted with EtOAc (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. Compound **60-3** (30.0 g, 99.2 mmol, 76.2% yield) was obtained as yellow oil. ^1H NMR: ET73270-54-P1A (400 MHz, CDCl_3) δ 2.79 (t, $J = 7.4$ Hz, 2H), 2.25 (s, 3H), 2.13 (t, $J = 7.4$ Hz, 2H), 1.45-1.50 (m, 5H), 1.37 (s, 10H), 1.25-1.27 (m, 3H), 1.21 (s, 9H).

Preparation of compound 60-5

To a solution of compound **60-3** (30.0 g, 99.2 mmol, 1.00 *eq*) in MeOH (150 mL) was added K_2CO_3 (27.4 g, 198 mmol, 2.0 *eq*). The suspension was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 8/1, R_f of compound **60-3** was 0.61, R_f of compound **60-5** was 0.57) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 100 mL). The resulting solution was extracted with EtOAc (200 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. Compound **60-5** (24.0 g, 92.2 mmol, 92.9% yield) was obtained as

colorless oil. ^1H NMR: ET73270-92-P1A (400 MHz, CDCl_3) δ 2.52-2.56 (m, 2H), 2.21 (t, $J = 7.4$ Hz, 2H), 1.55-1.61 (m, 4H), 1.45 (s, 9H), 1.32-1.35 (m, 2H), 1.30 (s, 8H).

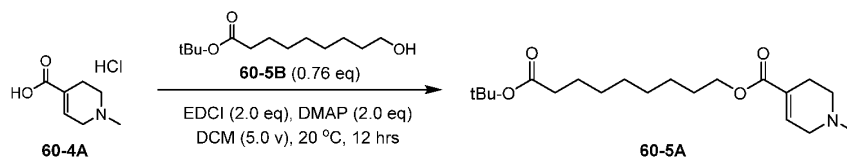
Preparation of compound 60-4A



A solution of compound **59-3** (18.0 g, 115.98 mmol, 1.0 eq) in HCl (6M, 90 mL, 4.66 eq) was stirred at 100 °C for 12 hrs. LC-MS showed the reaction was finished. Concentrated the reaction under vacuum to give the crude product as yellow solid. Compound **60-4A** (20.0 g, 113 mmol, 97.1% yield, HCl) was obtained as a yellow solid.

10 LCMS: ET73270-90-P1A, RT = 0.053, $\text{M} + \text{H}^+ = 142.08$.

Preparation of compound 60-5A

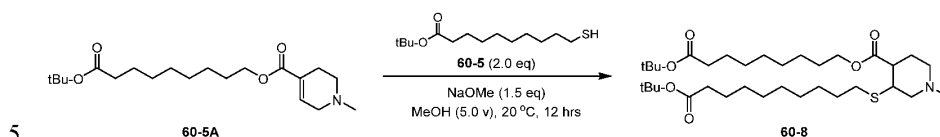


To a solution of compound **60-4A** (20.0 g, 113 mmol, 1.0 eq, HCl) in DCM (100 mL) was added EDCI (43.2 g, 225 mmol, 2.0 eq), DMAP (27.5 g, 225 mmol, 2.0 eq) and compound **60-5B** (19.7 g, 85.6 mmol, 0.76 eq). The solution was stirred at 20 °C for 12 hrs. TLC (Petroleum ether/Ethyl acetate = 8/1, R_f of compound **60-4A** was 0.63, R_f of compound **60-5A** was 0.00) showed the reaction was completed. The reaction was poured into aqueous solution of H_2O (aq. 20 mL). The resulting solution was extracted with DCM (10 mL x 2). The combined organic phase was washed with brine (50 mL x 2), dried over anhydrous Na_2SO_4 and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 1/0 to 0/1) to give colorless oil. Compound **60-5A** (6.00 g, 16.9 mmol, 15.1% yield) was obtained as colorless oil. ^1H NMR: ET73270-91-P1A (400 MHz, CDCl_3) δ 6.88 (t, $J = 1.6$ Hz, 1H), 4.14 (t, $J = 6.8$ Hz, 4H), 3.14 (s,

25

2H), 2.58-2.59 (m, 2H), 2.47 (s, 2H), 2.41 (s, 3H), 1.58-1.66 (m, 4H), 1.45 (s, 9H), 1.25-1.31 (m, 8H).

Preparation of compound 60-8



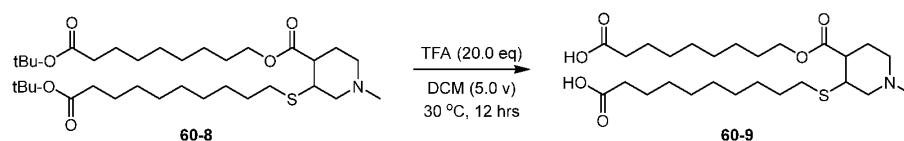
To a solution of compound **60-5A** (6.00 g, 16.9 mmol, 1.0 *eq*) in MeOH (30 mL) was added NaOMe (1.38 g, 25.5 mmol, 1.5 *eq*) and compound **60-5A** (8.84 g, 33.9 mmol, 2.0 *eq*). The suspension was stirred at 20 °C for 12 hrs. TLC (Dichloromethane/Methanol = 8/1, *R_f* of compound **60-5A** was 0.52, *R_f* of compound

10 **60-8** was 0.55) showed the reaction was completed. The reaction was poured into aqueous solution of H₂O (aq. 100 mL). The resulting solution was extracted with EtOAc (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum

15 ether/Ethyl acetate = 50/1 to 2/1) to give yellow oil. Compound **60-8** (8.00 g, 13.0 mmol, 76.8% yield) was obtained as yellow oil. ¹H NMR: ET73270-93-P1A (400 MHz, CDCl₃) δ 4.07-4.13 (m, 3H), 3.41 (s, 1H), 2.85 (s, 1H), 2.56 (s, 1H), 2.78 (s, 1H), 2.39-2.63 (m, 2H), 2.24 (s, 1H), 2.29 (s, 3H), 2.20 (t, *J* = 6.6 Hz, 4H), 1.94-2.03 (m, 3H), 1.76 (s, 1H), 1.52-1.57 (m, 9H), 1.44 (s, 18H), 1.26-1.35 (m, 22H).

20

Preparation of compound 60-9

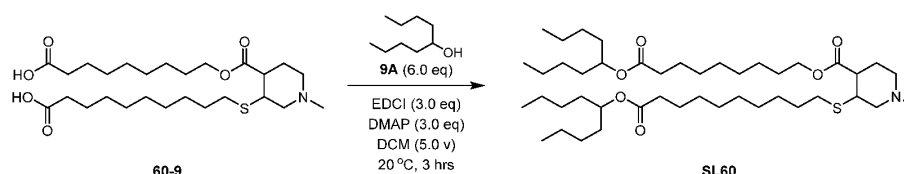


To a solution of compound **60-8** (8.00 g, 13.0 mmol, 1.0 *eq*) in DCM (40 mL) was added TFA (29.7 g, 261 mmol, 19.4 mL, 20 *eq*). The solution was stirred at 30 °C for 12

25 hrs. TLC (Dichloromethane/Methanol = 8/1, *R_f* of compound **60-8** was 0.53, *R_f* of

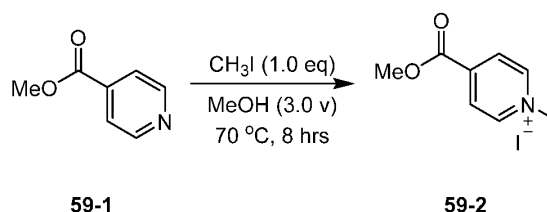
compound **60-9** was 0.00) showed the reaction was completed. Concentrated the reaction under vacuum to give the crude product as yellow oil. Compound **60-9** (4.00 g, 6.50 mmol, 49.9% yield, TFA) was obtained as yellow oil.

5 Preparation of **SL60**



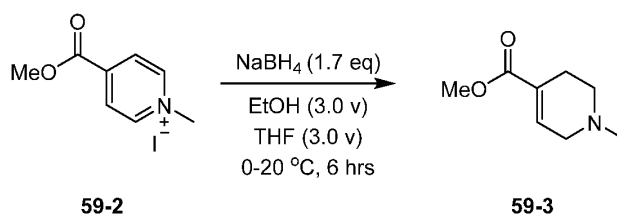
To a solution of compound **60-9** (4.00 g, 6.50 mmol, 1.0 *eq*, TFA) in DCM (20 mL) was added EDCI (3.74 g, 19.5 mmol, 3.0 *eq*), DMAP (2.38 g, 19.5 mmol, 3.0 *eq*) and compound **9A** (5.62 g, 38.9 mmol, 6.0 *eq*). The suspension was stirred at 20 °C for 3 hrs. TLC (Dichloromethane/Methanol = 8/1, R_f of compound **60-9** was 0.01, R_f of compound **SL60** was 0.68) showed the reaction was completed. The reaction was poured into aqueous solution of H₂O (aq. 100 mL). The resulting solution was extracted with DCM (100 mL x 2). The combined organic phase was washed with brine (100 mL x 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product as yellow oil. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 20/1 to 8/1) to give yellow oil. **SL60** (1.00 g, 1.33 mmol, 20.4% yield) was obtained as a light-yellow oil. ¹H NMR: ET73270-95-P1A (400 MHz, CDCl₃) δ 4.86-4.91 (m, 2H), 4.05-4.15 (m, 2H), 3.20 (s, 1H), 2.93 (s, 1H), 2.81 (s, 1H), 2.62 (s, 1H), 2.50-2.59 (m, 2H), 2.38 (s, 1H), 2.27-2.30 (m, 7H), 1.92-2.13 (m, 2H), 1.83-1.87 (m, 1H), 1.62-1.71 (m, 6H), 1.52-1.57 (m, 10H), 1.27-1.33 (m, 34H), 0.89 (t, *J* = 6.8 Hz, 12H).

Preparation of compound 59-2



To a solution of compound **59-1** (200 g, 1.46 mol, 172 mL, 1.0 eq) in MeOH (600 mL) was added CH₃I (207 g, 1.46 mol, 91 mL, 1.0 eq). The solution was stirred at 70 °C for 8 hrs. TLC showed the starting material was consumed completely. The reaction suspension was filtered and the filter cake was concentrated under reduced pressure to give compound **59-2** (400 g, 1.43 mol, 98.28% yield) as a yellow solid. ¹H NMR: ET73288-4-P1A (400 MHz, CDCl₃) δ 9.16-9.18 (d, *J* = 6.4 Hz, 2H), 8.49-8.51 (d, *J* = 6.4 Hz, 2H), 4.43 (s, 3H), 3.99 (s, 3H).

Preparation of compound 59-3

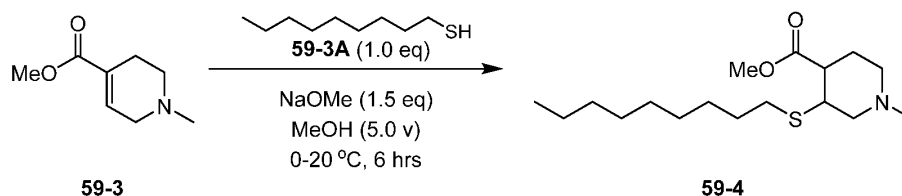


10

To a solution of compound **59-2** (90 g, 322.5 mmol, 1.0 eq) in EtOH (270 mL) and THF (270 mL) was added NaBH₄ (20.7 g, 548.2 mmol, 1.7 eq) at 0-10°C. The reaction solution was stirred at 20°C for 6 hrs. TLC showed the starting material was consumed completely. The solution was poured into NH₄Cl (sat. 300 mL) and extracted with EtOAc (200 mLx2). The combined organic layers were washed with brine (200 mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a compound **59-3** (45 g, 290 mmol, 89.9% yield) as yellow oil. ¹H NMR: ET73288-5-P1A (400 MHz, CDCl₃) δ 6.87-6.89 (m, 1H), 3.73 (m, 3H), 3.08 (q, *J* = 3.2 Hz, 2H), 2.54 (t, *J* = 5.6 Hz, 2H), 2.42-2.44 (m, 2H), 2.03 (s, 3H).

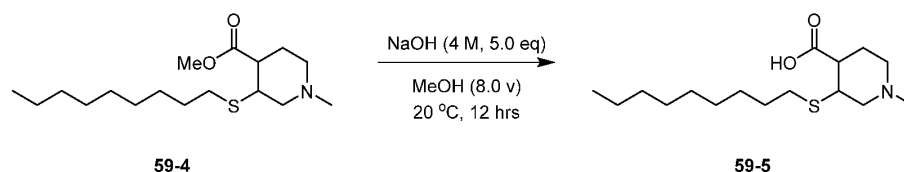
20

Preparation of compound 59-4



To a solution of compound **59-3** (10.0 g, 64.4 mmol, 1.0 eq) and compound **59-3A** (10.3 g, 64.4 mmol, 1.0 eq) in MeOH (50 mL) was added NaOMe (5.2 g, 96.7 mmol, 1.5 eq) at 0-10°C. The reaction solution was stirred at 20°C for 6 hrs. TLC showed the starting material was consumed completely. The solution was poured into NH₄Cl (sat. 50 mL) and extracted with EtOAc (50 mL x 2). The combined organic layers were washed with brine (50 mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give compound **59-4** (15 g, 47.5 mmol, 73.8% yield) was obtained as yellow oil. ¹H NMR: ET73288-18-P1A (400 MHz, CDCl₃) δ 3.71 (s, 3H), 3.21 (s, 1H), 2.91-2.94 (m, 1H), 2.72-2.79 (m, 1H), 2.54-2.56 (m, 1H), 2.49-2.52 (m, 2H), 2.43-2.49 (m, 1H), 2.29 (s, 3H), 2.02-2.13 (m, 2 H), 1.82-2.00 (m, 1H), 1.53-1.57 (m, 2H), 1.26-1.34 (m, 12H), 0.88 (t, *J* = 6.8 Hz, 3H).

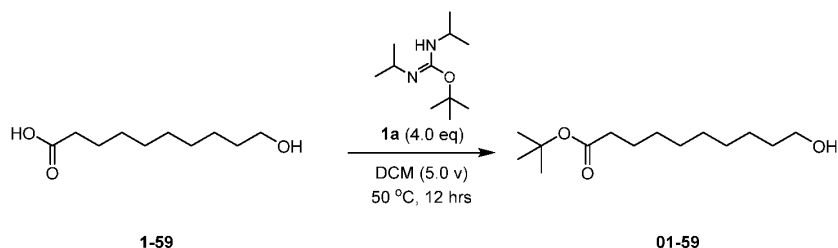
Preparation of compound 59-5



To a solution of compound **59-4** (5.0 g, 15.9 mmol, 1.0 eq) and NaOH (4 M, 19.8 mL, 5.0 eq) in MeOH (40 mL). The solution was stirred at 20 °C for 12 hrs. TLC showed the starting material was consumed completely. The solution was poured into HCl (4M, 100 mL) till PH = 2-3 and extracted with DCM (200 mLx2). The combined organic layers were washed with brine (200 mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give compound **59-5** (4.0 g, 13.3 mmol, 83.7% yield) as yellow oil. ¹H NMR: ET73288-20-P1A (400 MHz, CDCl₃) δ 6.90 (s, 1H), 3.54-3.76 (m, 1H), 3.46-3.52 (m, 2H), 3.14-3.22 (m, 1H), 2.81-2.86 (m, 2H), 2.63-2.68 (m, 2H), 2.09-2.53 (m, 3H), 1.57-1.60 (m, 2 H), 1.25-1.36 (m, 12H), 0.88 (t, *J* = 6.8 Hz, 3H).

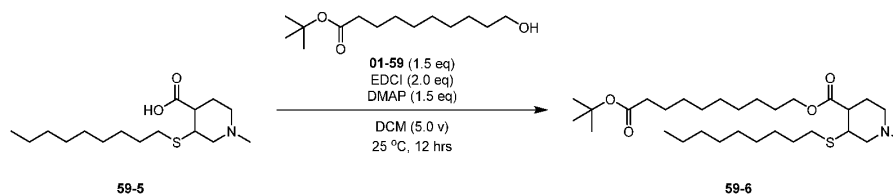
25 Preparation of compound 01-59

131



To a solution of compound **1a** (12.8 g, 63.7 mmol, 4.0 eq) in DCM (15 mL) was added compound **1-59** (3.0 g, 15.9 mmol, 1.0 eq). The solution was stirred at 50 °C for 12 hrs. TLC (DCM: Methanol = 8: 1, R_f of **1-59** was 0.70, R_f of **01-59** was 0.65) showed the starting material was consumed completely. The solution was concentrated under reduced pressure to give a residue. The residue was triturated with MTBE (20 mL) at 25 °C for 30 min. Then the suspension was filtered, and filtrate was concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether: Ethyl acetate=15: 1 to 2: 1) to give compound **01-59** (2.9 g, 11.9 mmol, 74.5% yield) as colorless oil. ¹H NMR: ET74125-6-P1A (400 MHz CDCl₃) δ 3.64-3.65 (m, 2H), 2.19-2.23 (t, J = 7.6 Hz, 2H), 1.55-1.61 (m, 4H), 1.41-1.47 (m, 10H), 1.28-1.33 (m, 9H).

Preparation of compound **59-6**

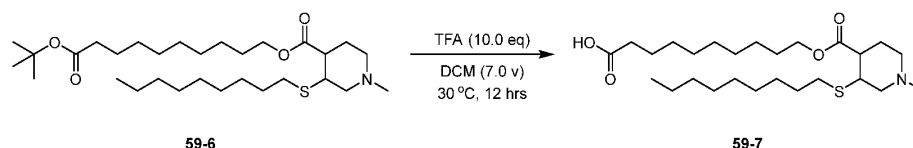


To a solution of **01-59** (2.4 g, 9.9 mmol, 1.5 eq) in DCM (10 mL) were added DMAP (1.2 g, 9.9 mmol, 1.5 eq), EDCI (2.5 g, 13.3 mmol, 2.0 eq) and compound **59-5** (2.0 g, 6.6 mmol, 1.0 eq). The solution was stirred at 25 °C for 12 hrs. TLC (DCM: Methanol = 8: 1, R_f of **59-5** was 0.05, R_f of **59-6** was 0.45) showed the starting material was consumed completely. The reaction solution was poured into H₂O (10 mL). The solution was extracted with DCM (10 mL) three times. The combined organic layer was dried over sodium sulfate and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether: Ethyl acetate =

132

10: 1 to 2: 1) to give compound **59-6** (2.0 g, 3.8 mmol, 57.1% yield) was obtained as colorless oil. ¹H NMR: ET74125-9-P1A (400 MHz CDCl₃) δ 4.05-4.13 (m, 2H), 3.16-3.25 (m, 1H), 2.91-3.00 (m, 1H), 2.80-2.85 (m, 1H), 2.50-2.55 (m, 2H), 2.34-2.41 (m, 1H), 2.50-2.57 (m, 2H), 2.35-2.40 (m, 1H), 2.27-2.31 (m, 2H) 2.18-2.22 (m, 3H), 1.96-2.06 (m, 2H), 1.75-1.92 (m, 2H), 1.62-1.66 (m, 2H), 1.55-1.59 (m, 4H), 1.45 (s, 9H), 1.26-1.30 (m, 22H), 0.87-0.90 (m, 3H).

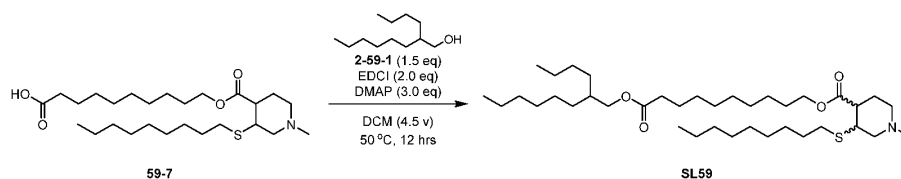
Preparation of compound **59-7**



10 To a solution of compound **59-6** (2.0 g, 3.8 mmol, 1.0 eq) in DCM (14 mL) was added TFA (4.3 g, 37.9 mmol, 2.81 mL, 10.0 eq). The solution was stirred at 30°C for 12 hrs. LC-MS (ET74125-11-P1, product: RT = 0.55, 0.68) showed the starting material was consumed completely. The solution concentrated under reduced pressure to give a compound **59-7** (2.2 g, 3.76 mmol, 99.1% yield, TFA) as colorless oil. ¹H NMR:

15 ET74125-11-P1A (400 MHz CDCl₃) δ 4.07-4.19 (m, 2H), 3.70-3.95 (m, 1H), 3.40-3.60 (m, 2H), 3.08-3.29 (m, 2H), 2.86 (s, 2H), 2.60-2.68 (m, 2H), 2.33-2.42 (m, 3H), 2.10-2.22 (m, 1H), 1.62-1.66 (m, 4H), 1.53-1.60 (m, 2H), 1.27-1.41 (m, 24H), 0.86-0.91 (m, 3H).

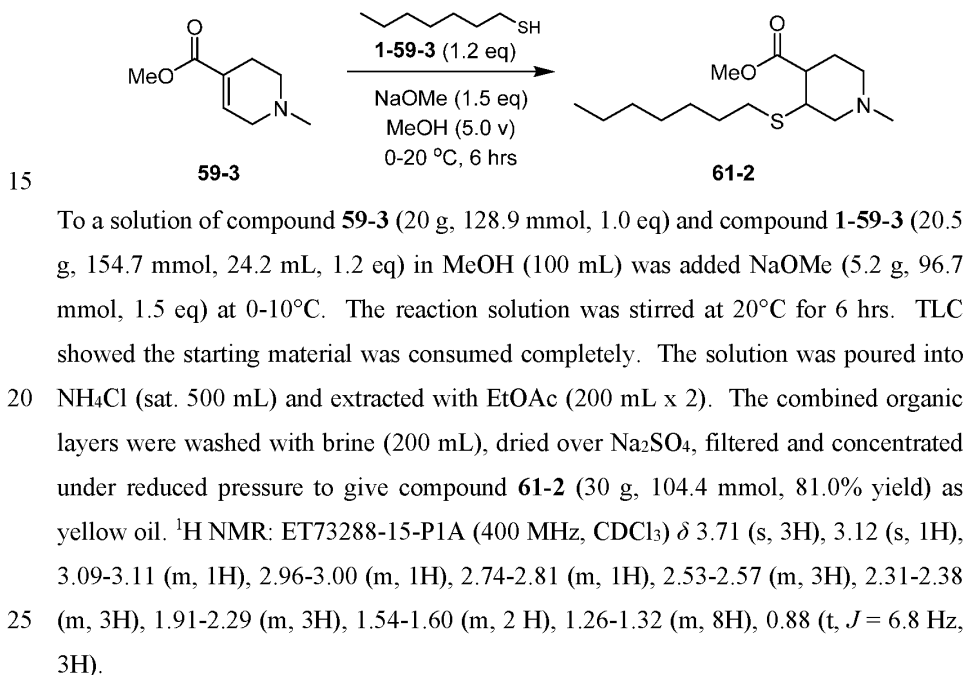
20 Preparation of compound **SL59**



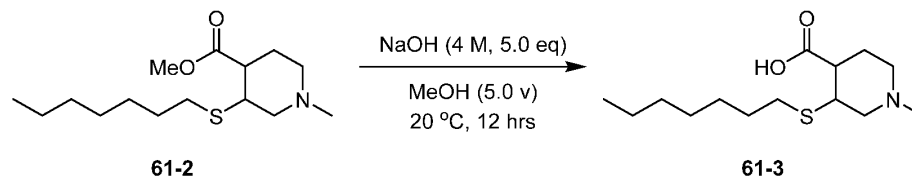
To a solution of compound **2-59-1** (1.1 g, 5.6 mmol, 1.5 eq) in DCM (10 mL) were added EDCI (1.4 g, 7.5 mmol, 2.0 eq), DMAP (1.4 g, 11.3 mmol, 3.0 eq) and compound **59-7**

(2.2 g, 3.8 mmol, 1.0 eq, TFA). The solution was stirred at 50 °C for 12 hrs. TLC (DCM: Methanol = 8: 1, R_f of **59-7** was 0.50, R_f of compound **SL59** was 0.47) showed the starting material was consumed completely. The reaction solution was poured into HCl (10 mL, 1 M). The solution was extracted with DCM (10 mL) three times. The combined organic layer was poured into Na₂CO₃ (aq., Sat., 20 mL). The solution was extracted with DCM (30 mL). The organic layer was dried over sodium sulfate and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether: Ethyl acetate = 20: 1 to 2: 1) to give compound **SL59** (1.0 g, 1.6 mmol, 41.6% yield) was obtained as a light-yellow oil. ¹H NMR: 5
 10 ET74125-11-P1A (400 MHz CDCl₃) δ 4.16-4.19 (m, 2H), 3.95-4.01 (d, J = 0.56 Hz 2H), 2.76-3.24 (m, 3H), 2.48-2.60 (m, 2H), 2.25-2.42 (m, 6H), 1.94-2.06 (m, 2H), 1.76-1.87 (m, 1H), 1.54-1.66 (m, 8H), 1.24-1.36 (m, 38H), 0.84-0.92 (m, 9H).

Preparation of compound 61-2

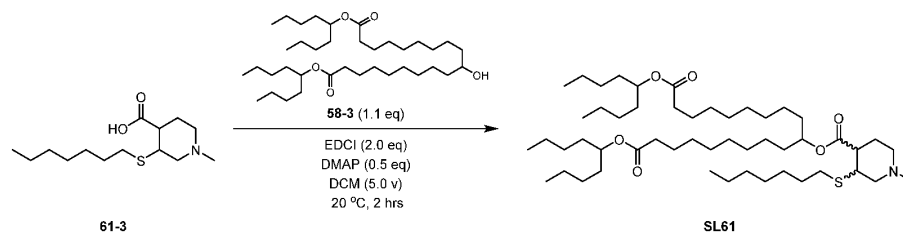


Preparation of compound 61-3



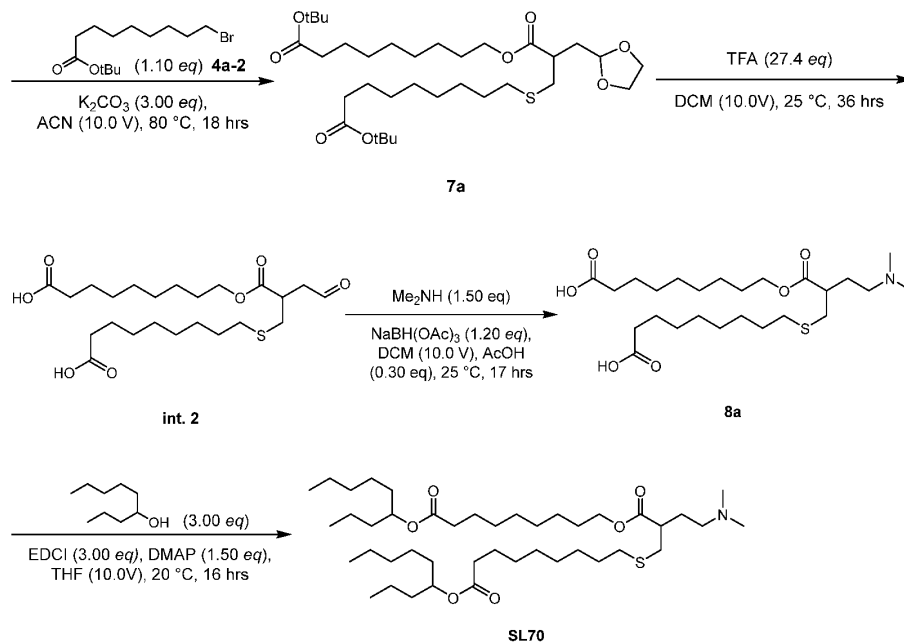
To a solution of compound **61-2** (3.0 g, 10.4 mmol, 1.0 eq) in MeOH (15 mL) was added NaOH (4 M, 13.1 mL, 5.0 eq) at 20°C. The solution was stirred at 20 °C for 12 hrs. TLC showed the starting material was consumed completely. The solution was poured into HCl (4M, 100 mL) till PH = 2-3 and extracted with DCM (200 mLx2). The combined organic layers were washed with brine (200 mL), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give compound **59-5** (4.0 g, 13.3 mmol, 83.7% yield) as yellow oil. LCMS: ET73288-23-P1A (R_T: 2.079 min and 2.133 min, MW=274.1). ¹H NMR: ET73479-14-P1A (400 MHz, CDCl₃) δ 10.60-10.45 (m, 1H), 3.75-3.60 (m, 2H), 3.47-3.45 (m, 2H), 3.02-2.85 (m, 5H), 2.65-2.62 (m, 3H), 2.34-2.22 (m, 2H), 1.59-1.54 (m, 2H), 1.34-1.27 (m, 8H), 0.89-0.85 (m, 3H).

Preparation of compound SL61

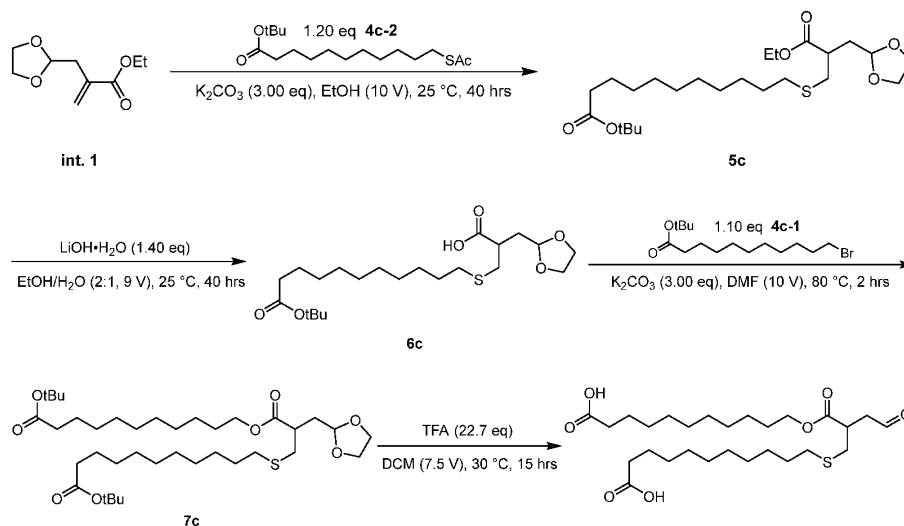


15 To a solution of compound **61-3** (2.00 g, 7.31 mmol, 1.00 eq) in DCM (10 mL) was added compound **58-3** (4.80 g, 8.05 mmol, 1.1 eq), DMAP (446 mg, 3.66 mmol, 0.50 eq) and EDCI (2.80 g, 14.63 mmol, 2.0 eq). The reaction was stirred at 20 °C for 2 hrs. TLC (Dichloromethane: Methanol = 10/1, R_f of product = 0.54) showed the starting material was consumed completely. The reaction solution was poured into the aqueous
20 of H₂O (80 mL), and the solution was extracted with DCM (20 mL x 2). The organic layer was dried over sodium sulfate and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 8/1 to 5/1) to give compound **SL61** (1.0 g, 1.17 mmol, 16.0% yield).

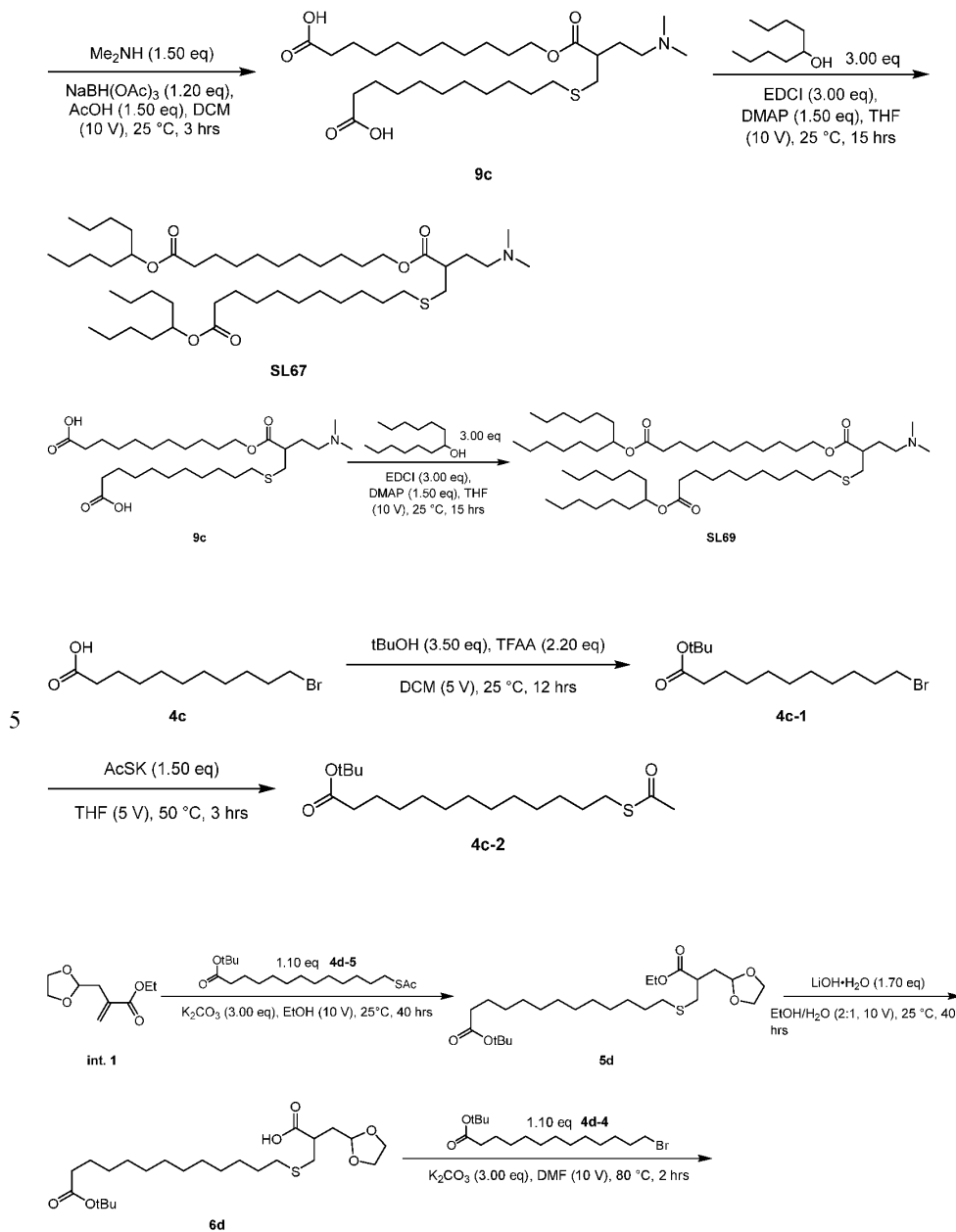
136



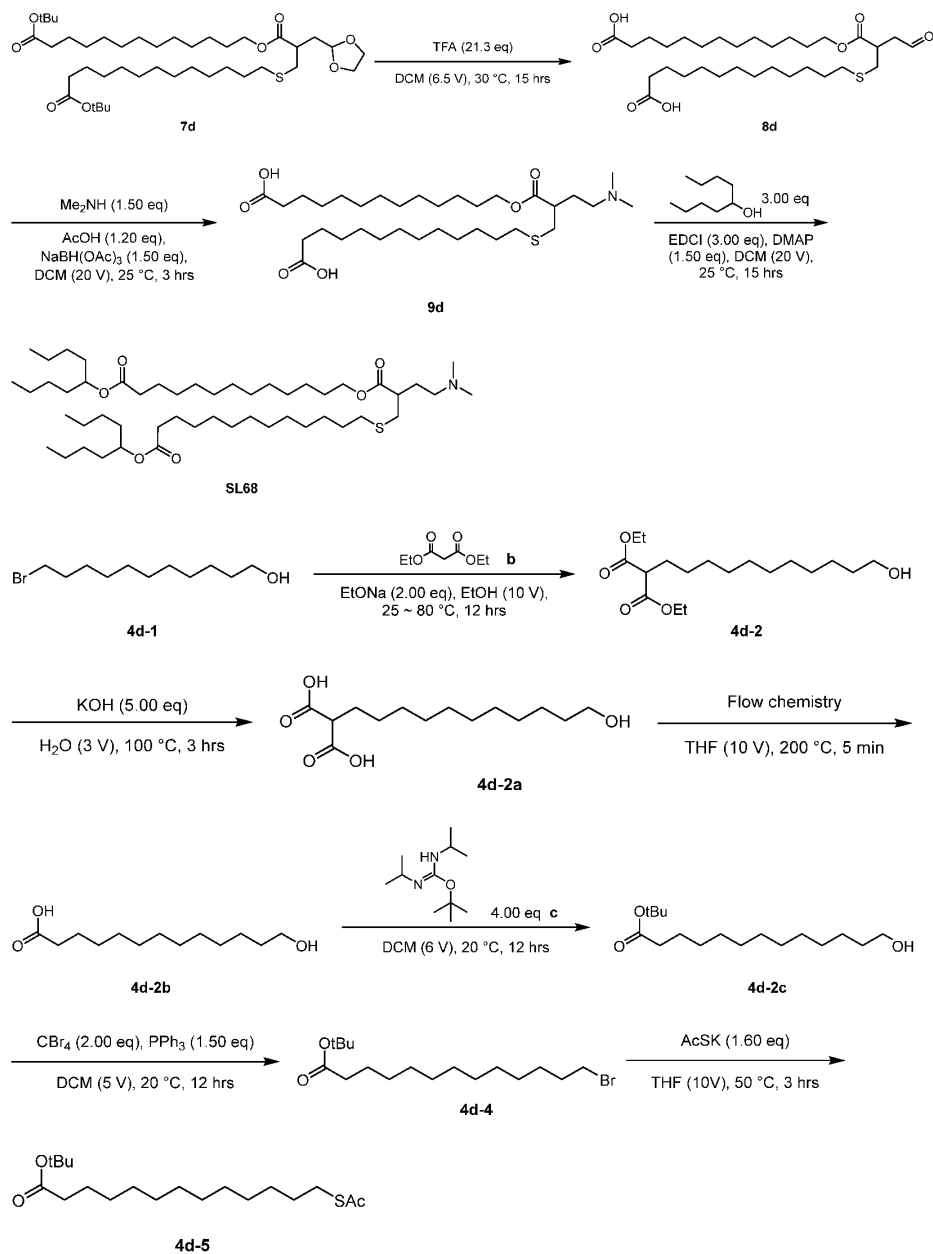
Compounds **SL62**, **SL63**, **SL64**, **SL65**, **SL71**, **SL72**, **SL73**, **SL74**, and **SL75** were synthesised via a comparable method, with derivatisation arising through the corresponding alkylanol reagent in the final reaction.



137

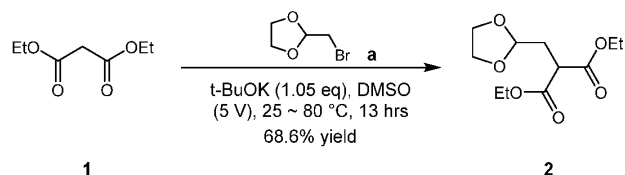


138



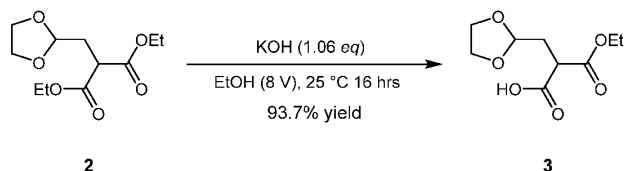
10 General procedure for preparation of compound 2

139



A mixture of compound **1** (457 g, 2.85 mol, 433 mL, 1.00 *eq*) in dimethylsulfoxide (2.50 L) was added portions t-BuOK (336 g, 3.00 mol, 59.9 mL, 1.05 *eq*) stirred for 1 h at 25 °C, then compound **a** (500 g, 3.00 mol, 307 mL, 1.05 *eq*) was added into by dropwise, the mixture was heated to 80 °C for 12 h. LCMS (EW49154-9-P1A, P1: Rt = 0.465 min) showed reactant (R_t = 0.460 min) consumed. TLC (Petroleum ether/ethyl acetate = 5/1, P1: R_f = 0.38). The crude was diluted with methyl tertiary butyl ether (3.00 L) and washed with a solution of NH₄Cl (aq. 5.00 L). The aqueous layer was extracted with methyl tertiary butyl ether (2.00 L × 3) and the combined organic layers were washed with brine (5.00 L × 3), dried over Na₂SO₄, filtered and concentrated. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 6/1 to 3/1) to give compound **2** (482 g, 1.96 mol, 68.6% yield) as a brown oil. LCMS: EW49154-9-P1B, Rt = 0.468 min, m/z = 269.0 (M+23).

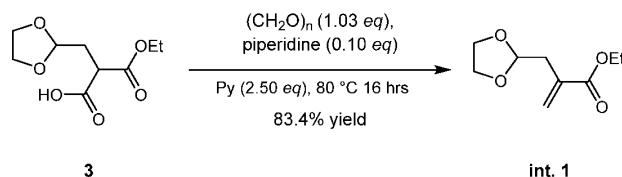
15 General procedure for preparation of compound **3**



Compound **2** (443 g, 1.80 mol, 1.00 *eq*) was dissolved in ethyl alcohol (1.50 L) and KOH (107 g, 1.91 mol, 1.06 *eq*) dissolved in ethyl alcohol (2.00 L) was added dropwise. The solution was stirred for 16 h at 25 °C. LCMS: (EW49333-3-P1B, P1: Rt = 0.365 min) showed 7.74% reactant (R_t = 0.466 min) remained. The mixture was concentrated to a quarter, dissolved in H₂O (1.50 L) which was made more basic with NH₃·H₂O (20.0 mL). The aqueous phase was washed with dichloromethane (500 mL), then made acidic with HCl (2M, 500 mL) to pH = 2, extracted with ethyl acetate (1.00 L × 3). The organic phase

was washed with brine (1.00 L), dried over Na₂SO₄ and concentrated to dryness to give compound **3** (368 g, 1.69 mol, 93.7% yield) as a brown oil which was used in next step without further purification. LCMS: EW49333-3-P1C, Rt = 0.365 min, m/z = 241.0 (M+23). ¹H NMR: EW49333-3-P1A (400 MHz, CDCl₃) δ 5.09 - 4.98 (m, 1H), 4.23 (q, J = 7.0 Hz, 2H), 3.92 - 4.02 (m, 2H), 3.81 - 3.91 (m, 2H), 3.63 (t, J = 7.0 Hz, 1H), 2.32 - 2.49 (m, 2H), 1.27 - 1.32 (m, 3H).

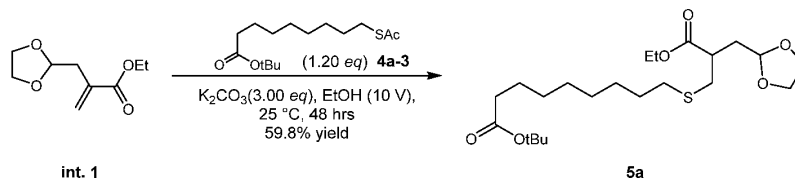
General procedure for preparation of compound int. 1



- 10 To a solution of compound **3** (358 g, 1.64 mol, 1.00 eq) in pyridine (324 g, 4.10 mol, 331 mL, 2.50 eq) was added piperidine (13.9 g, 164 mmol, 16.2 mL, 0.10 eq) and (CH₂O)_n (50.7 g, 1.69 mol, 1.03 eq), then the mixture was stirred at 80 °C for 16 h. LCMS: (EW49333-8-P1A, P1: Rt = 0.448 min) showed compound **3** (Rt = 0.365 min) was consumed, the mixture was diluted with H₂O (2.00 L), and extracted with methyl tertiary
- 15 butyl ether (2.00 L × 2). The organic layer was washed with HCl (2M, 2.00 L × 2), saturated sodium bicarbonate (2.00 L), and brine solution (1.50 L), dried over MgSO₄, and evaporated to dryness to give compound **int. 1** (255 g, 1.37 mol, 83.4% yield) as a brown oil which was used in next step without further purification. LCMS: EW49333-8-P1B (Rt = 0.447 min, m/z = 140.9 (M+1)). ¹H NMR: EW49333-8-P1B (400 MHz, CDCl₃)
- 20 δ 6.28 (d, J = 1.2 Hz, 1H), 5.73 (d, J = 1.2 Hz, 1H), 5.07 (t, J = 5.0 Hz, 1H), 4.22 (q, J = 7.0 Hz, 2H), 3.93 - 4.03 (m, 2H), 3.81 - 3.91 (m, 2H), 2.69 (d, J = 5.0 Hz, 2H), 1.30 (t, J = 7.0 Hz, 3H).

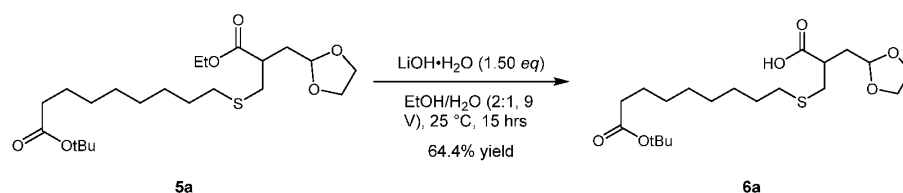
General procedure for preparation of compound 5a

141



To a solution of compound **int. 1** (100 g, 537 mmol, 1.00 *eq*) in ethyl alcohol (1.00 L) was added K_2CO_3 (222 g, 1.61 mol, 3.00 *eq*), then compound **4a-3** (185 g, 644 mmol, 1.20 *eq*) was added into the mixture was stirred at 25 °C for 24 h. Then the mixture was stirred at 25 °C for another 24 h. LCMS: (EW49333-26-P1C, P1: R_t = 0.667 min) showed compound **int.1** (R_t = 0.445 min) was consumed. The mixture was poured into water (1.50 L), extracted by ethyl acetate (2.00 L $\times$ 2), washed with brines (1.50 L $\times$ 2), dried by Na_2SO_4 , concentrated to dryness. The residue was purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 8/1 to 1/1, TLC (Petroleum ether/ethyl acetate = 5/1, P1: R_f = 0.43)) to give compound **5a** (159 g, 321 mmol, 59.8% yield, 87.5% purity) as a brown oil. LCMS: EW49333-26-P1D, R_t = 0.662 min, m/z = 433.1 (M+1). 1H NMR: EW49333-26-P1A (400 MHz, $CDCl_3$) δ 4.95 (t, J = 4.2 Hz, 1H), 4.18 (q, J = 7.0 Hz, 2H), 3.88 – 4.01 (m, 2H), 3.75 - 3.88 (m, 2H), 2.74 - 2.82 (m, 2H), 2.60 - 2.72 (m, 1H), 2.51 (t, J = 7.4 Hz, 2H), 2.20 (t, J = 7.6 Hz, 2H), 2.08 - 2.17 (m, 1H), 1.96 (td, J = 4.2, 14.2 Hz, 1H), 1.50 - 1.60 (m, 4H), 1.41 - 1.49 (m, 9H), 1.26 - 1.40 (m, 11H).

General procedure for preparation of compound 6a

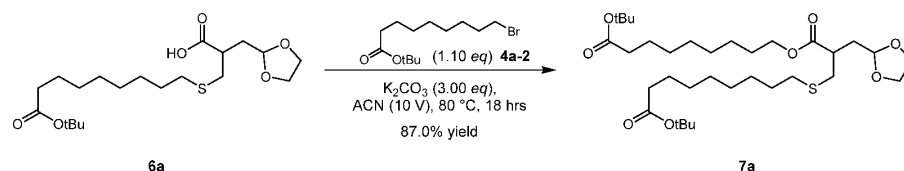


To a solution of compound **5a** (159 g, 367 mmol, 1.00 *eq*) in H_2O (477 mL) and ethyl alcohol (954 mL) was added $LiOH \cdot H_2O$ (23.1 g, 551 mmol, 1.50 *eq*). The mixture was stirred at 25 °C for 15 h. LCMS (EW49333-28-P1A, P1: R_t = 0.583 min) showed 3.44% compound **5a** (R_t = 0.660 min) remained. The mixture was adjusted pH = 3 by citric acid

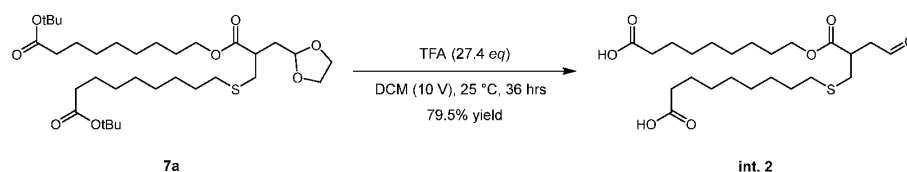
solution (500 mL), stirred for 10 min, extracted by ethyl acetate (2.00 L $\times$ 3), washed with brines (3.00 L), dried by Na₂SO₄, concentrated in vacuum. The residue was purified by column chromatography (SiO₂, Petroleum ether/ethyl acetate=3/1 to 1/1, TLC (petroleum ether/ethyl acetate= 1/1, P1: R_f = 0.58)) to give compound **6a** (106 g, 236 mmol, 64.4% yield, 90.4% purity) as a brown oil. LCMS: EW49333-28-P1D, Rt = 0.584 min, m/z = 427.1 (M+23). ¹H NMR: EW49333-28-P1A (400 MHz, CDCl₃) δ 5.01 (t, *J* = 4.0 Hz, 1H), 3.94 – 4.05 (m, 2H), 3.80 - 3.91 (m, 2H), 2.79 - 2.90 (m, 2H), 2.70 (d, *J* = 6.2 Hz, 1H), 2.48 - 2.58 (m, 2H), 2.18 - 2.24 (m, 2H), 2.14 - 2.18 (m, 1H), 2.00 - 2.05 (m, 1H), 1.58 (quin, *J* = 7.2 Hz, 4H), 1.45 (s, 9H), 1.28 - 1.39 (m, 8H).

10

General procedure for preparation of compound 7a

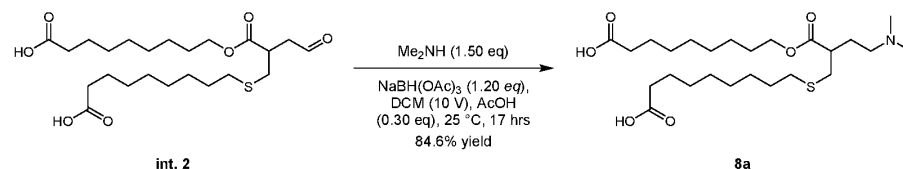


To a solution of compound **6a** (113 g, 279 mmol, 1.00 eq) in acetonitrile (1.13 L) was added K₂CO₃ (115 g, 837 mmol, 3.00 eq), then compound **4a-1** (90.1 g, 307 mmol, 1.10 eq) was added, the mixture was stirred at 80 °C for 18 h. LCMS: (EW49333-31-P1A, P1: Rt = 0.776 min) showed compound **6a** (Rt = 0.582 min) was consumed. The mixture was concentrated to remove acetonitrile, poured into H₂O (1.50 L), extracted by ethyl acetate (1.50 L $\times$ 3), washed with brines (1.50 L $\times$ 2), dried by Na₂SO₄, concentrated to dryness. The residue was purified by column chromatography (SiO₂, Petroleum ether/ethyl acetate = 5/1 to 3/1, TLC (Petroleum ether/ethyl acetate = 5/1, P1: R_f = 0.50)). compound **7a** (150 g, 243 mmol, 87.0% yield) was obtained as a brown oil. LCMS: EW49333-31-P1D, Rt = 0.775 min, m/z = 639.4 (M+23). ¹H NMR: EW49333-31-P1A (400 MHz, CDCl₃) δ 4.94 (t, *J* = 4.2 Hz, 1H), 4.10 (t, *J* = 6.8 Hz, 2H), 3.90 – 4.00 (m, 2H), 3.79 - 3.88 (m, 2H), 2.74 - 2.83 (m, 2H), 2.62 - 2.72 (m, 1H), 2.51 (t, *J* = 7.4 Hz, 2H), 2.20 (t, *J* = 7.4 Hz, 4H), 2.09 - 2.17 (m, 1H), 1.92 – 2.00 (m, 1H), 1.60 - 1.69 (m, 3H), 1.51 - 1.58 (m, 5H), 1.45 (s, 18H), 1.27 - 1.40 (m, 16H).

General procedure for preparation of compound int. 2

To a stirred solution of compound **7a** (75.0 g, 121 mmol, 1.00 eq) in dichloromethane (800 mL) was added TFA (379 g, 3.33 mol, 247 mL, 27.4 eq), the solution was stirred at 25 °C for 36 h. LCMS (EW49333-38-P1B, P1: R_t = 0.529 min) showed intermediate remained. Concentrated the reaction under vacuum to give the crude product as yellow oil. The mixture was purified by silica column (petroleum ether/tetrahydrofuran gradient), product come out at 29% tetrahydrofuran, TLC (petroleum ether/tetrahydrofuran = 2/1, P1: R_f = 0.24). Compound **int. 2** (48.0 g, 96.7 mmol, 79.5% yield, 92.8% purity) was obtained as a brown oil. LCMS: EW49333-38-P1C, R_t = 0.530 min, m/z = 461.2 (M+1). $^1\text{H NMR}$: EW49333-38-P1A (400 MHz, CDCl_3) δ 9.81 (s, 1H), 4.04 - 4.17 (m, 2H), 3.08 - 3.20 (m, 1H), 2.77 - 3.01 (m, 3H), 2.66 (dd, J = 8.0, 13.4 Hz, 1H), 2.50 (t, J = 7.2 Hz, 2H), 2.36 (t, J = 7.2 Hz, 4H), 1.52 - 1.71 (m, 8H), 1.33 (br s, 17H).

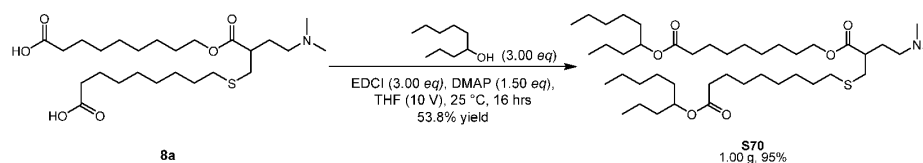
15

General procedure for preparation of compound 8a

To a solution of compound **int.2** (48.0 g, 104 mmol, 1.00 eq) in dichloromethane (400 mL) was added Me_2NH (2 M, 78.2 mL, 1.50 eq) and AcOH (1.88 g, 31.2 mmol, 1.79 mL, 0.30 eq), the solution was stirred at 25 °C for 1 h. Then the mixture was added $\text{NaBH}(\text{OAc})_3$ (26.5 g, 125 mmol, 1.20 eq), and the solution was stirred at 25 °C for 16 h. LCMS: (EW49333-39-P1A, P1: R_t = 0.457 min) showed compound **int.2** (R_t = 0.550 min) remained. The mixture was adjusted pH = 8 by NaHCO_3 solution (aq. 100 mL), extracted by dichloromethane (300 mL), then the aqueous phase was adjusted pH to 3 by

HCl (2M), suspension formed, extracted by tetrahydrofuran (500 mL $\times$ 4), combined the organic layer, dried by Na₂SO₄, concentrated to get an oil. The residue was purified by flash silica gel chromatography (ISCO®; 330 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 10% methyl alcohol/dichloromethane gradient @ 100 mL/min), product
 5 come out at 4% methyl alcohol, TLC (Dichloromethane/Methanol = 10/1, P1: R_f = 0.24). compound **8a** (46.0 g, 88.2 mmol, 84.6% yield, 93.9% purity) was obtained as a brown oil. LCMS: EW49333-39-P1D, R_t = 0.457 min, m/z = 490.3 (M+1). ¹H NMR: EW49333-39-P1A (400 MHz, CDCl₃) δ 4.13 (t, J = 6.2 Hz, 2H), 3.25 – 4.20 (m, 5H), 3.11 - 3.25 (m, 1H), 2.94 – 3.08 (m, 1H), 2.82 (s, 6H), 2.65 - 2.77 (m, 2H), 2.52 (t, J =
 10 7.2 Hz, 2H), 2.35 (t, J = 7.2 Hz, 3H), 2.11 - 2.25 (m, 2H), 1.51 - 1.86 (m, 7H), 1.20 - 1.45 (m, 13H).

Preparation of compound SL70



15 To a solution of compound **8a** (2.50 g, 5.11 mmol, 1.00 eq) and nonan-4-ol (2.21 g, 15.3 mmol, 3.00 eq) in tetrahydrofuran (25.0 mL) was added EDCI (2.94 g, 15.3 mmol, 3.00 eq) and DMAP (935 mg, 7.66 mmol, 1.50 eq). The solution was stirred at 25 °C for 16 h. LCMS: (EW49333-46-P1A) showed desired mass (R_t = 0.763 min, m/z = 742.8 (M+1)) was detected. The mixture was poured into NaHCO₃ solution (150 mL), extracted
 20 by ethyl acetate (250 mL), washed by HCl solution (1M, 150 mL), and followed by saturated NaHCO₃ solution (150 mL), brines (150 mL $\times$ 2), dried by Na₂SO₄, concentrated to dryness. The residue was purified by flash silica gel chromatography (ISCO®; 20 g Sepa Flash® Silica Flash Column, Eluent of 0~1% methanol/dichloromethane gradient @ 30 mL/min). The residue was further purified by
 25 flash silica gel chromatography (ISCO®; 20 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 25% tetrahydrofuran/petroleum ether gradient @ 30 mL/min), the residue was further purified by flash silica gel chromatography (ISCO®; 40 g Sepa Flash® Silica

Flash Column, Eluent of 0 ~ 1% Methanol/Dichloromethane gradient (@ 30 mL/min). compound **SL70** (2.04 g, 2.75 mmol, 53.8% yield) was obtained as a colorless oil. Special LCMS: EW49333-54-P1A, Rt = 3.688 min, m/z = 742.8 (M+1). ¹H NMR: EW49333-54-P1A (400 MHz, CDCl₃) δ 4.81 - 4.96 (m, 2H), 4.10 (t, J = 6.6 Hz, 2H),
5 2.73 - 2.84 (m, 1H), 2.59 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.23 - 2.35 (m, 6H), 2.21 (s, 6H), 1.80 (ddd, J = 2.4, 5.8, 8.0 Hz, 2H), 1.57 - 1.68 (m, 7H), 1.42 - 1.56 (m, 9H), 1.18 - 1.41 (m, 32H), 0.83 - 0.95 (m, 12H).

Preparation of compound SL62

- 10 Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL62** (1.50 g, colorless oil). Special LCMS: EW49333-55-P1A, Rt = 3.942 min, m/z = 798.9 (M+1). ¹H NMR: EW49333-55-P1A (400 MHz, CDCl₃) δ 4.84 - 4.91 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.72 - 2.85 (m, 1H), 2.58 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.23 - 2.33 (m, 6H), 2.21 (s, 6H), 1.72 - 1.89 (m, 2H), 1.57 - 1.67 (m, 7H), 1.44 - 1.56 (m, 9H),
15 1.19 - 1.42 (m, 40H), 0.84 - 0.95 (m, 12H).

Preparation of compound SL63

- Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL63** (2.00 g, colorless oil). Special LCMS: EW49333-56-P1B, Rt = 4.166 min, m/z = 855.0 (M+1). ¹H NMR: EW49333-56-P1A (400 MHz, CDCl₃) δ 4.84 - 4.90 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.71 - 2.83 (m, 1H), 2.59 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.24 - 2.33 (m, 6H), 2.21 (s, 6H), 1.73 - 1.89 (m, 2H), 1.58 - 1.69 (m, 7H), 1.45 - 1.56 (m, 9H), 1.19 - 1.40 (m, 48H), 0.83 - 0.94 (m, 12H).

25 *Preparation of compound SL64*

Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL64** (1.10 g, colorless oil). Special LCMS: EW49333-57-P1B, Rt = 4.382 min, m/z = 911.1 (M+1). ¹H NMR: EW49333-57-P1A (400 MHz, CDCl₃) δ 4.84 - 4.90 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.72 - 2.86 (m, 1H), 2.58 - 2.71 (m, 2H), 2.51 (t, J = 7.4 Hz, 2H), 2.23 -

146

2.36 (m, 6H), 2.21 (s, 6H), 1.74 - 1.88 (m, 2H), 1.59 - 1.69 (m, 7H), 1.46 - 1.57 (m, 9H), 1.20 - 1.41 (m, 56H), 0.82 - 0.94 (m, 12H).

Preparation of compound SL65

- 5 Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL65** (2.00 g, colorless oil). Special LCMS: EW49333-58-P1B, Rt = 3.213 min, m/z = 967.1 (M+1). ¹H NMR: EW49333-58-P1A (400 MHz, CDCl₃) δ 4.84 - 4.90 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.73 - 2.84 (m, 1H), 2.59 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.24 - 2.35 (m, 6H), 2.21 (s, 6H), 1.73 - 1.87 (m, 2H), 1.58 - 1.69 (m, 7H), 1.44 - 1.56 (m, 9H), 1.19 - 1.40 (m, 64H), 0.88 (t, J = 6.8 Hz, 12H).

Preparation of compound SL71

- Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL71** (1.10 g, colorless oil). Special LCMS: EW49333-59-P1B, Rt = 3.942 min, m/z = 798.9 (M+1). ¹H NMR: EW49333-59-P1A (400 MHz, CDCl₃) δ 4.84 - 4.90 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.73 - 2.84 (m, 1H), 2.59 - 2.71 (m, 2H), 2.51 (t, J = 7.4 Hz, 2H), 2.23 - 2.36 (m, 6H), 2.21 (s, 6H), 1.73 - 1.87 (m, 2H), 1.59 - 1.69 (m, 7H), 1.46 - 1.56 (m, 9H), 1.21 - 1.39 (m, 40H), 0.83 - 0.93 (m, 12H).

20 *Preparation of compound SL72*

- Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL72** (1.10 g, colorless oil). Special LCMS: EW49333-60-P1B, Rt = 4.383 min, m/z = 911.1 (M+1). ¹H NMR: EW49333-60-P1A (400 MHz, CDCl₃) δ 4.77 - 4.98 (m, 2H), 4.10 (t, J = 6.8 Hz, 2H), 2.72 - 2.84 (m, 1H), 2.57 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.23 - 2.34 (m, 6H), 2.21 (s, 6H), 1.73 - 1.88 (m, 2H), 1.59 - 1.68 (m, 7H), 1.44 - 1.56 (m, 9H), 1.18 - 1.39 (m, 56H), 0.84 - 0.92 (m, 12H).

Preparation of compound SL73

Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol.

SL73 (650 mg, colorless oil). Special LCMS: EW49594-40-P1A2, Rt = 2.308 min, m/z = 770.9 (M+1). ¹H NMR: EW49594-40-P1A (400 MHz, CDCl₃) δ 4.10 (t, *J* = 6.8 Hz, 2H), 3.98 (d, *J* = 5.8 Hz, 4H), 2.70 - 2.80 (m, 1H), 2.60 - 2.70 (m, 2H), 2.50 (t, *J* = 7.4 Hz, 2H), 2.30 - 2.40 (m, 5H), 2.20 - 2.30 (m, 1H), 2.21 (s, 6H), 1.70 - 1.90 (m, 2H), 1.50 - 1.70 (m, 10H), 1.20 - 1.40 (m, 40H), 0.80 - 1.00 (m, 12H).

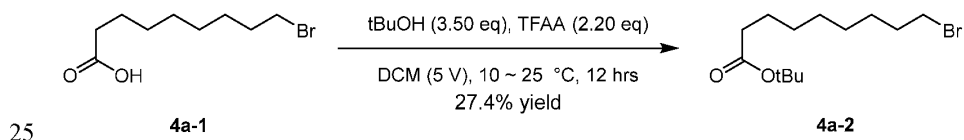
Preparation of compound **SL74**

Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL74** (2.00 g, colorless oil). Special LCMS: EW49333-61-P1B, Rt = 4.268 min, m/z = 883.0 (M+1). ¹H NMR: EW49333-61-P1A (400 MHz, CDCl₃) δ 4.10 (t, *J* = 6.8 Hz, 2H), 3.97 (d, *J* = 5.8 Hz, 4H), 2.73 - 2.83 (m, 1H), 2.59 - 2.69 (m, 2H), 2.50 (t, *J* = 7.4 Hz, 2H), 2.23 - 2.35 (m, 6H), 2.21 (s, 6H), 1.74 - 1.88 (m, 2H), 1.52 - 1.65 (m, 10H), 1.21 - 1.41 (m, 56H), 0.82 - 0.94 (m, 12H).

Preparation of compound **SL75**

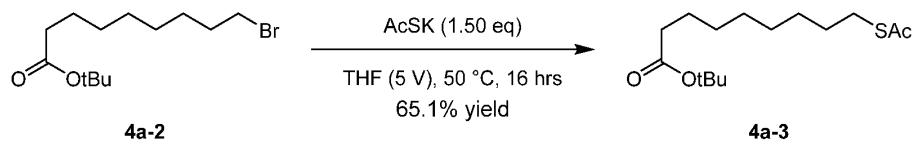
Procedure as per preparation of compound **SL70**, with reaction of relevant alkanol. **SL75** (2.20 g, colorless oil). Special LCMS: EW49594-41-P1A1, Rt = 4.268 min, m/z = 883.0 (M+1). ¹H NMR: EW49594-41-P1A (400 MHz, CDCl₃) δ 4.09 (t, *J* = 6.8 Hz, 2H), 3.97 (d, *J* = 5.6 Hz, 4H), 2.70 - 2.80 (m, 1H), 2.60 - 2.70 (m, 2H), 2.50 (t, *J* = 7.4 Hz, 2H), 2.20 - 2.40 (m, 6H), 2.20 (s, 6H), 1.70 - 1.90 (m, 2H), 1.50 - 1.70 (m, 10H), 1.27 (br s, 72H), 0.88 (t, *J* = 6.8 Hz, 12H).

General procedure for preparation of compound **4a-2**



To a solution of compound **4a-1** (460 g, 1.94 mol, 1.00 eq) in dichloromethane (2.30 L) was added t-BuOH (503 g, 6.79 mol, 649 mL, 3.50 eq) and TFAA (896 g, 4.27 mol, 593 mL, 2.20 eq) at 10 °C. The mixture was stirred at 25 °C for 12 h. TLC: (Petroleum ether/ethyl acetate = 10/1) indicated compound **4a-1** was remained, and two new spots
 5 (R_f = 0.00, 0.70) formed. The reaction was poured into aqueous solution of H₂O (1.00 L). The resulting solution was extracted with dichloromethane (1.00 L × 2). The combined organic phase was washed with brine (1.00 L × 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 1/0 to 10/1) (TLC:
 10 Petroleum ether/Ethyl acetate = 10/1, R_f = 0.70). Compound **4a-2** (156 g, 532 mmol, 27.4% yield) was obtained as a yellow oil. ¹H NMR: EW49525-1-P1A, (400 MHz, CDCl₃) δ 3.40 (t, J = 6.8 Hz, 2H), 2.20 (t, J = 7.2 Hz, 2H), 1.83 – 1.86 (m, 2H), 1.57 – 1.59 (m, 2H), 1.44 (s, 9H), 1.41 – 1.43 (m, 2H), 1.30 (s, 6H).

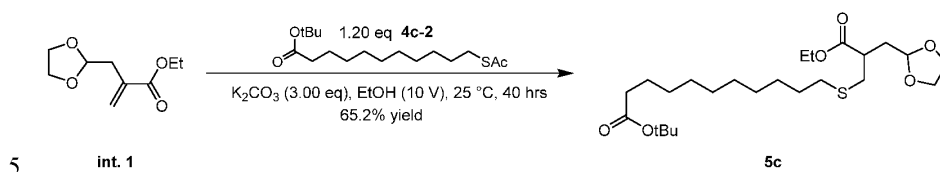
15 *General procedure for preparation of compound 4a-3*



To a solution of compound **4a-2** (221 g, 754 mmol, 1.00 eq) in tetrahydrofuran (1.00 L) was added potassium ethanethioate (129 g, 1.13 mol, 1.50 eq). The mixture was stirred at 50 °C for 16 h. LCMS: (EW49333-24-P1A, P1: Rt = 0.655 min) showed compound
 20 **4a-2** was consumed and one main peak with desired mass (Rt = 0.655 min, m/z = 311.0 (M+23)) was detected. The reaction was poured into aqueous solution of H₂O (1.60 L). The resulting solution was extracted with ethyl acetate (1.60 L × 3). The combined organic phase was washed with brine (1.60 L × 3), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give the crude product. The residue was purified by
 25 column chromatography (SiO₂, Petroleum ether/ethyl acetate = 10/1 to 8/1, (TLC: Petroleum ether/ethyl acetate = 8/1, P: R_f = 0.67)). Compound **4a-3** (167 g, 491 mmol, 65.1% yield, 84.8% purity) was obtained as a brown oil. LCMS: EW49333-24-P1D, Rt = 0.658 min, m/z = 311.1 (M+23). ¹H NMR: EW49333-24-P1B, (400 MHz, CDCl₃) δ

2.86 (t, J = 7.2 Hz, 2H), 2.33 (s, 3H), 2.20 (t, J = 7.6 Hz, 2H), 1.54 – 1.56 (m, 4H), 1.45 (s, 9H), 1.29 – 1.36 (m, 8H).

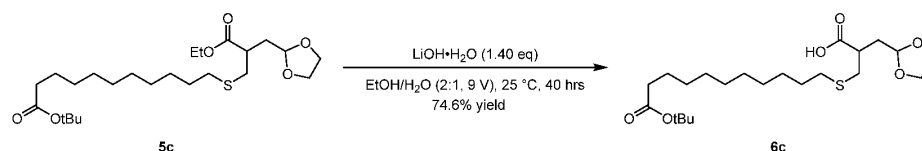
General procedure for preparation of compound 5c



To a solution of compound **int. 1** (8.70 g, 46.7 mmol, 1.00 eq) and compound **4c-2** (17.8 g, 56.1 mmol, 1.20 eq) in ethyl alcohol (90.0 mL) was added K₂CO₃ (19.4 g, 140 mmol, 3.00 eq). The reaction was stirred at 25 °C for 40 h. LCMS: (EW49594-4-P1B) showed desired mass (R_t = 0.714, m/z = 461.2 (M+1)) was detected. The reaction mixture was filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 330 g Sepa Flash ® Silica Flash Column, Eluent of 0 ~ 8% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 5/1, R_f = 0.28). Compound **5c** (17.2 g, 30.5 mmol, 65.2% yield, 81.6% purity) was obtained as colorless liquid. LCMS: EW49594-4-P1C1, R_t = 0.721, m/z = 461.3 (M+1). ¹H NMR: EW49594-4-P1A, (400 MHz, CDCl₃) δ 4.88 (t, *J* = 4.4 Hz, 1H), 4.00 - 4.20 (m, 2H), 3.80 - 4.00 (m, 2H), 3.70 - 3.80 (m, 2H), 2.60 - 2.80 (m, 2H), 2.61 - 2.63 (m, 1H), 2.44 (t, *J* = 7.4 Hz, 2H), 2.13 (t, *J* = 7.6 Hz, 2H), 2.06 (ddd, *J* = 4.2, 8.4, 14.2 Hz, 1H), 1.89 (td, *J* = 4.4, 14.4 Hz, 1H), 1.40 - 1.50 (m, 4H), 1.37 (s, 9H), 1.20 - 1.30 (m, 15H).

20

General procedure for preparation of compound 6c



To a solution of compound **5c** (17.0 g, 36.9 mmol, 1.00 eq) in ethyl alcohol (100 mL)/H₂O (50.0 mL) was added LiOH·H₂O (1.86 g, 44.3 mmol, 1.20 eq). The reaction mixture was stirred at 25 °C for 16 h. Additional LiOH·H₂O (310 mg, 7.38 mmol, 0.20 eq) was added. The reaction mixture was still stirred at 25 °C for 24 h. LCMS: (EW49594-8-P1C) showed desired mass (Rt = 0.644, m/z = 455.2 (M+23)) was detected and 1.30% of compound **5c** was remained. The reaction mixture was concentrated reduced pressure to remove ethyl alcohol, then added H₂O (50.0 mL), dropwise HCl (1M) to adjust the pH = ~3 and extracted with ethyl acetate (100 mL × 3). The combined organic layers were dried over MgSO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 330 g Sepa Flash ® Silica Flash Column, Eluent of 0 ~ 25% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 1/1, R_f = 0.50). Compound **6c** (12.2 g, 27.5 mmol, 74.6% yield, 97.6% purity) was obtained as colorless oil. LCMS: EW49594-8-P1C1, Rt = 0.633 min, m/z = 455.2 (M+23). ¹H NMR: EW49594-8-P1A, (400 MHz, CDCl₃) δ 4.93 (t, J = 4.2 Hz, 1H), 3.80 - 4.00 (m, 2H), 3.81 - 3.83 (m, 2H), 2.70 - 2.80 (m, 2H), 2.60 - 2.70 (m, 1H), 2.46 (t, J = 7.4 Hz, 2H), 2.13 (t, J = 7.6 Hz, 2H), 2.00 - 2.10 (m, 1H), 1.90 - 2.00 (m, 1H), 1.51 - 1.53 (m, 4H), 1.37 (s, 9H), 1.20 - 1.30 (m, 12H).

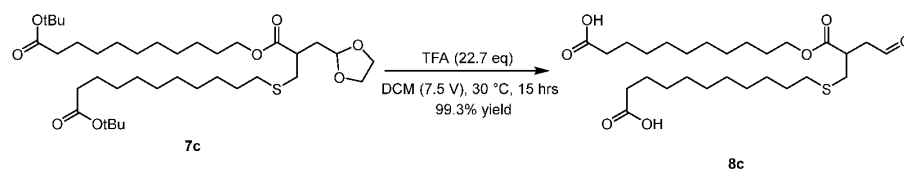
20 *General procedure for preparation of compound 7c*



To a solution of compound **6c** (12.2 g, 28.2 mmol, 1.00 eq) and compound **4c-1** (9.97 g, 31.0 mmol, 1.10 eq) in dimethylformamide (120 mL) was added K₂CO₃ (11.7 g, 84.6 mmol, 3.00 eq). The reaction was stirred at 80 °C for 2 h. LCMS: (EW49594-11-P1A) showed desired mass (Rt = 0.877 min, m/z = 695.5 (M=1)) was detected. The reaction mixture was added H₂O (100 mL) and extracted with ethyl acetate (100 mL × 2). The combined organic was washed with brine, dried over MgSO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel

chromatography (ISCO®; 330 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 8% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 5/1, R_f = 0.41). Compound **7c** (18.4 g, 27.3 mmol, 97.0% yield) was obtained as colorless liquid. ¹H NMR: EW49594-11-P1A, (400 MHz, CDCl₃) δ 4.87 (t, J = 4.4 Hz, 1H), 4.00 - 4.10 (m, 2H), 3.80 - 3.90 (m, 2H), 3.70 - 3.80 (m, 2H), 2.70 - 2.80 (m, 2H), 2.61 - 2.63 (m, 1H), 2.43 (t, J = 7.4 Hz, 2H), 2.13 (t, J = 7.6 Hz, 4H), 2.00 - 2.10 (m, 1H), 1.90 - 2.00 (m, 1H), 1.40 - 1.60 (m, 8H), 1.37 (s, 18H), 1.20 - 1.30 (m, 24H).

General procedure for preparation of compound 8c



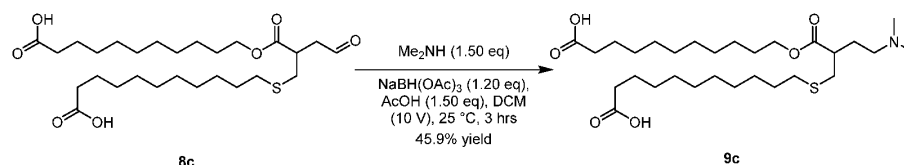
10

To a solution of compound **7c** (16.0 g, 23.8 mmol, 1.00 eq) in dichloromethane (120 mL) was added TFA (61.4 g, 538 mmol, 40.0 mL, 22.7 eq). The reaction was stirred at 30 °C for 15 h. LCMS: (EW19594-15-P1A) showed desired mass (R_t = 0.608 min, m/z = 517.3(M+1)) was detected. The reaction was concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 330 g Sepa Flash ® Silica Flash Column, Eluent of 0 ~ 30% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 1/1, P1: R_f = 0.55). Compound **8c** (12.2 g, 23.6 mmol, 99.3% yield) was obtained as white solid. ¹H NMR: EW49594-15-P1A, (400 MHz, CDCl₃) δ 9.73 (s, 1H), 8.72 (br s, 2H), 4.00 - 4.10 (m, 2H), 3.00 - 3.10 (m, 1H), 2.70 - 2.90 (m, 3H), 2.58 (dd, J = 8.4, 13.4 Hz, 1H), 2.42 (t, J = 7.4 Hz, 2H), 2.28 (t, J = 7.6 Hz, 4H), 1.40 - 1.60 (m, 8H), 1.20 - 1.30 (m, 24H).

20

General procedure for preparation of compound 9c

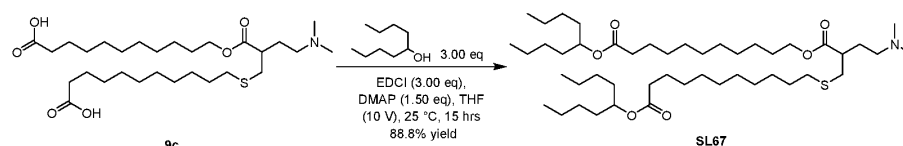
152



To a solution of compound **8c** (12.2 g, 23.6 mmol, 1.00 eq) in dichloromethane (120 mL) was added AcOH (2.13 g, 35.4 mmol, 2.03 mL, 1.50 eq) and Me₂NH (2 M, 17.7 mL, 1.50 eq). The solution was stirred at 25 °C for 2 h, then added NaBH(OAc)₃ (6.00 g, 28.3 mmol, 1.20 eq). The solution was stirred at 25 °C for 1 h. LCMS: (EW49594-16-P1A) showed desired mass (Rt = 0.507 min, m/z = 546.4 (M+1)) was detected. The reaction was poured into aqueous solution of NaHCO₃ (aq. 120 mL). The resulting solution was extracted with dichloromethane (100 mL × 2). Then acidified the aqueous phase with 2M HCl to pH = 3. The aqueous phase was extracted with dichloromethane (200 mL × 2). The combined organic layers were dried over MgSO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 120 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 5% methyl alcohol/dichloromethane gradient @ 85 mL/min, dichloromethane/methyl alcohol = 10/1, R_f = 0.50). Compound **9c** (5.90 g, 10.8 mmol, 45.9% yield, 100% purity) was obtained as white solid. LCMS: EW49594-16-P1C1: Rt = 0.510 min, m/z = 546.4 (M+1). ¹H NMR: EW49594-16-P1B, (400 MHz, CDCl₃) δ 10.4 (br s, 1H), 4.01 - 4.03 (m, 2H), 3.60 - 3.80 (m, 2H), 2.90 - 3.00 (m, 1H), 2.70 - 2.90 (m, 2H), 2.60 - 2.70 (m, 8H), 2.40 - 2.50 (m, 2H), 2.10 - 2.30 (m, 4H), 2.01 - 2.04 (m, 2H), 1.40 - 1.60 (m, 8H), 1.20 - 1.30 (m, 24H).

20

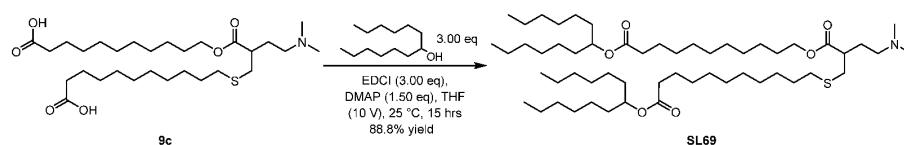
Preparation of compound **SL67**



To a solution of compound **9c** (2.5 g, 4.58 mmol, 1.00 eq) and nonan-5-ol (1.98 g, 13.7 mmol, 3.00 eq) in tetrahydrofuran (25.0 mL) was added EDCI (2.63 g, 13.7 mmol, 3.00

eq) and DMAP (839 mg, 6.87 mmol, 1.50 eq). The solution was stirred at 25 °C for 15 h. LCMS: (EW49594-19-P1A) showed desired mass ($R_t = 2.328$ min, $m/z = 798.6$ (M+1)) was detected. The reaction mixture was added NaHCO_3 (aq. 25.0 mL) and extracted with ethyl acetate (20.0 mL $\times$ 2). The combined organic was dried over MgSO_4 ,
 5 filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min, dichloromethane/methyl alcohol = 10/1, $R_f = 0.53$). Compound **SL-67** (2.80 g, 3.50 mmol, 76.5% yield, 99.9% purity) was obtained as a yellow oil. 1.8 g of compound **SL-**
 10 **67** was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% tetrahydrofuran/petroleum ether gradient @ 65 mL/min, petroleum ether/tetrahydrofuran = 1/1, $R_f = 0.21$) to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min,
 15 petroleum ether/tetrahydrofuran = 1/1, $R_f = 0.21$). Compound **SL67** (1.60 g, 2.00 mmol, 88.8% yield, 99.9% purity) was obtained as yellow oil. LCMS: EW49594-42-P1A1: $R_t = 2.485$ min, $m/z = 798.9$ (M+1). $^1\text{H NMR}$: EW49594-42-P1A, (400 MHz, CDCl_3) δ 4.86 (q, $J = 6.4$ Hz, 2H), 4.08 (t, $J = 6.8$ Hz, 2H), 2.70 - 2.80 (m, 1H), 2.60 - 2.70 (m, 2H), 2.49 (t, $J = 7.4$ Hz, 2H), 2.20 - 2.30 (m, 6H), 2.19 (s, 6H), 1.70 - 1.90 (m, 2H), 1.40 -
 20 1.60 (m, 16H), 1.20 - 1.40 (m, 40H), 0.88 (t, $J = 6.8$ Hz, 12H).

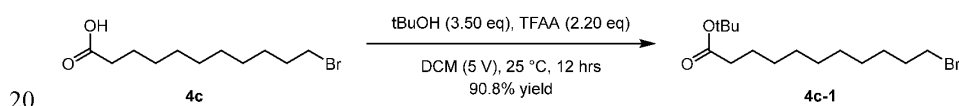
Preparation of compound **SL69**



To a solution of compound **9c** (2.50 g, 4.58 mmol, 1.00 eq) and tridecan-7-ol (2.75 g,
 25 13.7 mmol, 3.00 eq) in tetrahydrofuran (25.0 mL) was added EDCI (2.63 g, 13.7 mmol, 3.00 eq) and DMAP (839 mg, 6.87 mmol, 1.50 eq). The solution was stirred at 25 °C for 15 h. LCMS: (EW49594-20-P1A) showed desired mass ($R_t = 2.655$ min, $m/z = 910.7$ (M+1)) was detected. The reaction mixture was added NaHCO_3 (aq. 25.0 mL) and

extracted with ethyl acetate (20.0 mL $\times$ 2). The combined organic was dried over MgSO_4 , filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min, dichloromethane/methyl alcohol = 10/1, R_f = 0.53). Compound **SL69** (2.80 g, 3.50 mmol, 76.5% yield, 99.9% purity) was obtained as a colorless oil. Compound **SL69** (1.80 g, 1.98 mmol, 1.00 eq) was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 15% tetrahydrofuran/petroleum ether gradient @ 65 mL/min, petroleum ether/tetrahydrofuran = 1/1, R_f = 0.23) to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min, petroleum ether/tetrahydrofuran = 1/1, R_f = 0.23). Compound **SL69** (1.60 g, 1.76 mmol, 88.8% yield, 99.9% purity) was obtained as yellow oil. LCMS: EW49594-44-P1A1: R_t = 2.949 min, m/z = 911.0 (M+1). ^1H NMR: EW49594-44-P1B, (400 MHz, CDCl_3) δ 4.87 (q, J = 6.2 Hz, 2H), 4.09 (t, J = 6.8 Hz, 2H), 2.70 - 2.80 (m, 1H), 2.60 - 2.70 (m, 2H), 2.50 (t, J = 7.4 Hz, 2H), 2.20 - 2.40 (m, 6H), 2.20 (s, 6H), 1.70 - 1.90 (m, 2H), 1.50 - 1.70 (m, 16H), 1.27 (br d, J = 4.2 Hz, 56H), 0.80 - 1.00 (m, 12H).

General procedure for preparation of compound 4c-1



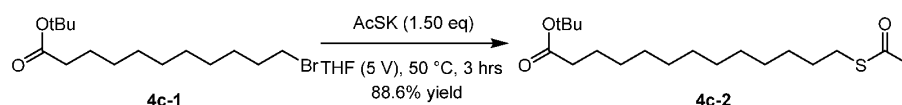
To a solution of compound **4c** (40.0 g, 151 mmol, 1.00 eq) in dichloromethane (200 mL) was added TFAA (69.7 g, 332 mmol, 46.1 mL, 2.20 eq) and t-BuOH (39.1 g, 528 mmol, 50.5 mL, 3.50 eq). The mixture was stirred at 25 °C for 12 h. TLC: (petroleum ether/ethyl acetate = 10/1) indicated compound **4c** was consumed completely and two new spots (R_f = 0.10, 0.70) formed. The reaction was poured into aqueous solution of H_2O (100 mL). The resulting solution was extracted with dichloromethane (100 mL $\times$ 2). The combined organic phase was washed with brine (100 mL $\times$ 2), dried over anhydrous Na_2SO_4 , and concentrated under vacuum to give a residue. The residue was purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 1/0 to 10/1) (TLC: petroleum

25

ether/ethyl acetate = 10/1, R_f = 0.70). Compound **4c-1** (44.0 g, 137 mmol, 90.8% yield) was obtained as a white oil. ^1H NMR: EW49144-10-P1A (400Hz, CDCl_3) δ 3.40 (t, J = 7.2 Hz, 2H), 2.20 (t, J = 7.6 Hz, 2H), 1.82 - 1.86 (m, 2H), 1.54 - 1.57 (m, 2H), 1.43 (s, 9H), 1.41 - 1.42 (m, 2H), 1.25 (s, 10H).

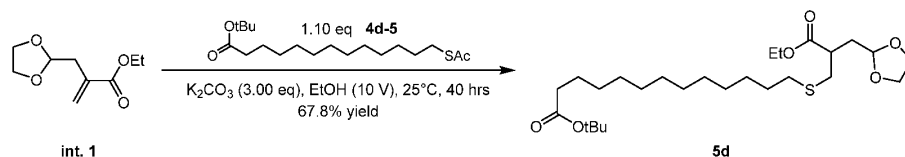
5

General procedure for preparation of compound 4c-2



To a solution of compound **4c-1** (20.0 g, 62.3 mmol, 1.00 eq) in tetrahydrofuran (100 mL) was added potassium; ethanethioate (10.7 g, 93.4 mmol, 1.50 eq). The mixture was stirred at 50 °C for 3h. TLC (petroleum ether/ethyl acetate = 10/1) indicated compound **4c-1** was consumed completely and one new spot (R_f = 0.60) formed. The reaction was poured into aqueous solution of H_2O (100 mL). The resulting solution was extracted with ethyl acetate (100 mL $\times$ 3). The combined organic phase was washed with brine (100 mL $\times$ 3), dried over anhydrous Na_2SO_4 , and concentrated under vacuum to give Compound **4c-2** (19.0 g, 55.14 mmol, 88.6% yield, 100% purity) as a colorless oil which was without purification. ^1H NMR: EW49525-3-P1A (400 MHz, CDCl_3) δ 2.86 (t, J = 7.2 Hz, 2H), 2.32 (s, 3H), 2.20 (t, J = 7.6 Hz, 2H), 1.55-1.58 (m, 4H), 1.44 (s, 9H), 1.24-1.28 (m, 16H).

20 *General procedure for preparation of compound 5d*



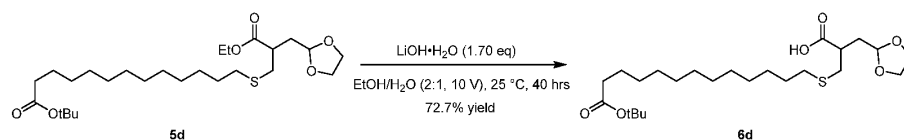
To a solution of compound **int.1** (5.00 g, 26.9 mmol, 1.00 eq) and compound **4d-5** (10.7 g, 29.5 mmol, 1.10 eq) in ethyl alcohol (50.0 mL) was added K_2CO_3 (11.1 g, 80.6 mmol, 3.00 eq). The reaction was stirred at 25 °C for 40 h. LCMS: (EW49594-18-P1A) showed desired mass (R_t = 0.754 min, m/z = 511.3 (M+1)) was detected. The reaction mixture

25

was filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 330 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 8% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 5/1, R_f = 0.28) to give compound **5d** (8.90 g, 18.2 mmol, 67.8% yield) as red liquid. ¹H NMR: EW49594-18-P1A, (400 MHz, CDCl₃) δ 4.88 (t, J = 4.4 Hz, 1H), 4.00 - 4.20 (m, 2H), 3.80 - 4.00 (m, 2H), 3.70 - 3.80 (m, 2H), 2.70 - 2.80 (m, 2H), 2.60 - 2.70 (m, 1H), 2.44 (t, J = 7.4 Hz, 2H), 2.13 (t, J = 7.6 Hz, 2H), 2.06 (ddd, J = 4.0, 8.4, 14.2 Hz, 1H), 1.91 - 1.93 (m, 1H), 1.40 - 1.60 (m, 4H), 1.37 (s, 9H), 1.10 - 1.30 (m, 19H).

10

General procedure for preparation of compound 6d



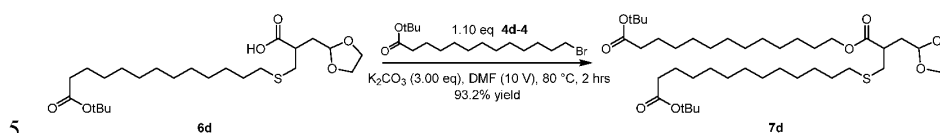
To a solution of compound **5d** (8.90 g, 18.2 mmol, 1.00 eq) in ethyl alcohol (60.0 mL) and H₂O (30.0 mL) was added LiOH·H₂O (1.07 g, 25.5 mmol, 1.40 q). The reaction mixture was stirred at 25 °C for 16 h. Then the mixture was added LiOH·H₂O (229 mg, 5.46 mmol, 0.30 eq). The reaction mixture was still stirred at 25 °C for 24 h. LCMS: (EW49594-21-P1B) showed compound **5d** was consumed completely and desired mass (R_t = 0.681 min, m/z = 483.2 (M+1)) was detected. The reaction mixture was concentrated reduced pressure to remove ethyl alcohol, added H₂O (50 mL), dropwise HCl (1M) to adjust the pH = ~3 and extracted with ethyl acetate (100 mL $\times$ 3). The combined organic layers were dried over MgSO₄, filtered, and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 120 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 25% ethyl acetate/petroleum ether gradient @ 100 mL/min, petroleum ether/ethyl acetate = 1/1, R_f = 0.50). Compound **6d** (6.10 g, 13.2 mmol, 72.7% yield) was obtained as orange oil. ¹H NMR: EW49594-21-P1A, (400 MHz, CDCl₃) δ 4.93 (t, J = 4.2 Hz, 1H), 3.80 - 4.00 (m, 2H), 3.70 - 3.80 (m, 2H), 2.70 - 2.80 (m, 2H), 2.60 - 2.70 (m, 1H), 2.46 (t, J =

25

157

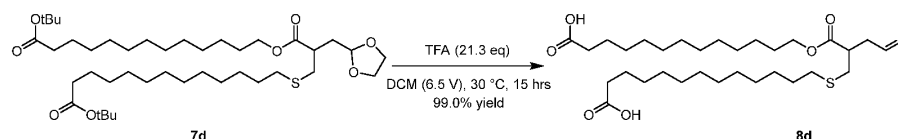
7.4 Hz, 2H), 2.13 (t, $J = 7.6$ Hz, 2H), 2.00 - 2.10 (m, 1H), 1.90 - 2.00 (m, 1H), 1.51 - 1.53 (m, 4H), 1.37 (s, 9H), 1.20 - 1.30 (m, 16H).

General procedure for preparation of compound 7d



To a solution of compound **6d** (6.10 g, 13.2 mmol, 1.00 *eq*) and compound **4d-4** (5.09 g, 14.6 mmol, 1.10 *eq*) in dimethylformamide (60.0 mL) was added K_2CO_3 (5.49 g, 39.7 mmol, 3.00 *eq*). The reaction was stirred at 80 °C for 2 h. TLC (petroleum ether/ethyl acetate = 5/1) showed compound **6d** ($R_f = 0.18$) was consumed completely and new spot ($R_f = 0.62$) was detected. The reaction mixture was added H_2O (100 mL) and extracted with ethyl acetate (100 mL $\times$ 2). The combined organic was washed with brine (100 mL), dried over $MgSO_4$, filtered, and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 120 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 8% ethyl acetate/petroleum ether gradient @ 85 mL/min, petroleum ether/ethyl acetate = 5/1, $R_f = 0.62$). Compound **7d** (9.00 g, 12.3 mmol, 93.2% yield) was obtained as yellow oil. 1H NMR: EW49594-24-P1A, (400 MHz, $CDCl_3$) δ 4.87 (t, $J = 4.4$ Hz, 1 H), 4.00 - 4.10 (m, 2 H), 3.80 - 3.90 (m, 2 H), 3.70 - 3.80 (m, 2 H), 2.70 - 2.80 (m, 2 H), 2.61 - 2.63 (m, 1 H), 2.43 (t, $J = 7.4$ Hz, 2 H), 2.13 (t, $J = 7.6$ Hz, 4 H), 2.00 - 2.10 (m, 1 H), 1.91 - 1.92 (m, 1 H), 1.50 - 1.60 (m, 8 H), 1.37 (s, 18 H), 1.20 - 1.30 (m, 32 H).

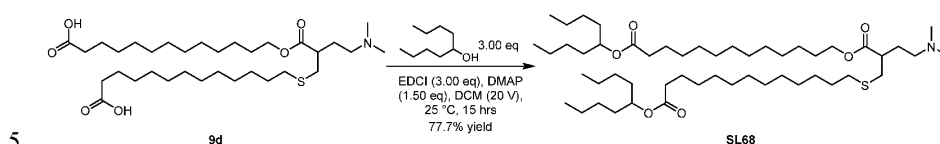
General procedure for preparation of compound 8d



To a solution of compound **7d** (9.00 g, 12.3 mmol, 1.00 *eq*) in dichloromethane (60.0 mL) was added TFA (30.0 g, 263 mmol, 19.6 mL, 21.3 *eq*). The reaction was stirred at

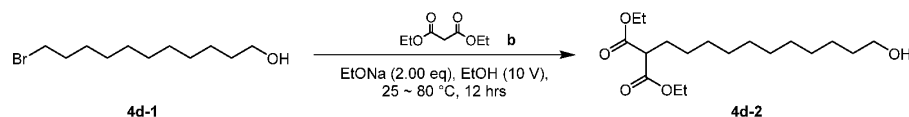
2.81 - 2.82 (m, 1H), 2.71 - 2.73 (m, 2H), 2.60 - 2.70 (m, 6H), 2.52 (br t, $J = 7.4$ Hz, 2H), 2.20 - 2.30 (m, 4H), 2.00 - 2.10 (m, 2H), 1.50 - 1.70 (m, 8H), 1.30 (br s, 32H).

Preparation of compound **SL68**

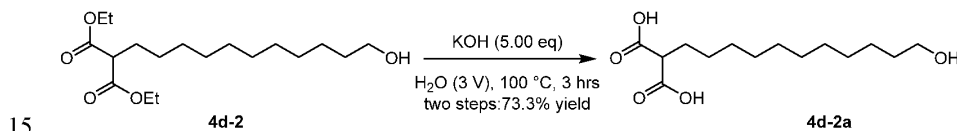


To a solution of compound **9d** (3.00 g, 4.98 mmol, 1.00 eq) and nonan-5-ol (2.16 g, 15.0 mmol, 3.00 eq) in tetrahydrofuran (60.0 mL) was added EDCI (2.87 g, 15.0 mmol, 3.00 eq) and DMAP (913 mg, 7.48 mmol, 1.50 eq). The solution was stirred at 25 °C for 15 h. LCMS: (EW49594-32-P1A) showed desired mass ($R_t = 2.050$, $m/z = 855.2$ ($M+1$))

- 10 was detected. The reaction mixture was added NaHCO_3 (aq., 50.0 mL) and extracted with ethyl acetate (50.0 mL $\times$ 2). The combined organic was dried over MgSO_4 , filtered, and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min, dichloromethane :
- 15 methyl alcohol = 10 : 1, $R_f = 0.53$). Compound **SL68** (1.80 g, 2.10 mmol, 42.2% yield, 99.8% purity) was obtained as a yellow oil. Compound **SL68** (1.80 g, 2.11 mmol, 1.00 eq) was purified by flash silica gel chromatography (ISCO ®; 80 g Sepa Flash ® Silica Flash Column, Eluent of 0 ~ 15% tetrahydrofuran/petroleum ether gradient @ 65 mL/min, petroleum ether/tetrahydrofuran = 1/1, $R_f = 0.20$) to give a residue. The residue
- 20 was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash ® Silica Flash Column, Eluent of 0 ~ 1% methyl alcohol/dichloromethane gradient @ 65 mL/min, petroleum ether/tetrahydrofuran = 1/1, $R_f = 0.20$). Compound **SL68** (1.40 g, 1.64 mmol, 77.7% yield, 99.9% purity) was obtained as yellow oil. LCMS: EW49594-43-P1A1, $R_t = 2.804$ min, $m/z = 855.0$ ($M+1$). ^1H NMR: EW49594-43-P1A, (400 MHz, CDCl_3) δ 4.87
- 25 (q, $J = 6.4$ Hz, 2H), 4.09 (t, $J = 6.8$ Hz, 2H), 2.70 - 2.80 (m, 1H), 2.60 - 2.70 (m, 2H), 2.50 (t, $J = 7.4$ Hz, 2H), 2.20 - 2.30 (m, 6H), 2.20 (s, 6H), 1.70 - 1.90 (m, 2H), 1.50 - 1.70 (m, 16H), 1.20 - 1.40 (m, 46H), 0.88 (t, $J = 6.8$ Hz, 12H).

General procedure for preparation of compound 4d-2

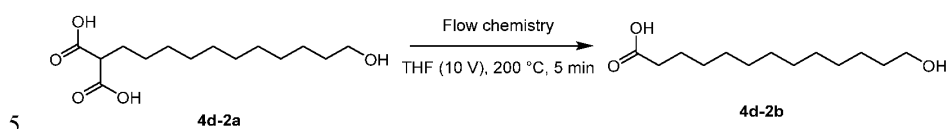
To a solution of compound **b** (191 g, 1.19 mol, 181 mL, 2.00 eq) and EtONa (406 g, 1.19 mol, 20% purity, 2.00 eq) in ethyl alcohol (1.50 L) was added drop-wise compound **4d-1** (150 g, 597 mmol, 1.00 eq) at 25 °C. Then the mixture was stirred at 80 °C for 12 h. TLC (Petroleum ether/Ethyl acetate = 5/1) showed compound **4d-1** (R_f = 0.40) was consumed and a new main spot (R_f = 0.30) was formed. The reaction was concentrated under vacuum to give a residue. The residue was poured into 3000 mL of ice water and extracted with methyl tertiary butyl ether (3.00 L $\times$ 2). The combined organic phase was washed with brine (2.00 L), dried with anhydrous Na_2SO_4 , filtered, and concentrated in vacuum to give compound **4d-2** (197.3 g, crude) as a yellow oil which was used directly without purification.

General procedure for preparation of compound 4d-2a

To a solution of compound **4d-2** (197 g, 597 mmol, 1.00 eq) in H_2O (600 mL) was added KOH (167 g, 2.99 mol, 5.00 eq). Then the mixture was stirred at 100 °C for 3 h. TLC (petroleum ether/ethyl acetate = 3/1) showed compound **4d-2** (R_f = 0.45) was consumed and a new main spot (R_f = 0.0) was formed. The reaction was cooled to 20 °C and poured into water (500 mL) and stirred for 10 min. The aqueous phase was extracted with methyl tertiary butyl ether (300 mL $\times$ 2). The combined aqueous phase was adjusted with 12 M HCl to pH = 1 ~ 2. The aqueous phase was extracted with methyl tertiary butyl ether (800 mL $\times$ 2), washed with brine (200 mL), dried with anhydrous Na_2SO_4 , filtered, and concentrated in vacuum to give compound **4d-2a** (120 g, 437 mmol, 73.3% yield) as an off-white solid which was used directly without purification. $^1\text{H NMR}$: EW49511-2-P1A

(400 MHz, dimethylsulfoxide- d_6) δ 12.6 (s, 2H), 3.75 - 4.03 (m, 1H), 3.38 - 3.44 (m, 2H), 3.25 (s, 1H), 1.67 - 1.69 (m, 2H), 1.37 - 1.39 (m, 2H), 1.23 (s, 16H).

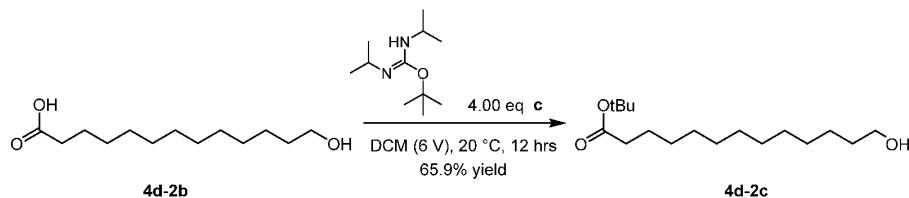
General procedure for preparation of compound 4d-2b



To a solution of compound **4d-2a** (120 g, 437 mmol, 1.00 eq) in tetrahydrofuran (1.20 L). The mixture was stirred at 200 °C for 5 min through flow chemistry. TLC (petroleum ether/ethyl acetate = 0/1) showed the compound **4d-2a** (R_f = 0.20) was consumed and a new spot (R_f = 0.50) was formed. The mixture was concentrated under vacuum to give compound **4d-2b** (100 g, crude) as a yellow solid which was used directly without purification.

10

General procedure for preparation of compound 4d-2c

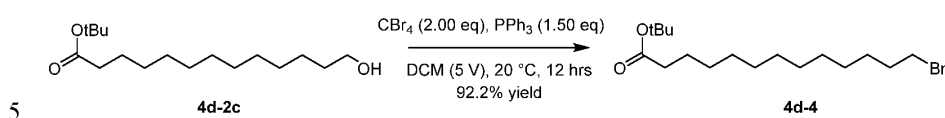


15 To a solution of compound **4d-2b** (75.0 g, 326 mmol, 1.00 eq) in dichloromethane (450 mL) was added compound **c** (261 g, 1.30 mol, 4.00 eq). The reaction was stirred at 20 °C for 12 h. TLC (petroleum ether/ethyl acetate = 3/1) showed compound **4d-2b** (R_f = 0.20) was consumed and a new spot (R_f = 0.50) was formed. The reaction was filtered and the cake was washed with dichloromethane (500 mL) and the filtrate was

20 concentrated under vacuum. The residue was purified by column chromatography (SiO₂, petroleum ether/ethyl acetate = 50/1 to 5/1, TLC (petroleum ether/ethyl acetate = 3/1, R_f = 0.50)) to give compound **4d-2c** (61.5 g, 215 mmol, 65.9% yield) as a yellow oil. ¹H

NMR: EW49511-12-P1A (400 MHz, CDCl₃) δ 3.64 (t, J = 6.4 Hz, 2H), 2.20 (t, J = 6.4 Hz, 2H), 1.55 - 1.59 (m, 4H), 1.44 (s, 9H), 1.27 - 1.34 (m, 16H).

General procedure for preparation of compound 4d-4

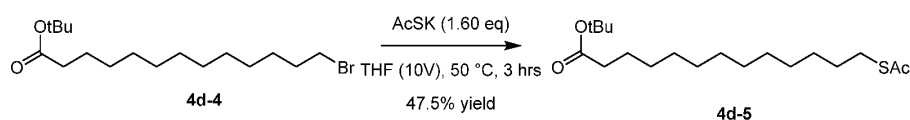


To a solution of compound **4d-2c** (61.0 g, 213 mmol, 1.00 eq) in dichloromethane (310 mL) was added CBr₄ (141 g, 426 mmol, 2.00 eq) and PPh₃ (83.8 g, 319 mmol, 1.50 eq). The solution was stirred at 20 °C for 12 h. TLC: (petroleum ether/ethyl acetate = 10/1) showed the compound **4d-2c** (R_f = 0.30) was consumed and a new spot (R_f = 0.70) was formed. The reaction was concentrated under vacuum and filtered and the filter cake was washed with petroleum ether (2.00 L). The filtrate was concentrated under vacuum. The residue was purified by silica gel chromatography (petroleum ether/ethyl acetate = 100/0 to 30/1, R_f = 0.70) to give compound **4d-4** (68.6 g, 196 mmol, 92.2% yield) as a yellow oil. ¹H NMR: EW49511-15-P1A (400 MHz, CDCl₃) δ 3.41 (t, J = 6.8 Hz, 2H), 2.20 (t, J = 7.6 Hz, 2H), 1.83 - 1.87 (m, 2H), 1.55 - 1.58 (m, 2H), 1.44 (s, 9H), 1.40 - 1.42 (m, 2H), 1.24 - 1.27 (m, 14H).

10

15

General procedure for preparation of compound 4d-5

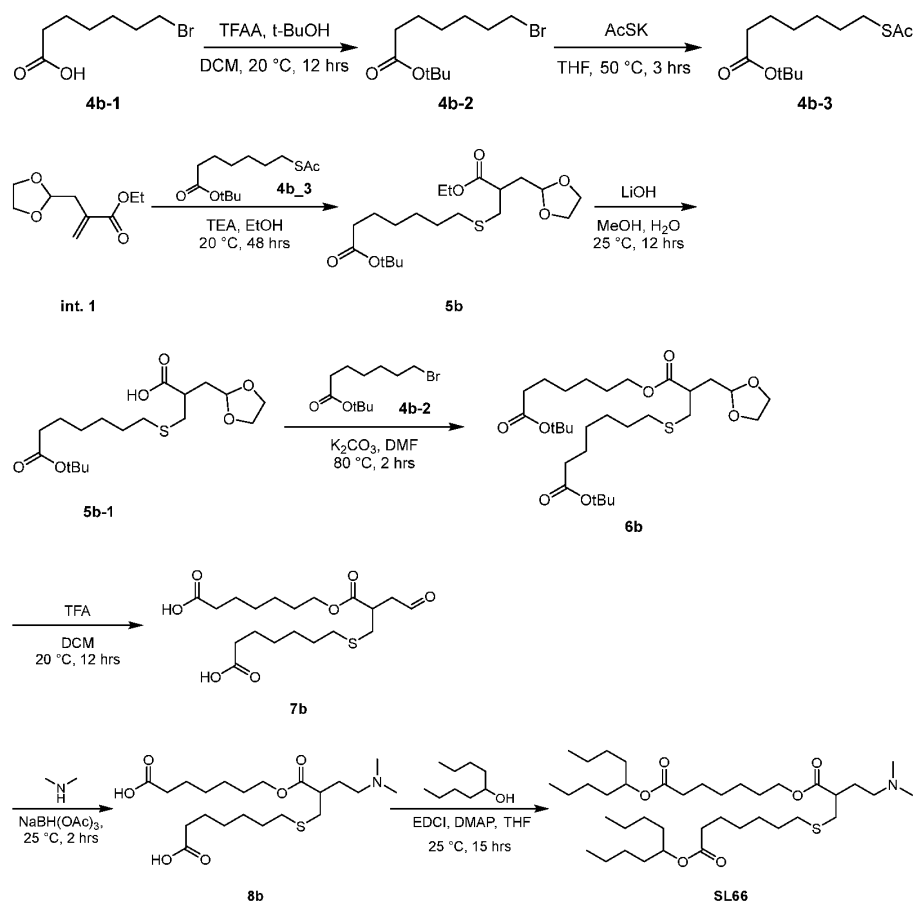


20 To a solution of compound **4d-4** (21.8 g, 62.4 mmol, 1.00 eq) in tetrahydrofuran (210 mL) was added potassium ethanethioate (11.4 g, 99.8 mmol, 1.60 eq). The mixture was stirred at 50 °C for 3 h. TLC: (petroleum ether/ethyl acetate = 10/1) showed compound **4d-4** (R_f = 0.50) was consumed and a new spot (R_f = 0.70) was formed. The reaction was poured into aqueous solution of H₂O (200 mL). The resulting solution was extracted with

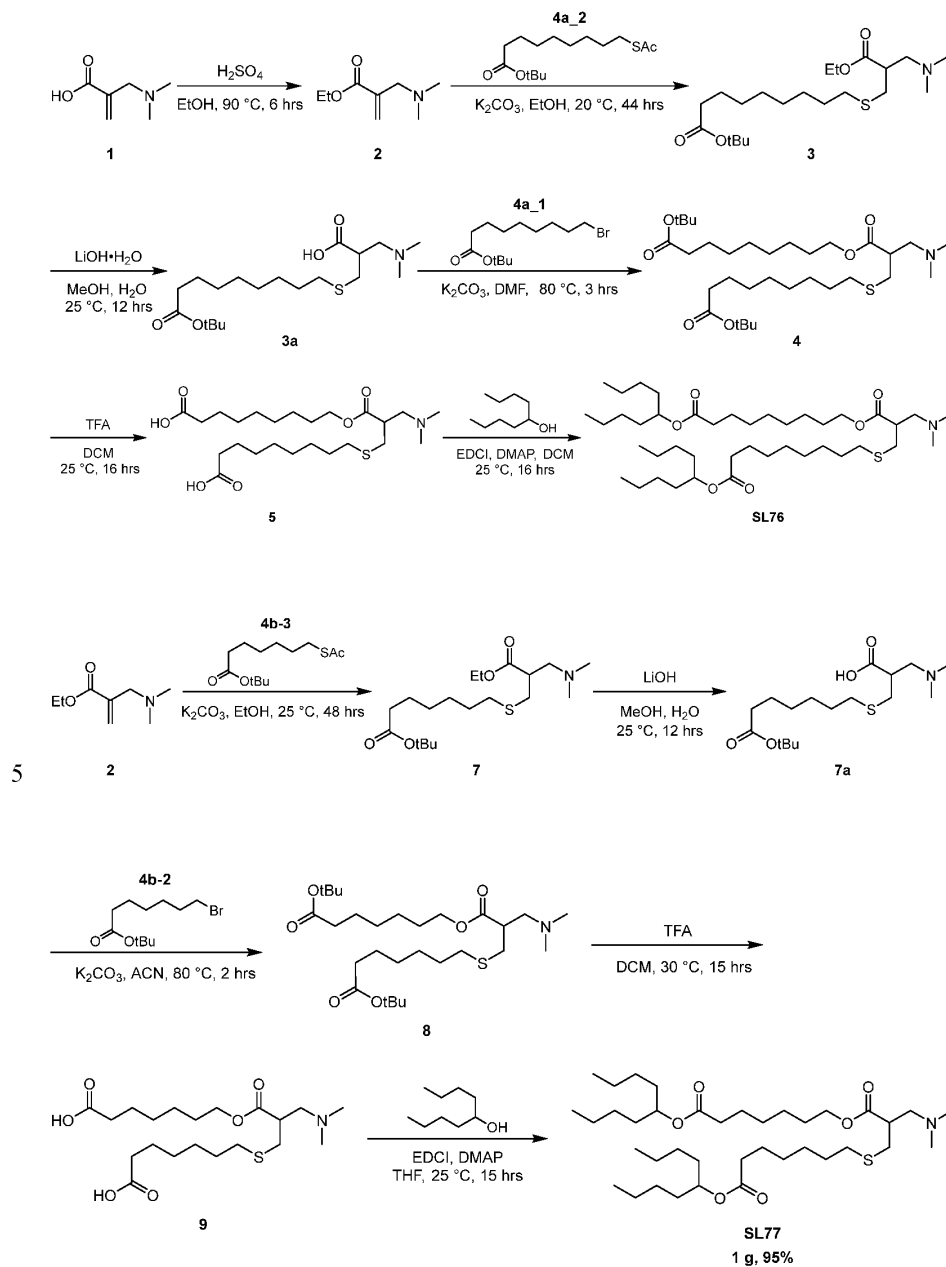
25 ethyl acetate (200 mL $\times$ 3). The combined organic phase was washed with brine (200

mL $\times$ 3), dried over anhydrous Na_2SO_4 , and concentrated under vacuum to give compound **4d-5** (10.8 g, 29.7 mmol, 47.5% yield, 94.6% purity) as a yellow oil which was used directly without purification. LCMS: EW49511-14-P1C, Rt = 0.759 min, m/z = 229.1 (M+1)⁺. ¹H NMR: EW49511-14-P1B (400 MHz, CDCl_3) δ 2.86 (t, J = 7.2 Hz, 2H), 2.32 (s, 3H), 2.20 (t, J = 7.6 Hz, 2H), 1.54 - 1.58 (4, 2H), 1.45 (s, 9H), 1.24 - 1.34 (m, 16H).

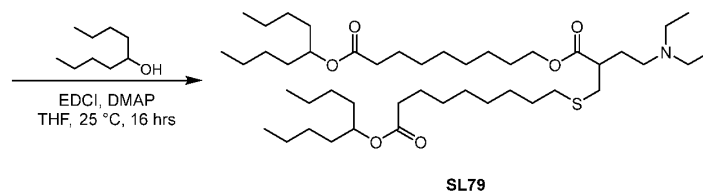
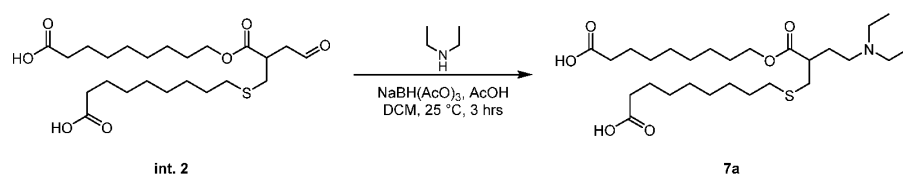
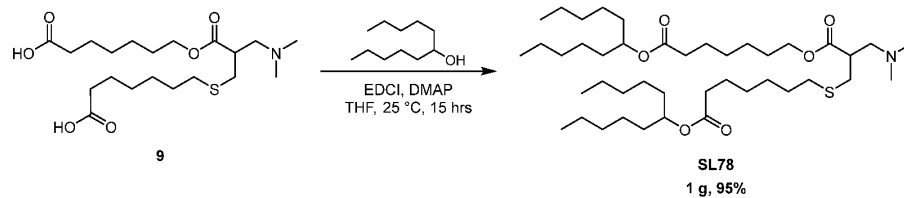
Example 1c: Synthetic Approach to Compounds SL66 and SL76-SL81



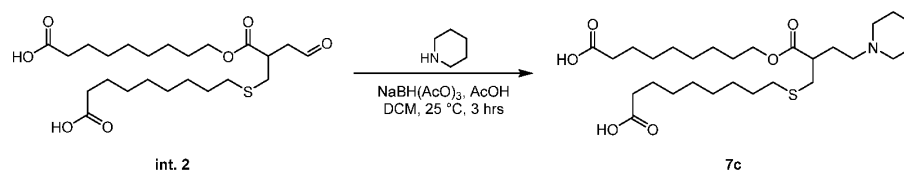
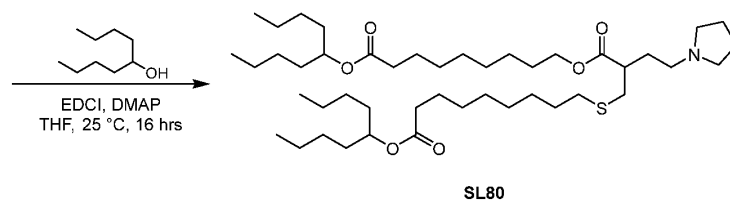
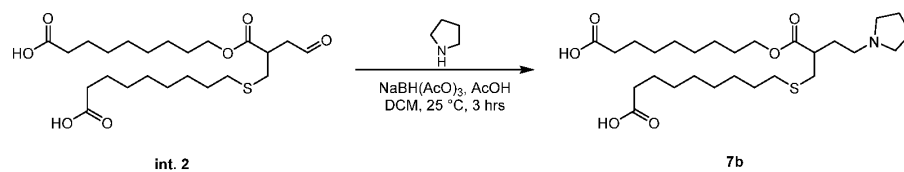
164



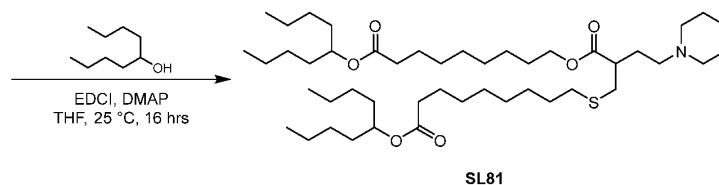
165



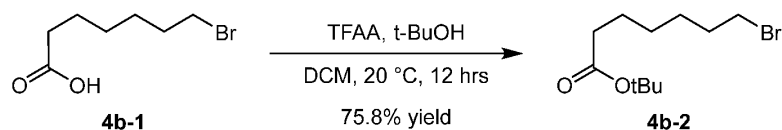
5



166

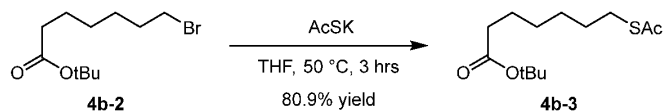


General procedure for preparation of compound 4b-2



- 5 To a solution of compound **4b-1** (185 g, 884 mmol, 1.00 eq) in dichloromethane (1.00 L) was added TFAA (408 g, 1.95 mol, 270 mL, 2.20 eq) and t-BuOH (229 g, 3.10 mol, 296 mL, 3.50 eq). The solution was stirred at 20 °C for 12 h. TLC (petroleum ether/ethyl acetate = 5/1) indicated one main spot (R_f = 0.80) was formed. The reaction mixture was quenched by NaHCO_3 solution (500 mL) and then extracted with ethyl acetate (500 mL
- 10 $\times$ 2). The combined organic layers were washed with brine (500 mL $\times$ 2), dried over Na_2SO_4 , filtered, and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO_2 , Dichloromethane, Petroleum ether/Ethyl acetate=5/1, R_f = 0.80) to give compound **4b-2** (178 g, 671 mmol, 75.8% yield) as a yellow oil. ^1H NMR: EW49391-9-P1A1 (400 MHz, CDCl_3) δ 3.37 - 3.41 (m, 2H), 2.18
- 15 - 2.22 (m, 2H), 1.83 - 1.85 (m, 2H), 1.53 - 1.65 (m, 2H), 1.35 - 1.48 (m, 11H), 1.28 - 1.32 (m, 2H).

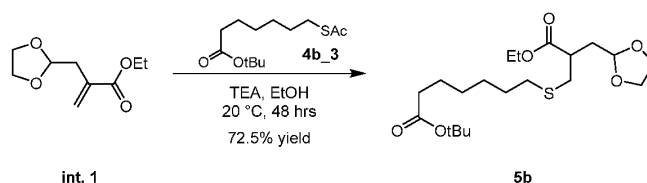
General procedure for preparation of compound 4b-3



- 20 To a solution of compound **4b-2** (150 g, 565 mmol, 1.00 eq) in tetrahydrofuran (750 mL) was added AcSK (119 g, 1.05 mol, 1.85 eq). The suspension was stirred at 50 °C for 3 h. TLC (petroleum ether/ethyl acetate = 10/1) showed one new spot (R_f = 0.70) was

formed. The reaction was poured into H₂O (200 mL) and extracted with ethyl acetate (200 mL × 2). The combined organic phase was washed with brine (200 mL × 2), dried over anhydrous Na₂SO₄ and concentrated under vacuum to give compound **4b-3** (140 g, 458 mmol, 80.9% yield, 85.2% purity) as a yellow oil. LCMS: EW49391-10-P1A1, Rt = 0.620 min, m/z = 283.0, M+Na⁺. ¹H NMR: EW49391-10-P1A2 (400 MHz, CDCl₃) δ 2.82 - 2.86 (m, 2H), 2.30 (s, 3H), 2.16 - 2.18 (m, 2H), 1.52 - 1.57 (m, 4H), 1.41 (s, 9H), 1.32 - 1.38 (m, 4H).

General procedure for preparation of compound 5b



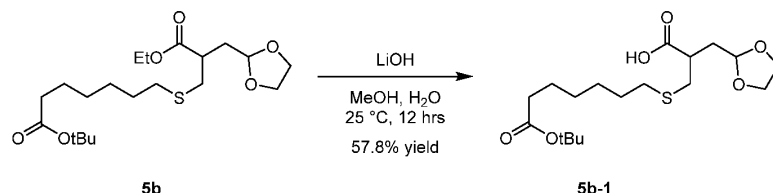
10

int. 1

5b

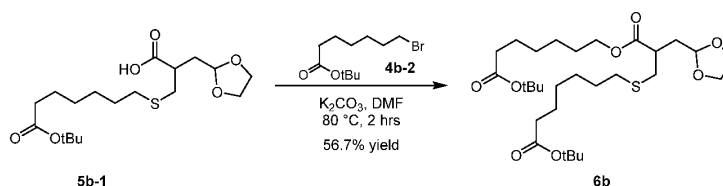
A mixture of compound **int.1** (12.0 g, 64.4 mmol, 1.00 eq) and compound **4b-3** (25.1 g, 77.3 mmol, 1.20 eq) in ethanol (120 mL) was added K₂CO₃ (17.8 g, 128 mmol, 2.00 eq) at 20 °C, the mixture was degassed and purged with N₂ for 3 times, and then the mixture was stirred at 20 °C for 48 h under N₂ atmosphere. TLC (petroleum ether/ethyl acetate = 10/1) indicated compound **int.1** (R_f = 0.60) was consumed completely and one main spot (R_f = 0.70) formed. The resulting product was dissolved in dichloromethane (100 mL) and filtered to removed the insoluble. The filter liquor was concentrated in vacuo. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 20/1 to 10/1, petroleum ether/ethyl acetate = 10/1, R_f = 0.10) to give compound **5b** (20.1 g, 46.7 mmol, 72.5% yield, 94.1% purity) as a yellow oil. LCMS: EW49391-7-P1A1, Rt = 0.635 min, m/z = 405.1, M+H⁺. ¹H NMR: EW49391-7-P1A2 (400 MHz, CDCl₃) δ 4.91 (t, J = 4.2 Hz, 1H), 4.14 (q, J = 14.0 Hz, 2H), 3.95 - 3.89 (m, 2H), 3.78 - 3.84 (m, 2H), 2.72 - 2.78 (m, 2H), 2.63 - 2.67 (m, 1H), 2.48 (t, J = 7.4 Hz, 2H), 2.17 (t, J = 7.4 Hz, 2H), 2.07 - 2.13 (m, 1H), 1.90 - 1.96 (m, 1H), 1.53 - 1.58 (m, 4H), 1.41 (s, 9H), 1.21 - 1.38 (m, 7H).

25

General procedure for preparation of compound 5b-1

To a solution of compound **5b** (13.0 g, 30.2 mmol, 1.00 eq) in methanol (97.5 mL) was added LiOH.H₂O (1.52 g, 36.2 mmol, 1.20 eq) in H₂O (32.5 mL). The mixture was stirred at 25 °C for 12 h. LCMS showed one peak with desired m/z was detected. The reaction mixture was concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 20/1 to 2/1, Petroleum ether/Ethyl acetate = 5/1, R_f = 0.20) to give Compound **5b-1** (7.24 g, 17.5 mmol, 57.8% yield, 91.0% purity) as a yellow oil. LCMS: EW49391-16-P1B1, Rt = 0.552 min, m/z = 399.1, M+Na⁺. LCMS: EW49391-16-P1C1, Rt = 0.558 min, m/z = 399.1, M+Na⁺. ¹H NMR: EW49391-16-P1C3 (400 MHz, CDCl₃) δ 7.85 - 9.48 (m, 1H), 4.98 (t, J = 4.2 Hz, 1H), 3.90 - 4.03 (m, 2H), 3.80 - 3.89 (m, 2H), 2.77 - 2.86 (m, 2H), 2.63 - 2.73 (m, 1H), 2.51 (t, J = 6.0 Hz, 2H), 2.10 - 2.21 (m, 3H), 1.96 - 2.03 (m, 1H), 1.52 - 1.58 (m, 4H), 1.43 (s, 9H), 1.27 - 1.40 (m, 4H).

15

General procedure for preparation of compound 6b

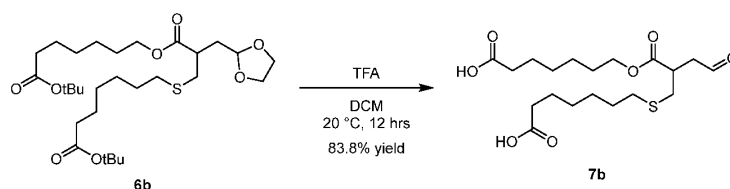
To a solution of compound **5b-1** (5.00 g, 13.2 mmol, 1.00 eq) in dimethylformamide (50.0 mL) was added compound **4b-2** (5.28 g, 19.9 mmol, 1.50 eq) and K₂CO₃ (3.67 g, 26.5 mmol, 2.00 eq). The mixture was stirred at 80 °C for 2 h. TLC (petroleum ether/ethyl acetate = 5/1) indicated one main spot (R_f = 0.50) formed. The reaction mixture was diluted with water (50.0 mL) and extracted with ethyl acetate (50.0 mL × 2). The

20

combined organic layers were washed with brine (50.0 mL $\times$ 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether/ethyl acetate = 20/1 to 5/1, Petroleum ether/Ethyl acetate = 5/1, R_f = 0.50) to give Compound **6b** (4.38 g, 7.54 mmol, 56.7% yield, 96.5% purity) as a yellow oil. LCMS: EW49391-18-P1A1, Rt = 0.712 min, m/z = 583.3, M+Na⁺. ¹H NMR: EW49391-18-P1A1 (400 MHz, CDCl₃) δ 4.93 (t, J = 4.4 Hz, 1H), 4.09 (t, J = 6.6 Hz, 2H), 3.90 - 4.00 (m, 2H), 3.80 - 3.88 (m, 2H), 2.73 - 2.81 (m, 2H), 2.63 - 2.72 (m, 1H), 2.50 (t, J = 7.2 Hz, 2H), 2.07 - 2.23 (m, 5H), 1.90 - 2.01 (m, 1H), 1.52 - 1.67 (m, 8H), 1.44 (s, 18H), 1.28 - 1.41 (m, 8H).

10

General procedure for preparation of compound 7b

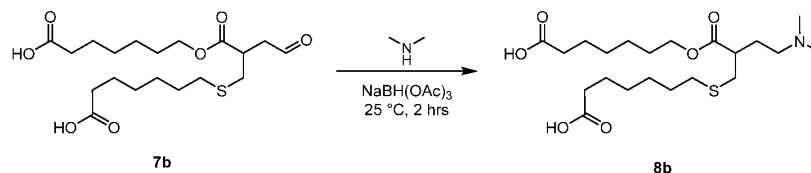


To a solution of compound **6b** (4.38 g, 7.81 mmol, 1.00 eq) in dichloromethane (43.8 mL) was added TFA (17.8 g, 156 mmol, 11.6 mL, 20.0 eq). The solution was stirred at 20 °C for 12 h. TLC (petroleum ether/ethyl acetate = 1/2) indicated compound **6b** (R_f = 0.80) was consumed completely and one new spot (R_f = 0.40) formed. The reaction mixture was concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether/ethyl acetate = 5/1 to 1/1, petroleum ether/ethyl acetate = 1/2, R_f = 0.40) to give Compound **7b** (2.65 g, 6.55 mmol, 83.8% yield, 100% purity) as a yellow oil. LCMS: EW49391-21-P1A2, Rt = 0.593 min, m/z = 405.1, M+H⁺. ¹H NMR: EW49391-21-P1A2 (400 MHz, CDCl₃) δ 9.78 (s, 1H), 8.62 - 9.05 (m, 2H), 3.09 - 3.13 (m, 1H), 2.73 - 3.00 (m, 3H), 2.61 - 2.66 (m, 1H), 2.49 (t, J = 7.24 Hz, 2H), 2.35 (td, J = 7.2, 2.4 Hz, 4H), 1.51 - 1.73 (m, 8H), 1.30 - 1.45 (m, 8H).

25

General procedure for preparation of compound 8b

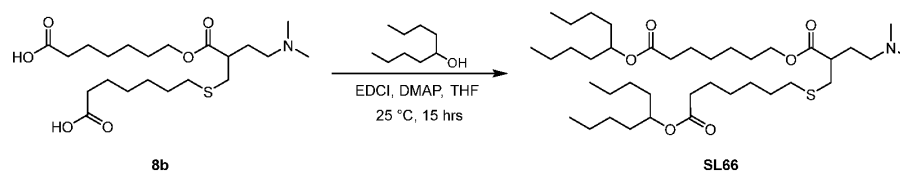
170



To a solution of compound **7b** (5.80 g, 14.3 mmol, 1.00 eq) in dichloromethane (100 mL) was added AcOH (1.03 g, 17.2 mmol, 984 μ L, 1.20 eq) and N-methylmethanamine (2 M, 10.7 mL, 1.50 eq). The solution was stirred at 25 $^{\circ}$ C for 1 h, then added NaBH(OAc)₃ (4.56 g, 21.5 mmol, 1.50 eq). The solution was stirred at 25 $^{\circ}$ C for 1 h. LCMS (EW49594-35-P1B) showed desired Ms was detected. The reaction was poured into aqueous solution of NaHCO₃ (aq. 120 mL). The resulting solution was extracted with tetrahydrofuran (100 mL x 2). Then acidified the aqueous phase with 2 M HCl to pH = 3. The aqueous phase was extracted with tetrahydrofuran (200 mL x 2). The combined organic layers were washed with brine (200 mL), dried over MgSO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g SepaFlash® Silica Flash Column, Eluent of 0~8% methanol/dichloromethane gradient @ 65 mL/min, dichloromethane : methanol = 10:1, P1:R_f = 0.37) to give Compound **8b** (1.50 g, 3.46 mmol, 24.1% yield) as a yellow gum. LCMS: EW49594-35-P1B, Rt = 0.402 min, m/z = 434.2, M+Na⁺. ¹H NMR: EW49594-35-P1B (400 MHz, CDCl₃) δ 4.87 (quin, *J* = 6.0 Hz, 2H), 4.09 (t, *J* = 6.8 Hz, 2H), 2.73 - 2.80 (m, 1H), 2.59 - 2.67 (m, 2H), 2.50 (t, *J* = 6.8 Hz, 2H), 2.23 - 2.35 (m, 6H), 2.20 (s, 6H), 1.74 - 1.83 (m, 2H), 1.49 - 1.70 (m, 16H), 1.18 - 1.44 (m, 24H), 0.89 (t, *J* = 6.8 Hz, 12H).

20

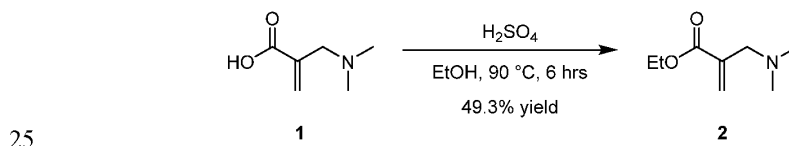
Preparation of compound SL66



To a solution of compound **8b** (1.40 g, 3.23 mmol, 1.00 eq) and nonan-5-ol (1.40 g, 9.69 mmol, 3.00 eq) in tetrahydrofuran (30.0 mL) was added EDCI (1.86 g, 9.69 mmol, 3.00

eq) and DMAP (788 mg, 6.46 mmol, 2.00 eq). The solution was stirred at 25 °C for 15 h. LCMS (EW49594-37-P1A) showed desired Ms was detected. The reaction mixture was added NaHCO₃ (aq. 50.0 mL) and extracted with ethyl acetate (50.0 mL x 2). The combined organic was dried over MgSO₄, filtered and concentrated under reduced pressure to give a crude product. The crude product was purified by flash silica gel chromatography (ISCO®; 40 g SepaFlash® Silica Flash Column, Eluent of 0~15% tetrahydrofuran/petroleum ether gradient @ 35mL/min, petroleum ether : tetrahydrofuran = 1:1, R_f = 0.17) to obtain 1.5 g of the residue. 1.5 g of the residue was purified by prep-HPLC (column: Phenomenex Luna C18 150*25mm*10um; mobile phase: [water (TFA) - methanol]; gradient:80%-98% B over 15 min) to obtain 750 mg of compound **SL66**. 750 mg of the product was added ethyl acetate (20.0 mL) and washed NaHCO₃ (aq. 10.0 mL x 5), dried over MgSO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 40 g SepaFlash® Silica Flash Column, Eluent of 0~1% methanol/dichloromethane gradient @ 35 mL/min, petroleum ether : tetrahydrofuran=1:1, R_f = 0.17) to give compound **SL66** (650 mg, 946 μmol, 29.3% yield, 99.9% purity) as a yellow oil. LCMS: EW49594-37-P1A, R_t = 0.708 min, m/z = 686.6, M+H⁺. Special LCMS: EW49594-37-P1C, R_t = 1.814 min, m/z = 686.6, M+H⁺. ¹H NMR: EW49594-37-P1A (400 MHz, CDCl₃) δ 4.87 (q, *J* = 6.0 Hz, 2H), 4.09 (t, *J* = 6.8 Hz, 2H), 2.73 - 2.80 (m, 1H), 2.59 - 2.67 (m, 2H), 2.50 (t, *J* = 6.8 Hz, 2H), 2.23 - 2.35 (m, 6H), 2.20 (s, 6H), 1.74 - 1.83 (m, 2H), 1.49 - 1.70 (m, 16H), 1.18 - 1.44 (m, 24H), 0.89 (t, *J* = 6.8 Hz, 12H).

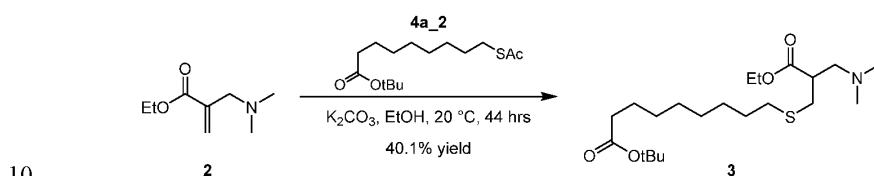
General procedure for preparation of compound 2



To a solution of compound **1** (200 g, 1.55 mol, 1.00 eq) in ethanol (2000 mL) and H₂SO₄ (400 mL) was stirred at 90 °C for 6 h. LCMS (EW49393-11-P1A) showed the desired mass (R_t = 0.332 min, m/z = 158.1, M+H⁺) was detected. The reaction mixture was

cooled to 25 °C and concentrated under reduced pressure to remove ethanol, and then adjust pH = 9 with aq. Na₂CO₃, extracted with dichloromethane 3000 mL (1000 mL x 3). The combined organic layers were washed with brine 1000 mL, dried with anhydrous Na₂SO₄, filtered and concentrated in vacuum to give compound **2** (120 g, 763 mmol, 49.3% yield) as a yellow oil. LCMS: EW49393-11-P1A, Rt = 0.332 min, m/z = 158.1 (M+H⁺). ¹H NMR: EW49393-11-P1B (400 MHz, CDCl₃) δ 6.26 (s, 1H), 5.72 (s, 1H), 4.21 (q, J = 7.2 Hz, 2H), 2.13 (s, 2H), 2.24 (s, 6H), 1.29 (t, J = 7.2 Hz, 3H).

General procedure for preparation of compound 3



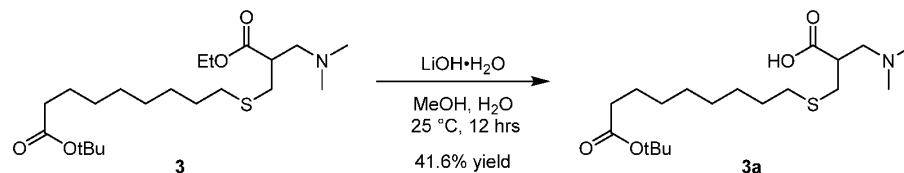
To a solution of compound **4a_2** (68.6 g, 238 mmol, 1.10 eq) in ethanol (340 mL) was added K₂CO₃ (59.8 g, 432 mmol, 2.00 eq) and compound **2** (34.0 g, 216 mmol, 1.00 eq). The mixture was stirred at 20 °C for 44 h. LCMS (EW49393-30-P1A) showed compound **2** was consumed completely. The reaction mixture was filtered and concentrated in vacuum to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether/ethyl acetate = 40/1 to 5/1, petroleum ether/ethyl acetate = 0:1, R_f = 0.14) to give compound **3** (35.0 g, 86.7 mmol, 40.1% yield) as a yellow oil. LCMS: EW49393-30-P1A, Rt = 0.507 min, m/z = 404.4 (M+H⁺). ¹H NMR: EW49393-30-P1A (400 MHz, CDCl₃) δ 4.15 - 4.21 (m, 2H), 2.75 - 2.78 (m, 1H), 2.70 - 2.72 (m, 2H), 2.57 - 2.62 (m, 1H), 2.51 (t, J = 7.2 Hz, 2H), 2.38 - 2.43 (m, 1H), 2.18 - 2.22 (m, 7H), 1.52 - 1.60 (m, 4H), 1.44 (s, 9H), 1.24 - 1.29 (m, 12H).

15

20

General procedure for preparation of compound 3a

173



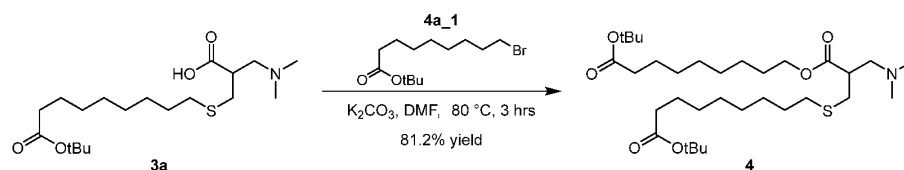
To a solution of compound **3** (35.0 g, 86.7 mmol, 1.00 eq) in methanol (245 mL) was added LiOH·H₂O (4.37 g, 104 mmol, 1.20 eq) in H₂O (100 mL). The mixture was stirred at 25 °C for 12 h. TLC (dichloromethane : methanol = 10 : 1) indicated compound **3** was

5 remained (R_f = 0.24) and the desired product was formed (R_f = 0.02). The reaction mixture was concentrated under reduced pressure to remove methanol. The residue was diluted with ethyl acetate 300 mL and extracted with H₂O 300 mL. The water layer was adjusted pH = 5 with aq. citric acid and extracted with ethyl acetate 300 mL and extracted with dichloromethane 300 mL, washed with brine 300 mL, dried over Na₂SO₄, filtered

10 and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 8/1, dichloromethane : methanol = 10:1, R_f = 0.02) to give compound **3a** (13.9 g, 36.0 mmol, 41.6% yield, 97.4% purity) as a yellow oil. LCMS: EW49393-33-P1C, R_t = 0.479 min, m/z = 376.1 ($M+H^+$). ¹H NMR: EW49393-33-P1A (400 MHz, CDCl₃) δ 10.39 (s, 1H), 3.41 - 3.49

15 (m, 1H), 3.06 - 3.20 (m, 2H), 2.95 - 2.98 (m, 1H), 2.89 (s, 6H), 2.53 - 2.59 (m, 3H), 2.20 (t, J = 7.2 Hz, 2H), 1.54 - 1.59 (m, 4H), 1.44 (s, 9H), 1.34 - 1.37 (m, 2H), 1.29 (s, 6H).

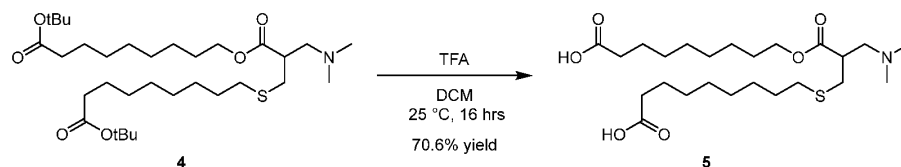
General procedure for preparation of compound 4



20 To a solution of compound **3a** (11.8 g, 31.4 mmol, 1.00 eq) in dimethylformamide (120 mL) was added K₂CO₃ (8.68 g, 62.84 mmol, 2.00 eq) and compound **4a_1** (13.8 g, 47.1 mmol, 1.50 eq). The mixture was stirred at 80 °C for 3 h. LCMS (EW49393-35-P1A) showed compound **3a** was consumed completely. The reaction mixture was diluted with ethyl acetate 200 mL and extracted with H₂O 200 mL, washed with brine 200 mL, dried

over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether/Ethyl acetate = 40/1 to 5/1, Petroleum ether/Ethyl acetate = 1:1, R_f = 0.14) to give compound **4** (15.0 g, 25.5 mmol, 81.2% yield) as a yellow oil. LCMS: EW49393-35-P1C, Rt = 0.588 min, m/z = 588.4 (M+H⁺). ¹H NMR: EW49393-35-P1A (400 MHz, CDCl₃) δ 4.12 (t, *J* = 7.2 Hz, 2H), 2.70 - 2.79 (m, 3H), 2.57 - 2.62 (m, 1H), 2.51 (t, *J* = 7.2 Hz, 2H), 2.38 - 2.43 (m, 1H), 2.22 (s, 6H), 2.19 - 2.20 (m, 3H), 1.53 - 1.66 (m, 9H), 1.45 (s, 18H), 1.30 - 1.38 (m, 17H).

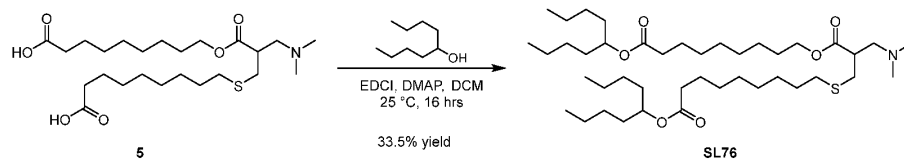
10 General procedure for preparation of compound **5**



To a solution of compound **4** (7.00 g, 11.9 mmol, 1.00 eq) in dichloromethane (70.0 mL) was added TFA (35.0 mL). The mixture was stirred at 25 °C for 16 h. LCMS (EW49393-37-P1A) showed compound **4** was consumed completely. The reaction mixture was concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 20/1, dichloromethane : methanol = 10:1, R_f = 0.02) to give compound **5** (4.00 g, 8.41 mmol, 70.6% yield) as a yellow oil. LCMS: EW49393-37-P1A, Rt = 0.442 min, m/z = 476.2 (M+H⁺). ¹H NMR: EW49393-37-P1A (400 MHz, CDCl₃) δ 4.15 - 4.20 (m, 2H), 3.50 - 3.55 (m, 2H), 3.24 - 3.26 (m, 1H), 2.72 - 2.92 (m, 8H), 2.53 (t, *J* = 7.2 Hz, 2H), 2.35 (t, *J* = 7.2 Hz, 4H), 1.53 - 1.69 (m, 9H), 1.45 - 1.26 (m, 17H).

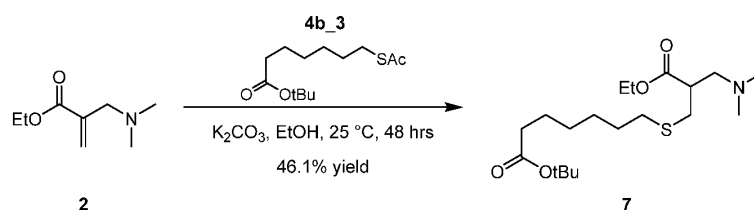
Preparation of compound **SL76**

175



- To a solution of compound **5** (3.00 g, 6.31 mmol, 1.00 eq), nonan-5-ol (2.73 g, 18.9 mmol, 3.00 eq) in tetrahydrofuran (30.0 mL) was added EDCI (3.63 g, 18.9 mmol, 3.00 eq) and DMAP (1.54 g, 12.6 mmol, 2.00 eq). The mixture was stirred at 25 °C for 16 h.
- 5 LC-MS (EW49393-38-P1A) showed compound **5** was consumed completely. The reaction mixture was partitioned between NaHCO₃ (25.0 mL) and EtOAc 25.0 mL × 2. The organic phase was separated, washed with brine 25.0 mL, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 100/1,
- 10 dichloromethane : methanol= 10:1, R_f = 0.02) to give **SL76** (1.54 g, 2.11 mmol, 33.5% yield, 99.9% purity) as a yellow oil. LCMS: EW49393-38-P1A, R_t = 0.761 min, m/z = 728.5 (M+H⁺). Special LCMS: EW49393-38-P1A, R_t = 2.073 min, m/z = 728.8 (M+H⁺). ¹H NMR: EW49393-38-P1B (400 MHz, CDCl₃) δ 4.86 - 4.91 (m, 2H), 4.11 (t, J = 6.8 Hz, 2H), 2.70 - 2.79 (m, 3H), 2.57 - 2.62 (m, 1H), 2.51 (t, J = 7.2 Hz, 2H), 2.38 -
- 15 2.43 (m, 1H), 2.28 (t, J = 7.2 Hz, 4H), 2.22 (s, 6H), 1.48 - 1.62 (m, 15H), 1.21 - 1.38 (m, 33H), 0.89 (t, J = 7.2 Hz, 12H).

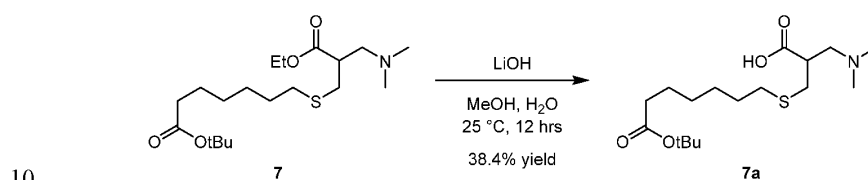
General procedure for preparation of compound 7



- 20 To a solution of compound **4b_3** (107 g, 350 mmol, 1.10 eq) in ethanol (500 mL) was added K₂CO₃ (87.9 g, 636 mmol, 2.00 eq) and compound **2** (50.0 g, 318 mmol, 1.00 eq). The mixture was stirred at 20 °C for 48 h. TLC (petroleum ether/ethyl acetate = 0:1) showed compound **2** was consumed completely. The reaction mixture was filtered and

concentrated in vacuum to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether/ethyl acetate = 40/1 to 5/1, petroleum ether / ethyl acetate = 0:1, R_f = 0.14) to give compound **7** (55.0 g, 146 mmol, 46.1% yield) as a yellow oil. ¹H NMR: EW49393-15-P1A (400 MHz, CDCl₃) δ 4.11 - 4.21 (m, 2H), 2.75 - 2.80 (m, 1H), 2.70 - 2.72 (m, 2H), 2.57 - 2.62 (m, 1H), 2.51 (t, *J* = 7.2 Hz, 2H), 2.38 - 2.43 (m, 1H), 2.18 - 2.22 (m, 7H), 1.54 - 1.61 (m, 4H), 1.44 (s, 9H), 1.31 - 1.42 (m, 4H), 1.24 (t, *J* = 6.0 Hz, 3H).

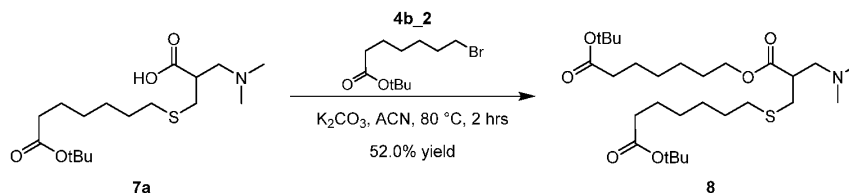
General procedure for preparation of compound 7a



To a solution of compound **7** (20.0 g, 53.2 mmol, 1.00 eq) in methanol (140 mL) was added LiOH·H₂O (2.68 g, 63.9 mmol, 1.20 eq) in H₂O (80.0 mL). The mixture was stirred at 25 °C for 12 h. TLC (dichloromethane : methanol = 10 : 1) indicated compound **7** was remained (R_f = 0.24) and the desired product was formed (R_f = 0.02). The reaction mixture was concentrated under reduced pressure to remove methanol. The residue was diluted with ethyl acetate 100 mL and extracted with H₂O 100 mL. The water layer was adjusted pH= 5 with aq. citric acid and extracted with ethyl acetate 100 mL and extracted with dichloromethane 300 mL, washed with brine 100 mL, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give compound **7a** (7.11 g, 20.5 mmol, 38.4% yield) as a yellow oil. ¹H NMR: EW49393-17-P1A (400 MHz, CDCl₃) δ 6.05 (s, 1H), 3.08 - 3.14 (m, 2H), 2.93 - 2.98 (m, 1H), 2.69 (s, 6H), 2.51 - 2.59 (m, 3H), 2.20 (t, *J* = 7.2 Hz, 2H), 1.54 - 1.63 (m, 4H), 1.44 (s, 9H), 1.24 - 1.40 (m, 2H).

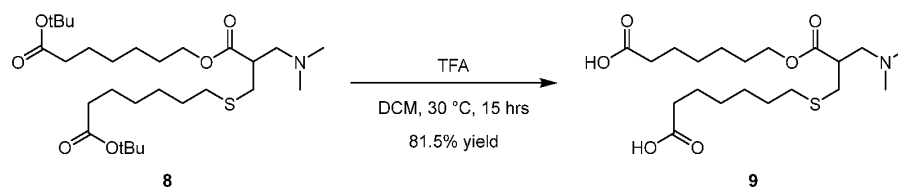
General procedure for preparation of compound 8

177



To a solution of compound **7a** (14.2 g, 40.9 mmol, 1.00 eq) in dimethylformamide (150 mL) was added K_2CO_3 (11.3 g, 81.7 mmol, 2.00 eq) and compound **4b_2** (16.3 g, 61.3 mmol, 1.50 eq). The mixture was stirred at 80 °C for 3 h. LCMS (EW49393-24-P1A) showed compound **7a** was consumed completely. The reaction mixture was diluted with ethyl acetate 200 mL and extracted with H_2O 200 mL, washed with brine 200 mL, dried over Na_2SO_4 , filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO_2 , Petroleum ether/Ethyl acetate = 40/1 to 5/1, Petroleum ether /Ethyl acetate = 1:1, R_f = 0.14) to give compound **8** (11.3 g, 21.3 mmol, 52.0% yield) as a yellow oil. LCMS: EW49393-24-P1A, R_t = 0.563 min, m/z = 532.5 ($\text{M}+\text{H}^+$). ^1H NMR: EW49393-24-P1A (400 MHz, CDCl_3) δ 4.12 (t, J = 6.8 Hz, 2H), 2.70 - 2.81 (m, 3H), 2.57 - 2.62 (m, 1H), 2.51 (t, J = 7.2 Hz, 2H), 2.38 - 2.42 (m, 1H), 2.19 - 2.23 (m, 9H), 1.54 - 1.69 (m, 9H), 1.45 (s, 18H), 1.28 - 1.41 (m, 9H).

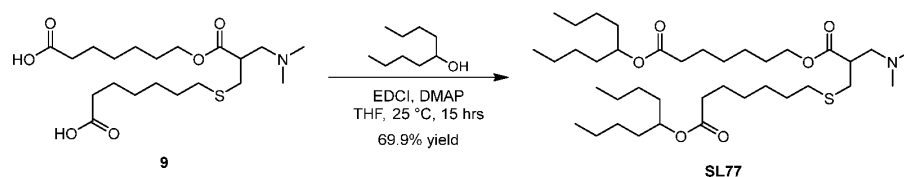
15 General procedure for preparation of compound **9**



To a solution of compound **8** (5.60 g, 10.5 mmol, 1.00 eq) in dichloromethane (45.0 mL) was added TFA (24.0 g, 210 mmol, 15.6 mL, 20.0 eq). The reaction was stirred at 30 °C for 15 h. The reaction was concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g SepaFlash® Silica Flash Column, Eluent of 0~8% ethyl acetate/ petroleum ether @ 65 mL/min, dichloromethane: methanol = 10: 1, P1: R_f = 0.39) to give compound **9** (3.60 g, crude) as a yellow gum. ^1H NMR: EW49594-31-P1A (400 MHz, CDCl_3) δ 4.03 - 4.08 (m, 2H),

3.36 - 3.37 (m, 1H), 3.34 - 3.36 (m, 1H), 3.26 - 3.27 (m, 1H), 2.74 - 2.79 (m, 1H), 2.48 (s, 6H), 2.45 - 2.47 (m, 1H), 2.28 - 2.29 (m, 2H), 2.25 - 2.26 (m, 4H), 1.58 - 1.60 (m, 8H), 1.31 - 1.36 (m, 8H).

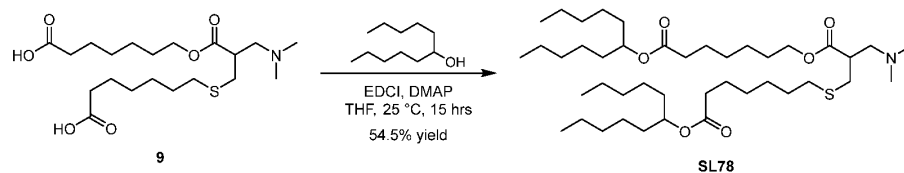
5 Preparation of compound **SL77**



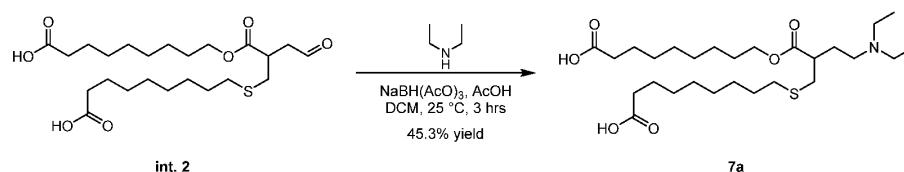
To a solution of compound **9** (2.50 g, 5.96 mmol, 1.00 eq) and nonan-5-ol (2.58 g, 17.9 mmol, 3.00 eq) in tetrahydrofuran (60.0 mL) was added EDCI (3.43 g, 17.9 mmol, 3.00 eq) and DMAP (1.46 g, 11.9 mmol, 2.00 eq). The solution was stirred at 25 °C for 15 h.

10 The reaction mixture was partitioned between NaHCO₃ 25.0 mL and EtOAc 25.0 mL ×
2. The organic phase was separated, washed with brine 25.0 mL, dried over Na₂SO₄,
filtered and concentrated under reduced pressure to give a residue. The residue was
purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash ® Silica Flash
Column, Eluent of 0~1% methanol / dichloromethane gradient @ 65 mL/min, R_f = 0.47).
15 **SL77** (2.00 g, 2.98 mmol, 49.9% yield) was obtained as yellow gum and further purified
by flash silica gel chromatography (ISCO®; 80 g Sepa Flash ® Silica Flash Column,
Eluent of 0~15% tetrahydrofuran / petroleum ether gradient @ 65 mL/min, petroleum
ether: tetrahydrofuran = 1: 1, R_f = 0.47) to give **SL77** (1.40 g, 2.08 mmol, 69.9% yield,
99.8% purity) as a yellow oil. Specical LCMS: EW49594-45-P1A1: Rt = 1.784 min, m/z
20 = 672.1 (M+H⁺). ¹H NMR: EW49594-45-P1C (400 MHz, CDCl₃) δ 4.84 - 4.91 (m, 2H),
4.12 (t, *J* = 6.8 Hz, 2H), 2.72 - 2.81 (m, 1H), 2.51 - 2.71 (m, 2H), 2.50 - 2.51 (m, 1H),
2.30 - 2.50 (m, 2H), 2.29 - 2.30 (m, 1H), 2.27 - 2.29 (m, 4H), 2.22 (s, 6H), 1.51 - 1.64
(m, 16H), 1.27 - 1.33 (m, 24H), 0.89 (t, *J* = 7.2 Hz, 12H).

25 Preparation of compound **SL78**



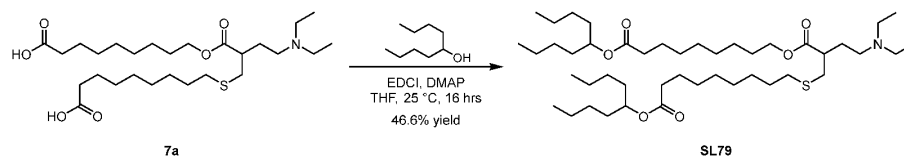
To a solution of compound **9** (2.50 g, 5.96 mmol, 1.00 eq) and undecan-6-ol (3.08 g, 17.88 mmol, 3 eq) in tetrahydrofuran (60.0 mL) was added EDCI (3.43 g, 17.9 mmol, 3.00 eq) and DMAP (1.46 g, 11.9 mmol, 2.00 eq). The solution was stirred at 25 °C for 15 h. The reaction mixture was partitioned between NaHCO₃ 25.0 mL and EtOAc 25.0 mL × 2. The organic phase was separated, washed with brine 25.0 mL, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash ® Silica Flash Column, Eluent of 0~1% methanol / dichloromethane gradient @ 65 mL/min, R_f = 0.66). **SL78** (2.20 g, 3.02 mmol, 50.7% yield) was obtained as yellow gum and further purified by flash silica gel chromatography (ISCO®; 80 g Sepa Flash ® Silica Flash Column, Eluent of 0~15% tetrahydrofuran / Petroleum ether gradient @ 65 mL/min, Petroleum ether: tetrahydrofuran = 1: 1, R_f = 0.66) to give **SL78** (1.20 g, 1.65 mmol, 54.5% yield, 99.9% purity) as a yellow oil. Specical LCMS: EW49594-46-P1A2: Rt = 2.202 min, m/z = 728.8 (M+H⁺). ¹H NMR: EW49594-46-P1C (400 MHz, CDCl₃) δ 4.85 - 4.91 (m, 2H), 4.12 (t, J = 6.8 Hz, 2H), 2.72 - 2.81 (m, 1H), 2.51 - 2.71 (m, 2H), 2.50 - 2.51 (m, 1H), 2.30 - 2.50 (m, 2H), 2.29 - 2.30 (m, 1H), 2.27 - 2.29 (m, 4H), 2.22 (s, 6H), 1.51 - 1.64 (m, 16H), 1.28 - 1.37 (m, 32H), 0.89 (t, J = 7.2 Hz, 12H).

20 *General procedure for preparation of compound 7a*

To a solution of compound **int.2** (5.00 g, 10.8 mmol, 1.00 eq), piperidine (793 mg, 70.8 mmol, 1.12 mL, 1.00 eq) in dichloromethane (50.0 mL) was added AcOH (782 mg, 13.0 mmol, 745 μ L, 1.20 eq), the mixture was stirred at 25 °C for 1.5 h, then NaBH(OAc)₃

(2.76 g, 13.0 mmol, 1.20 eq) was added to the mixture. The mixture was stirred at 25 °C for 1.5 h. TLC (dichloromethane/methanol = 10/1, R_f = 0.3) indicated compound **int.2** was consumed completely and two new spots formed. The reaction mixture was diluted with H₂O 100 mL and extracted with dichloromethane 300 mL (100 mL × 3). The combined organic layers were washed with brine 100 mL, dried with anhydrous Na₂SO₄, filtered and concentrated in vacuum to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 80/1, dichloromethane: Methanol= 10:1, R_f = 0.30) to give compound **7a** (4.05 g, 4.90 mmol, 45.3% yield, 66.8% purity) as a colorless oil. LCMS: EW49391-28-P1C1, R_t = 0.443 min, m/z = 518.3 (M+H⁺). ¹H NMR: EW49391-28-P1C2 (400 MHz, CDCl₃) δ 9.5 - 10.7 (s, 2H), 4.09 - 4.14 (m, 2H), 3.05 - 3.20 (m, 5H), 2.62 - 2.72 (m, 2H), 2.48 - 2.51 (m, 2H), 2.30 - 2.67 (m, 4H), 1.96 - 2.06 (m, 2H), 1.50 - 1.67 (m, 9H), 1.33 - 1.43 (m, 17H).

Preparation of compound **SL79**

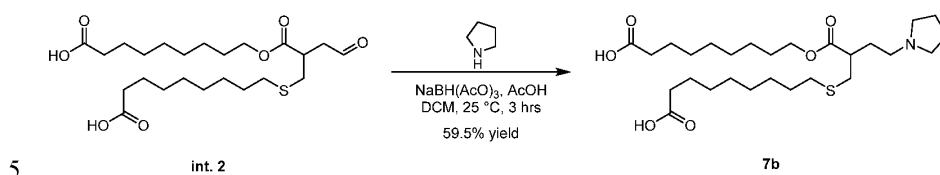


To a solution of compound **7a** (2.40 g, 4.64 mmol, 1.00 eq) nonan-5-ol (2.01 g, 13.9 mmol, 3.00 eq) in tetrahydrofuran (48.0 mL) was added EDCI (2.67 g, 13.9 mmol, 3.00 eq) and DMAP (1.13 g, 9.27 mmol, 2.00 eq). The mixture was stirred at 25 °C for 16 h. TLC (Dichloromethane/Methanol = 10/1) indicated compound **7a** (R_f = 0.30) was consumed completely and one main spot (R_f = 0.50) formed. The reaction mixture was partitioned between NaHCO₃ 25.0 mL and EtOAc 25.0 mL × 2. The organic phase was separated, washed with brine 25.0 mL, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 100/1, Dichloromethane/Methanol = 10/1, R_f = 0.50) to give compound **SL79** (1.67 g, 2.16 mmol, 46.6% yield, 99.7% purity) as a yellow oil. ELSD: EW49391-30-P1A2, R_t = 2.037 min, m/z = 771.9 (M+H⁺). ¹H NMR: EW49391-30-P1A4 (400 MHz, CDCl₃) δ 4.84 - 4.89 (m, 2H), 4.08 (t, J = 6.8 Hz, 2H), 2.72 - 2.83 (m, 1H), 2.45 - 2.67 (m, 10H),

181

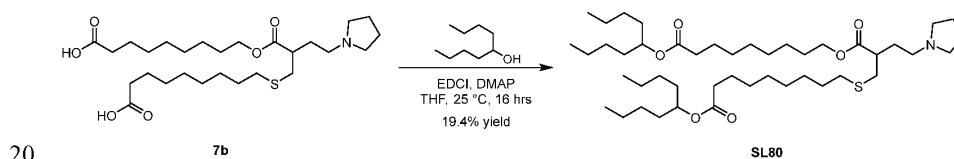
2.27 (t, J = 7.2 Hz, 4 H), 1.72 - 1.91 (m, 2 H), 1.44 - 1.69 (m, 15 H), 1.19 - 1.41 (m, 33 H), 1.03 (t, J = 7.2 Hz, 6 H), 0.88 (t, J = 6.8 Hz, 12H).

General procedure for preparation of compound 7b



To a solution of compound **int.2** (15.0 g, 32.5 mmol, 1.00 eq), piperidine (2.32 g, 32.5 mmol, 2.72 mL, 1.00 eq) in dichloromethane (150 mL) was added AcOH (2.35 g, 39.0 mmol, 2.24 mL, 1.20 eq), the mixture was stirred at 25 °C for 1.5 h, then NaBH(OAc)₃ (8.28 g, 39.0 mmol, 1.20 eq) was added to the mixture. The mixture was stirred at 25 °C for 1.5 h. TLC (dichloromethane/methanol = 10/1, R_f = 0.3) indicated compound **int.2** was consumed completely and three new spots formed. The reaction mixture was diluted with H₂O 100 mL and extracted with dichloromethane 300 mL (100 mL × 3). The combined organic layers were washed with brine 100 mL, dried with anhydrous Na₂SO₄, filtered and concentrated in vacuum to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 80/1, Dichloromethane/Methanol = 10/1, R_f = 0.30) to give compound **7b** (10.0 g, 19.4 mmol, 59.5% yield) as a yellow oil.

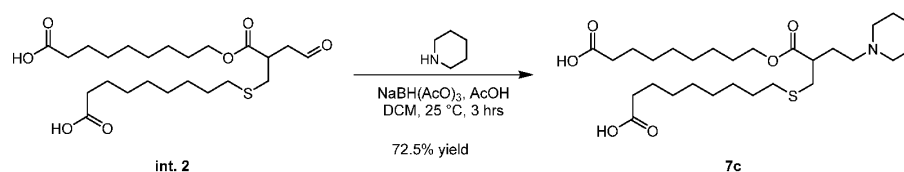
Preparation of compound SL80



To a solution of compound **7b** (5.00 g, 9.69 mmol, 1.00 eq) nonan-5-ol (4.20 g, 29.0 mmol, 3.00 eq) in tetrahydrofuran (100 mL) was added EDCI (5.58 g, 29.0 mmol, 3.00 eq) and DMAP (2.37 g, 19.3 mmol, 2.00 eq). The mixture was stirred at 25 °C for 16 h.

TLC (dichloromethane/methanol = 10/1 indicated compound **7b** (R_f = 0.30) was consumed completely and one main spot (R_f = 0.50) formed. The reaction mixture was partitioned between NaHCO_3 100 mL and EtOAc 50.0 mL $\times$ 2. The organic phase was separated, washed with brine 100 mL, dried over Na_2SO_4 , filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 100/1, dichloromethane/methanol = 10/1, R_f = 0.50) to give compound **SL80** (1.45 g, 1.88 mmol, 19.4% yield, 99.8% purity) as a yellow oil. ELSD: EW49391-32-P1A1, Rt = 2.231 min, m/z = 768.9 ($\text{M}+\text{H}^+$). ¹H NMR: EW49391-32-P1A2 (400 MHz, CDCl_3) δ 4.84 - 4.91 (m, 2H), 4.08 (t, J = 6.8 Hz, 2H), 2.72 - 2.84 (m, 1H), 2.46 - 2.69 (m, 10H), 2.27 (t, J = 7.6 Hz, 4H), 1.77 - 1.98 (m, 6H), 1.48 - 1.63 (m, 1H), 1.43 - 1.69 (m, 17H), 1.16 - 1.42 (m, 34H), 0.88 (t, J = 6.8 Hz, 12H).

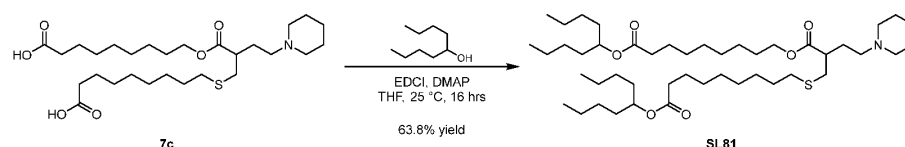
General procedure for preparation of compound 7c



To a solution of compound **int.2** (3.00 g, 6.51 mmol, 1.00 eq), piperidine (554 mg, 6.51 mmol, 643 μL , 1.00 eq) in dichloromethane (30.0 mL) was added AcOH (469 mg, 7.82 mmol, 447 μL , 1.20 eq), the mixture was stirred at 20 $^\circ\text{C}$ for 1.5 h, then NaBH(OAc)_3 (1.66 g, 7.82 mmol, 1.20 eq) was added to the mixture. The mixture was stirred at 20 $^\circ\text{C}$ for 1.5 h. LC-MS (EW49393-31-P1A) showed compound **int.2** was consumed completely. The reaction mixture was diluted with H_2O 100 mL and extracted with dichloromethane 300 mL (100 mL $\times$ 3). The combined organic layers were washed with brine 100 mL, dried with anhydrous Na_2SO_4 , filtered and concentrated in vacuum to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 8/1, dichloromethane: Methanol = 10:1, R_f = 0.16) to give compound **7c** (2.50 g, 4.72 mmol, 72.5% yield, 100% purity) as a colorless oil. LCMS: EW49393-31-P1A, Rt = 0.467 min, m/z = 530.4 ($\text{M}+\text{H}^+$). LCMS: EW49393-31-P1B, Rt = 0.484 min, m/z = 530.3 ($\text{M}+\text{H}^+$). ¹H NMR: EW49393-31-P1B (400 MHz,

CDCl₃) δ 11.27 (s, 2H), 4.06 - 4.16 (m, 2H), 2.97 - 3.06 (m, 4H), 2.74 - 2.81 (m, 2H), 2.63 - 2.68 (m, 2H), 2.47 - 2.54 (m, 2H), 2.28 (t, $J = 7.2$ Hz, 4H), 2.01 - 2.06 (m, 5H), 1.52 - 1.62 (m, 10H), 1.33 - 1.43 (m, 16H).

5 Preparation of compound **SL81**



To a solution of compound **7c** (2.30 g, 4.34 mmol, 1.00 eq) nonan-5-ol (1.88 g, 13.0 mmol, 3.00 eq) in tetrahydrofuran (25.0 mL) was added EDCI (2.50 g, 13.0 mmol, 3.00 eq) and DMAP (1.06 g, 8.68 mmol, 2.00 eq). The mixture was stirred at 25 °C for 16 h.

LCMS (EW49393-36-P1A) showed compound **7c** was consumed completely. The reaction mixture was partitioned between NaHCO₃ 25.0 mL and EtOAc 25.0 mL × 2. The organic phase was separated, washed with brine 25.0 mL, dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, dichloromethane/methanol = 1/0 to 100/1, dichloromethane : Methanol = 10:1, R_f = 0.43). to give compound **SL81** (2.17 g, 2.77 mmol, 63.8% yield, 99.8% purity) as a yellow oil. LCMS: EW49393-36-P1A, R_t = 0.748 min, m/z = 782.7 (M+H⁺). ELSD: EW49393-36-P1A, R_t = 2.188 min, m/z = 782.9 (M+H⁺). ¹H NMR: EW49393-36-P1A (400 MHz, CDCl₃) δ 4.84 - 4.91 (m, 2H), 4.09 (t, *J* = 4.8 Hz, 2H), 2.74 - 2.78 (m, 1H), 2.62 - 2.66 (m, 2H), 2.30 (t, *J* = 7.2 Hz, 2H), 2.33 - 2.39 (m, 5H), 2.28 (t, *J* = 7.2 Hz, 5H), 1.80 - 1.90 (m, 3H), 1.49 - 1.63 (m, 17H), 1.43 - 1.45 (m, 3H), 1.21 - 1.34 (m, 33H), 0.89 (t, *J* = 6.8 Hz, 12H).

Example 2: Formulation of ionizable cationic lipids into LNPs with a mRNA

[0195] This is a general description of how each ionizable cationic lipid will be used to formulate an samRNA LNP. The ionizable cationic lipid, 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), cholesterol, and 1,2-dimyristoyl-rac-glycero-3-methylpolyoxyethylene (DMG-PEG) will be combined in a 40:10:48:2 molar ratio in

ethanol at a concentration of 3.2 mM (for 0.5 mg samRNA scale) or 6.4 mM (for 1.5 mg samRNA scale). A solution of samRNA expressing the antigen of interest in 50 mM citrate buffer at pH 6 with 10 mM Tris (2-carboxyethyl) phosphine (TCEP) will be prepared. When mixing at an aqueous to ethanol flow rate ratio (FRR) of 2:1, the RNA concentration will be 0.025 mg/mL (for 0.5 mg samRNA scale) or 0.050 mg/mL (for 1.5 mg samRNA scale). When mixing at an aqueous to ethanol flow rate ratio (FRR) of 3:1, the RNA concentration will be 0.017 mg/mL (for 0.5 mg samRNA scale). The lipid solution in ethanol will then be rapidly mixed with the samRNA solution using a Knauer's benchtop IJM NanoScaler system at a flow rate ratio (FRR) of either 2:1 or 3:1 aqueous buffer to ethanol. This mixing ratio will result in a 8:1 molar ratio of ionizable cationic lipid (see Table 1) to samRNA phosphate groups and a total lipid to samRNA mass ratio of 37:1. The resulting mixed solution will then be diluted 10-fold into 50 mM citrate buffer at pH 6 with 10 mM TCEP and subjected to tangential flow filtration (TFF) using a 300k molecular weight cut-off membrane (mPES) until concentrated to the original volume. Subsequently, the citrate buffer will be replaced with a buffer containing 20 mM Tris buffer at pH 7.5, 80 mM sodium chloride, and 3% sucrose using diafiltration with 10 diavolumes. The LNP solution will be concentrated to a volume between 4-10 mL, filtered using a 0.2 micron PES syringe filter, aliquoted into vials, and frozen at 1°C/min using a Corning® CoolCell® LX Cell Freezing Container until the samples reached -80 °C. Samples will be stored at -80 °C and thawed on wet ice before analysis or use. The total RNA concentration, the percentage of input RNA recovered (% recovery), and the encapsulation efficiency (%EE) will be determined using a Ribogreen assay, which is described elsewhere. The Z-avg diameter (nm) and polydispersity index (PDI) will be measured using dynamic light scattering (Malvern Zetasizer) following a 1:100 dilution in phosphate buffered saline (PBS).

[0196] Results of the formulation studies are depicted in the table below:

LNP #	Ioniz-able Cationic Lipid	Flow-rate ratio (FRR, Aq: EtOH)	Citrate buffer pH	Encap-sulation efficie-ncy (EE, %)	Total RNA conc-entration (µg/mL)	RNA reco-very (%)	Z-average diameter by DLS (nm)	PDI by DLS	pKa
-------	---------------------------	---------------------------------	-------------------	------------------------------------	----------------------------------	-------------------	--------------------------------	------------	-----

1	LKY750	2:1	6	96.9	40.5	68.8	80.18	0.138	6.45
2	SL56	3:1	6	97.6	44.7	75.1	138.9	0.211	7.1
3	SL57	3:1	6	96.1	52.7	92.4	91.37	0.082	6.83
4	SL58	3:1	4	95.9	43.7	72.2	69.61	0.215	5.67
5	SL59	2:1	4	95.2	41.2	66.7	93.5	0.217	6.24
6	SL60	2:1	4	93.8	41.7	73	97.99	0.078	6.17
7	SL61	2:1	4	92.4	45.5	81.8	73.85	0.127	4.88
8	SL62	3:1	6	94.3	30.2	56.2	79.34	0.141	7.04
9	SL63	3:1	4->6	88.7	26.9	54.5	79.50	0.104	7.03
10	SL64	3:1	4->6	79.7	57.9	41.7	89.86	0.207	6.61
11	SL65	3:1	4->6	90.1	37.4	68	71.87	0.122	6.70
12	SL66	3:1	6	94.3	34.4	62.7	136.9	0.075	6.83
13	SL67	3:1	6	95.3	42.7	74.4	88.25	0.099	7.05
14	SL68	3:1	6	97.5	35.4	61.9	144.9	0.195	7.51
15	SL69	3:1	6	81.7	43.1	69.4	104.7	0.090	6.77
16	SL70	3:1	6	86.85	59.6	35	91.2	0.106	6.90
17	SL71	3:1	6	76.49	57.7	43.7	90.1	0.128	6.85
18	SL72	3:1	4->6	98.48	44.0	33.3	84.7	0.220	6.58
19	SL73	3:1	6	91.17	60.7	42.2	88.8	0.081	6.89
20	SL74	3:1	6	83.6	54.8	38.1	93.0	0.106	6.79
21	SL75	3:1	4->6	86.74	41.7	45.3	89.2	0.217	6.59
22	SL76	3:1	4->6	90.86	61.1	36.3	88.1	0.122	5.25
23	SL77	3:1	4->6	91.74	54.9	46.4	85.5	0.083	5.25
24	SL78	3:1	4->6	86.53	57.8	43.8	85.5	0.216	5.12
25	SL79	3:1	6	92.95	61.9	44.2	94.8	0.091	6.48
26	SL80	3:1	6	94	49.9	80.2	85.87	0.109	6.85
27	SL81	3:1	4->6	88.1	35.4	62.1	101.6	0.284	5.96

Biophysical characterisation of lipid nanoparticles

[0197] The messenger RNA-containing LNP composition will be characterized using analytical methods to determine the loading of messenger RNA, the percentage of messenger RNA that is encapsulated, and the size of the particles. The total amount of messenger RNA contained in the sample and the percentage of that messenger RNA that is encapsulated will be determined using a fluorescence assay employing Ribogreen, a dye that becomes more emissive upon binding messenger RNA. The total amount of messenger RNA will be determined by disrupting the LNP with 1wt% Triton-X 100 to expose the encapsulated messenger RNA, adding the dye, and comparing the emission intensity against a standard curve prepared using ribosomal RNA. The amount of unencapsulated messenger RNA will be measured in a similar manner with the detergent disruption of the LNP is omitted. With the total amount of messenger RNA known and the amount of unencapsulated messenger RNA known, the percent encapsulated messenger RNA will be calculated thus:

$$\text{Percent Encapsulation (\%)} = ((\text{RNA}_{\text{TOTAL}} - \text{RNA}_{\text{UNENCAPSULATED}}) / \text{RNA}_{\text{TOTAL}}) \times 100$$

where $\text{RNA}_{\text{TOTAL}}$ and $\text{RNA}_{\text{UNENCAPSULATED}}$ are, respectively, the concentrations of total messenger RNA and unencapsulated messenger RNA. The size of LNP will be measured using dynamic light scattering of a sample diluted 1:100 in PBS buffer.

pKa Protocol

[0198] In a black 96-well plate, solutions of samRNA-LNP (final assay concentration 2 $\mu\text{g/mL}$ total RNA) in a series of buffers ranging from pH 4 to 9.5 will be prepared. Buffers between pH 4 to 7.6 will be prepared from disodium phosphate and citric acid. Buffers from 7.8 to 9.5 will be prepared by titrating tris buffer with 10 N sodium hydroxide. To each well will be added 6-(p-Toluidino)-2-naphthalenesulfonic acid sodium salt (TNS) in water to a final assay concentration of 6 μM . The fluorescence will be read on a plate reader at 25 °C with an excitation setting of 321 nm and an

emission setting of 445 nm. The intensity values will be plotted as a function of pH using GraphPad Prism and be fit with a sigmoidal dose response curve. The apparent pKa will be determined as the EC50 of this curve where half of the ionizable amines are expected to be protonated.

5

Characterization of the *in vitro* potency of LNPs

[0199] The ability of the samRNA-LNPs to transfect cultured cells was characterized using an in vitro assay based on the percentage of cells expressing the antigen of interest. Specifically, for each LNP, 1 million BHK-21 cells were co-incubated
10 with samRNA-LNP of varying concentration in 2 mL of media for 17-19 hours at 37°C with 5% CO₂. Subsequently, the cells were treated with TrypLE (Gibco) to detach from the dish to form a single-cell suspension, fixed and permeabilized (BD Cytofix/Perm kit), and then stained with fluorophore-labeled antigen-specific antibodies against H5 and N1. The percentage of double antigen positive cells was quantified using a BD Accuri flow
15 cytometer.

[0200] The results of the study are shown in the below table:

Lipid	[Potency]/[Potency LKY750]	R square
SL57	1.029	0.9990
SL62	1.867	0.9941
SL63	1.242	0.9985
SL67	5.282	0.9903
SL69	1.647	0.9915
SL70	1.021	0.9962
SL71	1.283	0.9962
SL73	3.204	0.9973
SL74	1.802	0.9800
SL79	1.187	0.9994
SL80	4.132	0.9995
LKY750	1	0.9907

[0201] The ability of the samRNA-LNPs to transfect cultured cells was characterized using a high throughput microscopy imaging-based assay. In summary,
20 4×10⁵ BHK-21 cells were incubated with a concentration range of the LNPs varying from

0-1 ng in presence of 4% FBS and the DMEM media for over 18 hr in a PS-96 well flatbottom plates at 37 °C with 5% CO₂. Subsequently, the cells are fixed and permeabilized using the BD Cytfix/Perm kit) reagent, followed by staining with the (in-house) human-anti-H5 specific antibody using the Thermofisher cy5-zenon-anti-human labelling kit and DAPI. The plates are further imaged using Agilent high throughput cytation5 instrument. The analysis is performed by counting the DAPI spots as the total cell number and the cy-5 red spots as the number of antigen positive cells. To increase the linearity range of the assay, the absolute value of natural log of percentage of negative cells i.e., $\ln(1 - (\text{total red spot})/(\text{total blue spots}))$ was calculated. Reference lipid LKY750 (2,5-bis((9z,12z)-octadeca-9,12,dien-1-yloxy)benzyl-4-(dimethylamino)butanoate; WO 2016/037053 herein incorporated by reference) was always included as a positive control for normalization. The normalized absolute numbers of lipids to LKY750 was reported. The higher the normalized absolute value for a sample, the higher the transfection rate of that particular LNP chemistry is. The assay exhibits average value of 20% with minimum and maximum of 10 and 30% variability in nature, respectively.

[0202] The results are depicted in Figures 1 and 2, in which the potency of each of the lipid formulations is assessed relative to LKY750 control. Both the FACS and Cytation 5 potency results indicate the assessed lipid formulations show improved, or at least comparable (i.e., CY089 with SL57), *in vitro* potency to the LKY750 control CY087.

Characterization of the *in vivo* potency of LNPs

[0203] The ability of mRNA-LNPs to act as a vaccine was evaluated by measuring the antibody- and cell-based immune response following a prime-boost vaccination schedule. A priming vaccination was given to Balb/c mice on Day 0 via intramuscular injection (i.m.) and followed 21 days later with a boosting vaccination. After an additional 21 days (Day 42 of the experiment), the mice were sacrificed and serum and splenocytes collected for further analysis.

[0204] Serum was analysed for vaccine-specific antibody response using an IgG enzyme linked immunosorbent assay (IgG ELISA), a pseudovirus microneutralization

(MN) assay, a hemagglutination inhibition (HAI) assay, and a neuraminidase inhibition enzyme-linked lectin assay (ELLA). In HAI, NA inhibition and MN assays as shown in FIG 3A-D, SL57 gave comparable activity to the comparator lipid LKY750. In MN assay, SL60 gave comparable activity to LKY750. In ELISA, HAI, MN assays, both
5 SL57 and SL60 gave comparable activity to the adjuvanted inactivated virus vaccine (aH5N1).

[0205] The results will be compared to an adjuvanted inactivated virus vaccine (aH5N1) tested in the same experiment using a one-way ANOVA statistical test.

[0206] Splenocytes from the experiment described above will be analysed for
10 antigen-specific cytokine production using intracellular cytokine staining by flow cytometry following *in vitro* peptide stimulation. Splenocytes will be pooled (n=5/group) and stimulated *ex vivo* in the absence or presence of H5 or N1 peptides in duplicate. The mean level of response, measured by interferon gamma, interleukin-2, and/or tumor necrosis factor alpha production, will be displayed and error bars display the
15 measurement precision.

[0207] Taken together, these *in vivo* results are expected to demonstrate that LNPs prepared using the lipids SL56-SL81 are immunogenic and effective as an influenza vaccine in this preclinical model.

20 Characterization of the *in vivo* transfection efficiency of LNPs

[0208] The *in vivo* potency of mRNA-LNPs will be evaluated using samRNA expressing the reporter protein firefly luciferase in order to quantify the location, relative amount, and the duration of protein expression. LNPs formulated with a mRNA expressing luciferase will be injected into mice, for example intramuscularly in the hind
25 leg. At defined time points, such as daily, the mice will be administered luciferin and the bioluminescence will be imaged and quantified.

Characterization of genotoxic potential, tolerability, biodistribution, and biodegradability

[0209] A standard test battery for prediction of genotoxic potential (damage of DNA) will be used, as no single test can detect all genotoxic mechanisms leading to tumorigenicity. As an example, when a positive result is seen in *in vitro* mammalian cell assay, clearly negative results in two *in vivo* assays, in appropriate tissues and demonstrated sufficient test substance exposure, will be considered evidence for lack of genotoxic potential *in vivo*. Relevant guidelines, such as S2(R1) Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use”, will be followed.

- Test for gene mutation in bacteria.
- *In vitro* cytogenetic test for chromosomal damage (metaphase chromosome aberration or micronucleus test) or mouse lymphoma Tk gene mutation assay.
- *In vivo* genotoxicity test for chromosomal damage using rodent hematopoietic cells either for micronuclei or chromosomal aberrations in metaphase cells.

In silico screening of novel lipids for toxicity

[0210] Compounds of Formula I and/or Formula II will be screened for potential toxicity and mutagenicity using commercially available computational toxicology assessment products and/or services to satisfy International “Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) M7(R1) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals To Limit Potential Carcinogenic Risk”. For example, two complementing Quantitative (Q)SAR methods, expert rule-based Derek Nexus and statistical-based Sarah Nexus (or Leadscope), will be used to predict outcome of bacterial mutagenicity. Classification will be based on Class 1 to 5 with respect to mutagenic and carcinogenic potential from known mutagenic carcinogen (Class 1) to lack of mutagenicity or carcinogenicity (Class 5). The outcome of computer-based analysis will be reviewed for relevance of positive, negative, conflicting, or inconclusive prediction and rationale for conclusion provided.

[0211] As most structural alerts are based on bacterial mutagenicity, compounds with structural alert can be detected in standard test battery. In addition, some chemical classes are more easily detected in mammalian cell chromosome damage assays than

bacterial mutation assays. A negative result from a compound with structural alert in either test battery should be considered not genotoxic.

In vitro screening for genotoxicity

- 5 [0212] Compounds of Formula I and/or Formula II will be assessed first for mutagenicity in bacterial reverse gene mutation test (Ames), which detects relevant genetic changes and most genotoxic rodent and human carcinogens and then for genotoxicity *in vitro* micronucleus (MN) assay using mammalian cells:

- [0213] Ames assay will follow Organization for Economic Co-operation and
10 Development (OECD) Guidelines for the Testing of Chemicals No. 471 Bacterial Reverse Mutation Test and S2(R1) Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use to assess mutagenic potential. At least five strains of bacteria *Salmonella typhimurium* TA98, TA100, TA1537, TA1535 and TA102 or *E. coli* WP2 will be exposed to test substance in the
15 presence and absence of metabolic activation system rat liver metabolizing system (S-9).

- [0214] *In vitro* mammalian cell micronucleus genotoxicity test will follow OECD Guidelines for the Testing of Chemicals No. 487. This test should detect MN in the cytoplasm of interphase cells and chromosome damaging potential (aneugens and clastrogens). Sufficiently validated and appropriate are mouse lymphoma L5178Y cell
20 Tk (thymidine kinase) gene mutation assay (MLA) and human lymphocyte micronucleus assay (HLM) treated in the absence and presence of rat liver metabolising system (S-9).

In vivo screening for genotoxicity

- [0215] Compounds of Formula I and/or Formula II will be assessed *in vivo* in either
25 acute or repeat-dose rat study, as some agents are mutagenic *in vivo* but not *in vitro*. The choice of analysis will be micronuclei in erythrocytes (in blood or bone marrow) or chromosome aberrations in metaphase cells in bone marrow.

Biodistribution and Biodegradability assessment

[0216] ***Ex vivo* biodegradability assessment:** Compounds of Formula I and/or Formula II will be screened for enzymatic biodegradation using enzyme-containing solutions prepared from relevant species, such as human, mouse, and rat. Lipids will be screened both in their neat form (unformulated) and when incorporated into LNPs that do or do not contain RNA. For example, a novel ionizable cationic lipid will be diluted at an approximate concentration of 1.0 to 0.001 mg/mL in an aqueous solution of human or rat liver microsomes and incubated at temperatures in the range 25-37 °C for durations from 0.1-24 hours. Subsequently, the amount of intact lipid remaining will be measured using liquid chromatography (LC) with evaporative light scattering (ELS) or mass spectrometry (MS) detection and compared to control samples that were not treated with enzymes or were incubated at 2-8 °C to inhibit enzymatic activity. Furthermore, the appearance of new peaks in the chromatograph, which presumably represent degradation products, will be investigated using MS to confirm identity. Without wishing to be bound by theory, the expectation is that degradation may occur via ester hydrolysis and that the lipid's structure will influence the rate of ester hydrolysis.

[0217] ***In vivo* biodistribution, pharmacokinetics, and biodegradability assessment:** Assessment of biodistribution and pharmacokinetics [(Absorption, Distribution, Metabolism, and Excretion) (ADME)] will be performed in rat, which will receive the novel lipids (in their neat form and/or incorporated into LNPs that do or do not contain RNA) intramuscularly or intravenously.

[0218] Assessment of biodistribution will be based on the principles of ICH-M3(R2) and World Health Organization (WHO) guidelines on nonclinical evaluation of vaccines, WHO Technical Report Series No. 927, Annex 1. Blood samples will be collected multiple times for lipid, mRNA and immunogenicity analysis. Tissues will be collected at predefined days and lipid analyses performed with qualified liquid chromatography-mass spectrometry method (LS-MS/MS). If needed, RT qPCR analyses of the nucleic acid payload will be performed.

[0219] Lipid pharmacokinetics in plasma will be evaluated with qualified LS-MS/MS.

[0220] ***In vitro* biodegradability assessment:** Compounds of Formula I and/or II will be screened by testing the stability of the test compound in the presence of cryopreserved hepatocytes. For example, a suspension of cryopreserved hepatocytes (final cell density 0.5×10^6 viable cells/mL in Williams E media supplemented with 2 mM L glutamine and 25 mM HEPES) are pre-incubated at 37 °C prior to the addition of
5 test compound (final substrate concentration 1 μ M; final DMSO concentration 0.25 %) to initiate the reaction. Two control compounds are included with each species. Each compound is incubated for 0, 5, 10, 20, 40 and 60 min at 37 °C. The reactions are stopped by transferring incubate into acetonitrile at the appropriate time points, in a 1:3 ratio. The
10 termination plates are centrifuged at 3,000 rpm for 30 min at 4 °C to precipitate the protein. Following protein precipitation, the sample supernatants are combined in cassettes of up to four compounds, internal standard is added and samples analysed using Cypotex generic LC MS/MS conditions.

15 *In vivo* tolerability assessment

[0221] The tolerability of Compounds of Formula I and/or Formula II will be evaluated in appropriate species, such as mice or rats. For example, groups of Sprague-Dawley rats will be administered novel-lipid-containing RNA LNP vaccines at doses ranging from 0.1 to 30 μ g total RNA per animal via intramuscular injection. Each animal
20 will be given between 1-3 injections and the effect of treatment on weight loss, food intake, and body temperature will be measured. In addition, the injection site will be monitored using the Draize dermal irritation scoring system. In addition, serum will be collected to evaluate the vaccine immune response using ELISA and functional antibody assays. In addition, the complete blood count, blood chemistry, and blood coagulation
25 will be assessed at one or more timepoint during the study. At the end of the study during necropsy select organs may be macroscopically assessed and weighed. Novel lipids associated with RNA LNP vaccines that maintain high immunogenicity while causing the least undesired responses, such as weight loss, fever, or injection site reactogenicity, will be considered more promising candidates over lipids with a less attractive
30 tolerability profile.

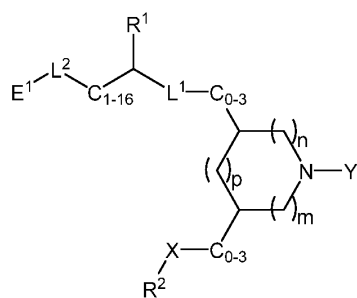
Toxicity assessment

[0222] ***Repeat-dose general toxicity study:*** Toxicity of the candidate Compounds of Formula I and/or Formula II will be assessed in a pharmacological relevant nonclinical
5 species (e.g. rat). Novel-lipid-containing RNA/LNP vaccines will be administered in Sprague-Dawley rats once every three weeks (total of 3 doses) via intramuscular injection. The reversibility or persistence of any effects will be assessed after a 3-week recovery phase.

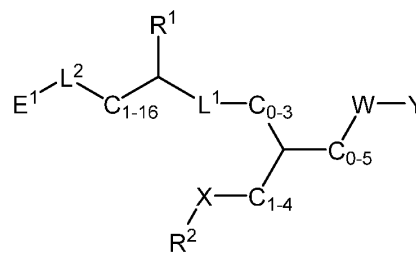
[0223] Assessment of toxicity will be based on mortality, clinical observations,
10 body weights, food consumption, ophthalmic observations, dose site (dermal) observations, body temperatures, and clinical and anatomic pathology, micronucleus analysis. Blood samples will be collected for immunogenicity analysis.

Itemized Listing of Embodiments

1. A compound of Formula I or a compound of Formula II:



Formula I;



Formula II;

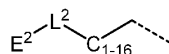
- 5 or a pharmaceutically acceptable salt, prodrug, or stereoisomer thereof, wherein:

X is selected from the group consisting of -S-, -O- and -C-;

E¹ is a linear or branched -C₁₋₃₀-alkyl;

R¹ is selected from the group consisting of -H and Formula IA;

R² is selected from the group consisting of linear or branched -C₁₋₃₀-alkyl and
 10 Formula IA;



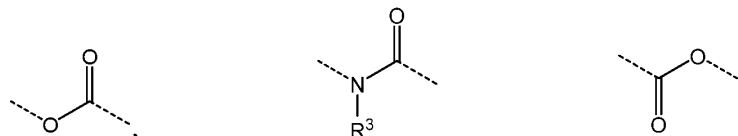
Formula IA;

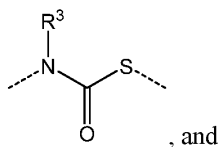
E², if present, is a linear or branched -C₁₋₃₀-alkyl;

m and n are each independently an integer from 0 to 3;

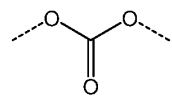
- 15 p is an integer from 0 to 2;

L¹ is selected from the group consisting of:





, and



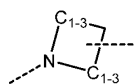
;

R³, if present, is selected from the group consisting of -H and linear or branched -C₁₋₈-alkyl;

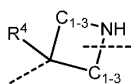
L² is selected from -OC(O)- and -C(O)O-;

W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:

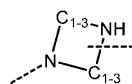
5



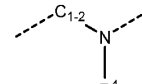
Formula IIA;



Formula IIB;



Formula IIC;



Formula IID;

R⁴, if present, is selected from the group consisting of -H and linear or branched -C₁₋₅-alkyl;

Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol; and

10 dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

2. The compound of item 1, wherein X is -S-.

3. The compound of item 1 or item 2, wherein E¹ is a linear -C₁₋₃₀-alkyl.

4. The compound of item 1 or item 2, wherein E¹ is a branched -C₁₋₃₀-alkyl.

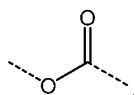
15 5. The compound of any one of items 1 to 4, wherein R¹ is -H.

6. The compound of any one of items 1 to 4, wherein R¹ is Formula IA.

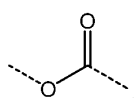
7. The compound of any one of items 1 to 4 and 6, wherein R¹ is Formula IA, and E² is a linear -C₁₋₃₀-alkyl.

8. The compound of any one of items 1 to 4 and 6, wherein R^1 is formula IA, and E^2 is a branched $-C_{1-30}$ -alkyl.

9. The compound of any one of items 1 to 8, wherein L^1 is:



10. The compound of any one of items 1 to 9, wherein X is $-S-$, and L^1 is:



11. The compound of any one of items 1 to 10, wherein Y is a linear or branched $-C_{1-5}$ -alkyl.

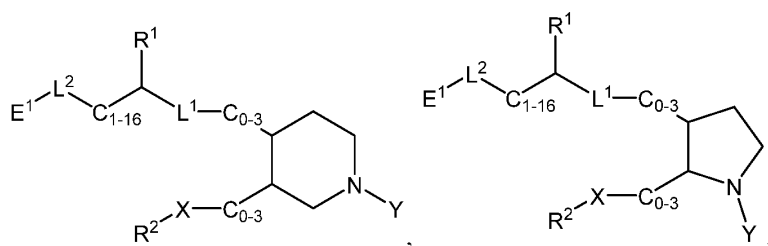
12. The compound of any one of items 1 to 10, wherein Y is a linear or branched C_{1-5} -alcohol.

13. The compound of any one of items 1 to 10 and 11, wherein Y is a linear C_{1-5} -alkyl.

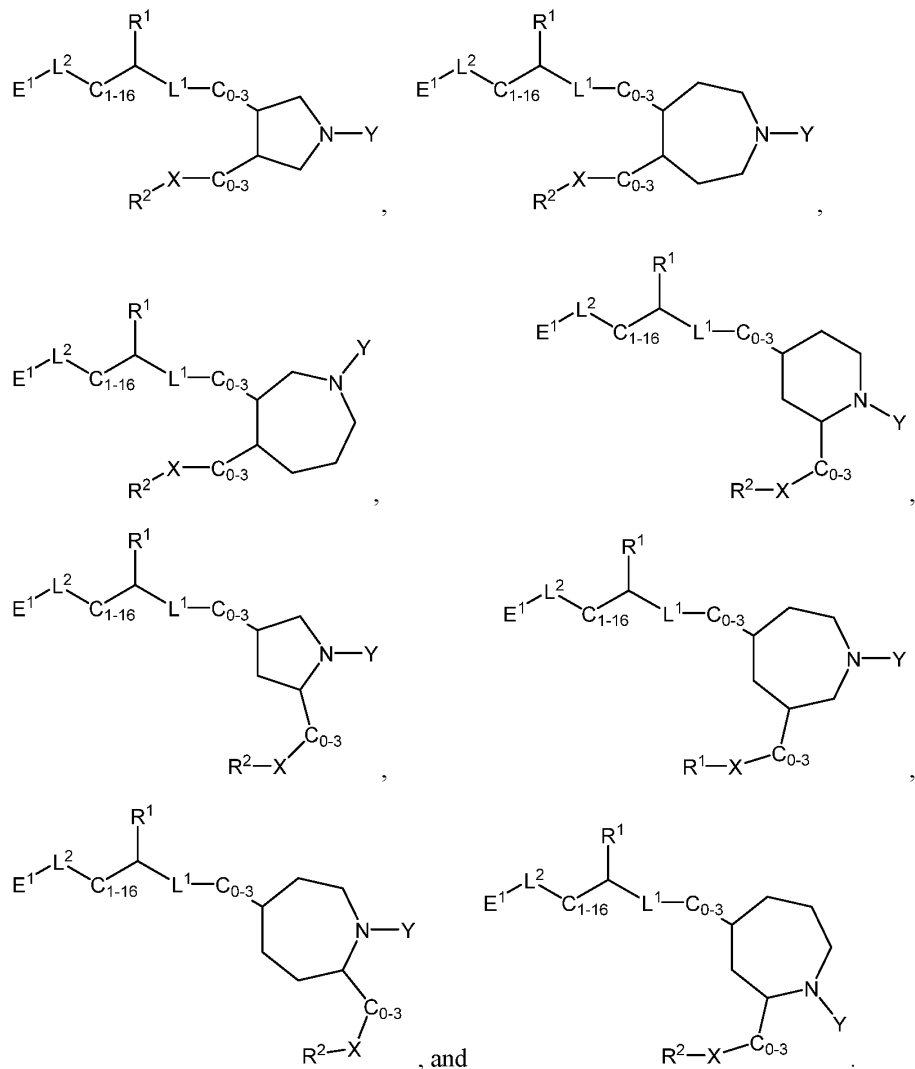
14. The compound of any one of items 1 to 10 and 12, wherein Y is a linear C_{1-5} -alcohol.

15. The compound of item 1, wherein the compound is a compound of Formula I.

16. The compound of item 1 and item 15, wherein the compound of Formula I is selected from the group consisting of:

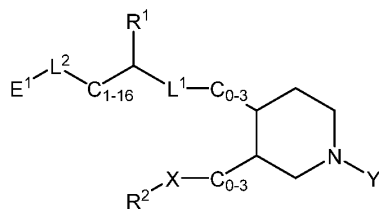


198

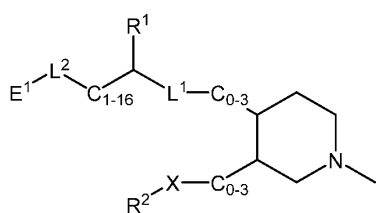


- 5 17. The compound of any one of items 1, 15, and 16, wherein the compound of Formula I is:

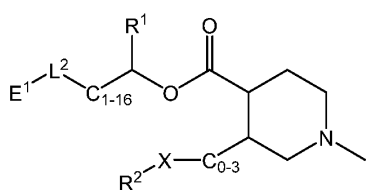
199



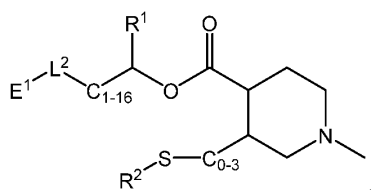
18. The compound of any one of items 1 and 15 to 17, wherein the compound of Formula I is:



5 19. The compound of any one of items 1 and 15 to 18, wherein the compound of Formula I is:



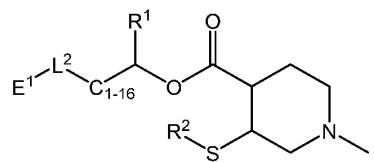
20. The compound of any one of items 1 and 15 to 19, wherein the compound of Formula I is:



10

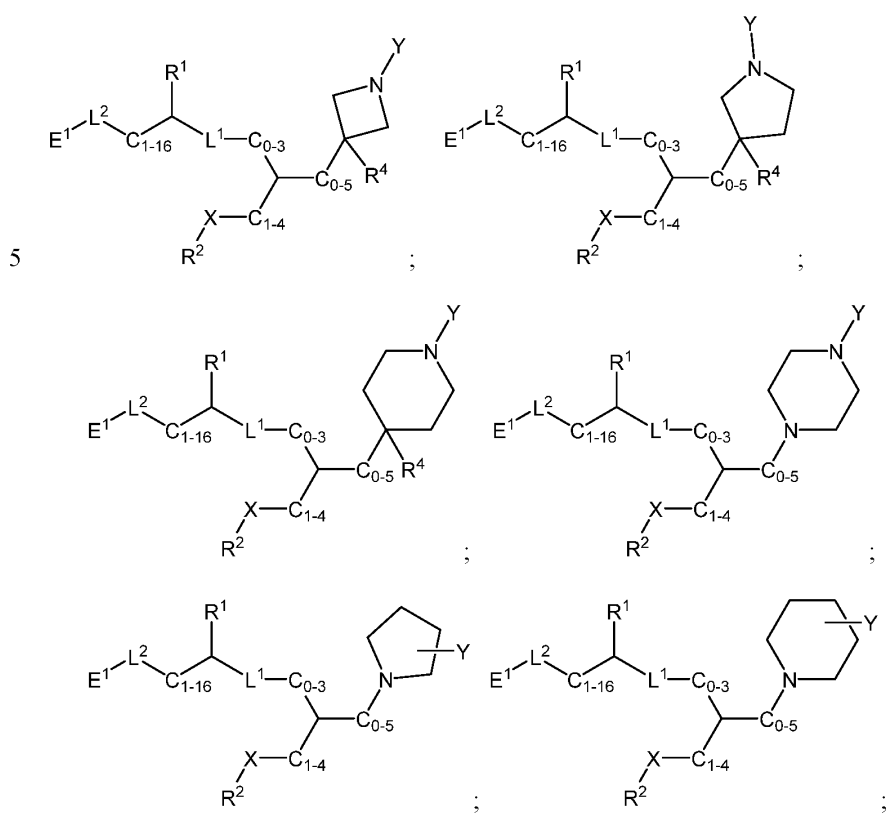
21. The compound of any one of items 1 and 15 to 20, wherein the compound of Formula I is:

200

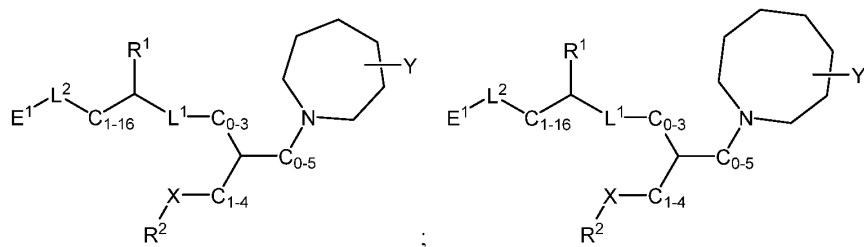


22. The compound of item 1, wherein the compound is a compound of Formula II.

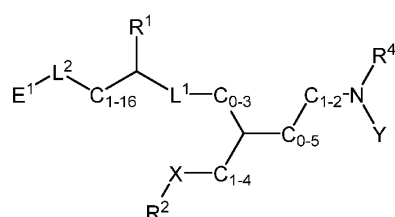
23. The compound of item 1 and item 22, wherein the compound of Formula II is selected from the group consisting of:



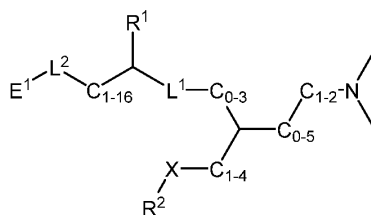
201



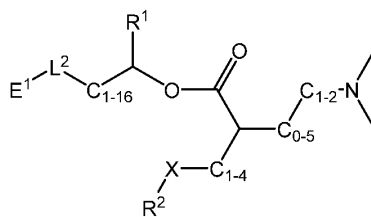
and



24. The compound of any one of items 1, 22, and 23, wherein the compound of
5 Formula II is:

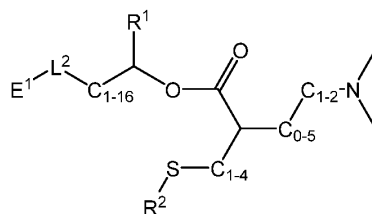


25. The compound of any one of items 1 and 22 to 24, wherein the compound of Formula II is:

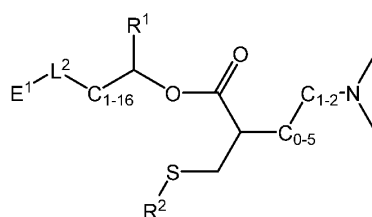


- 10 26. The compound of any one of items 1 and 22 to 25, wherein the compound of Formula II is:

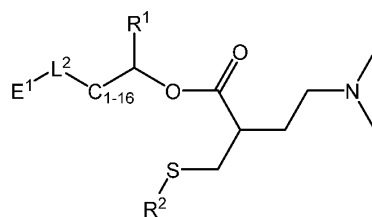
202



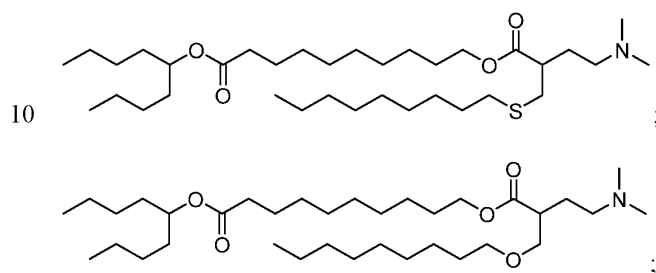
27. The compound of any one of items 1 and 22 to 26, wherein the compound of Formula II is:

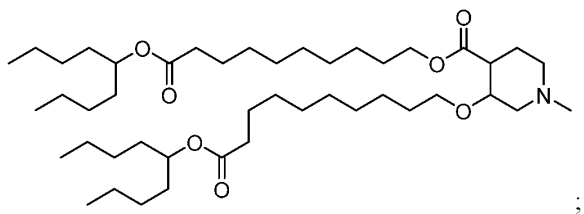
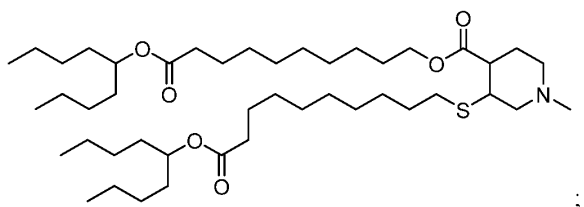
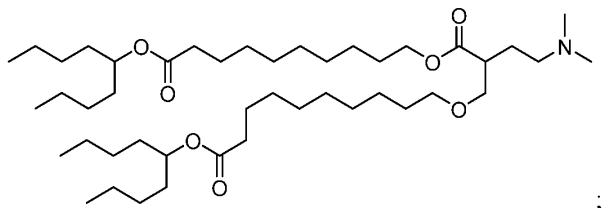
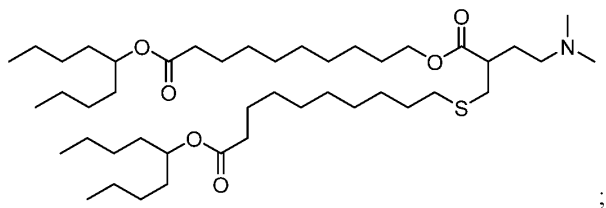
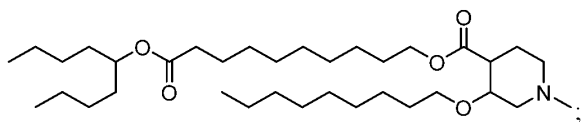
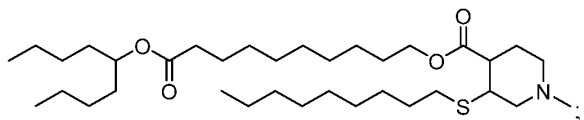


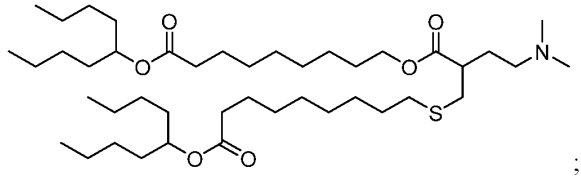
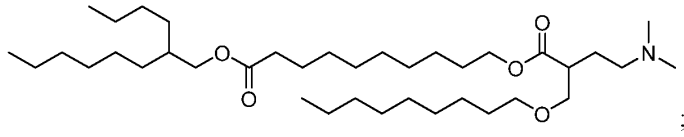
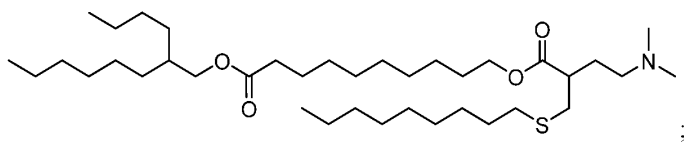
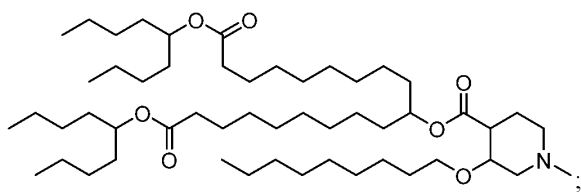
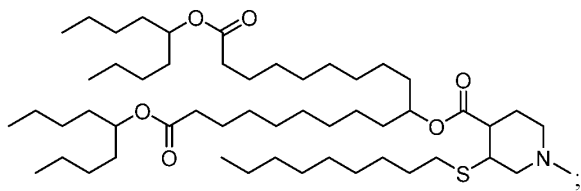
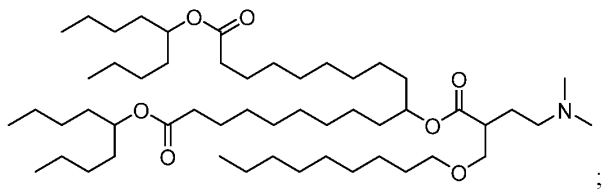
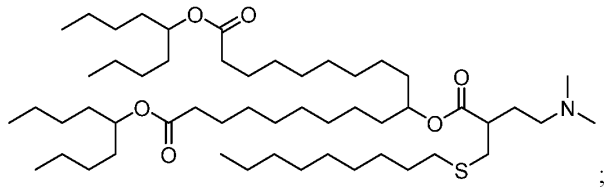
28. The compound of any one of items 1 and 22 to 27, wherein the compound of Formula II is:



29. The compound of item 1, wherein the compound is selected from the group consisting of:

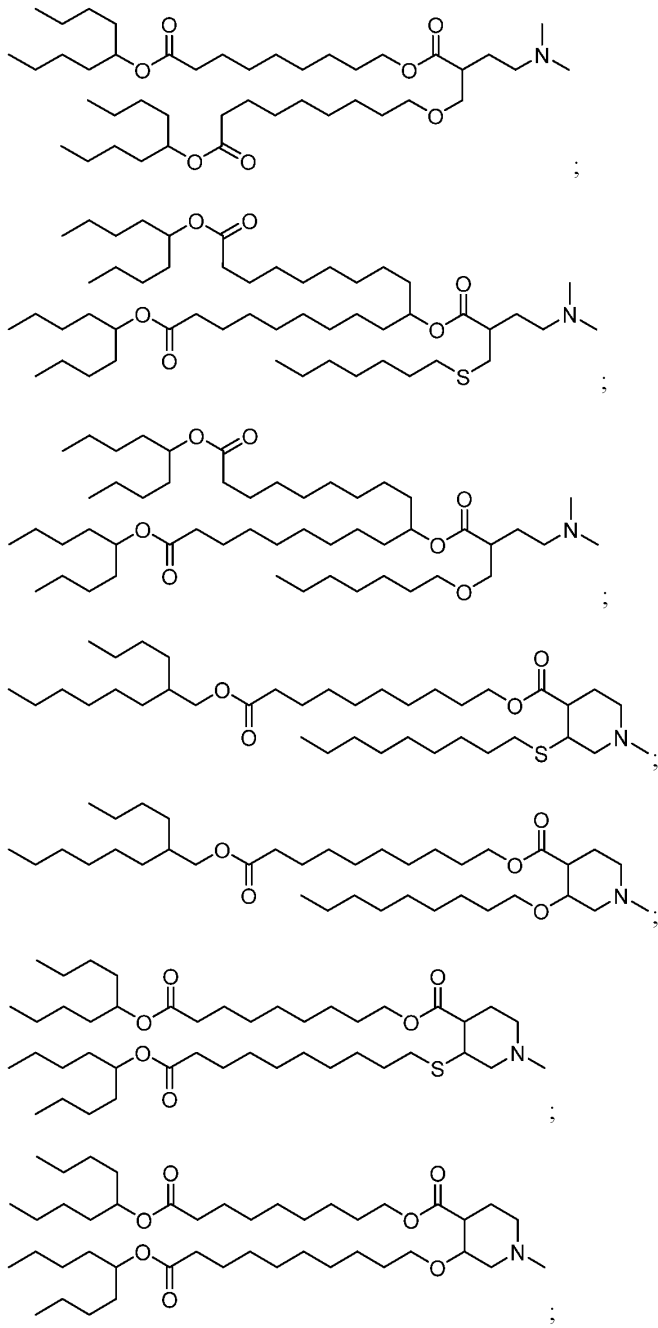




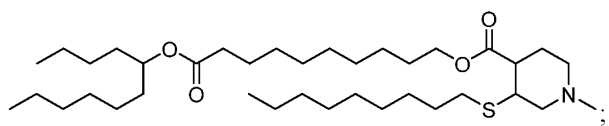
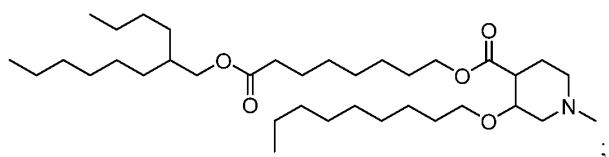
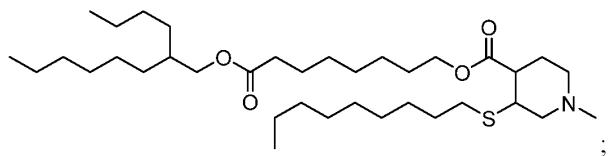
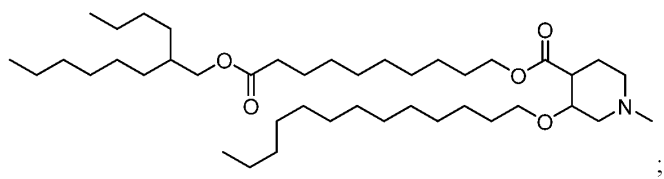
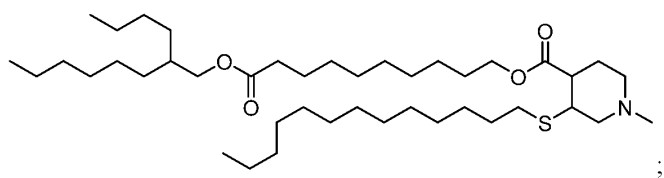
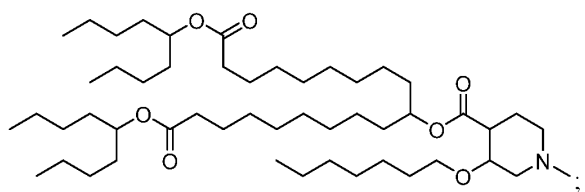
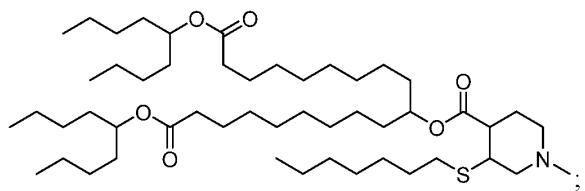


5

205

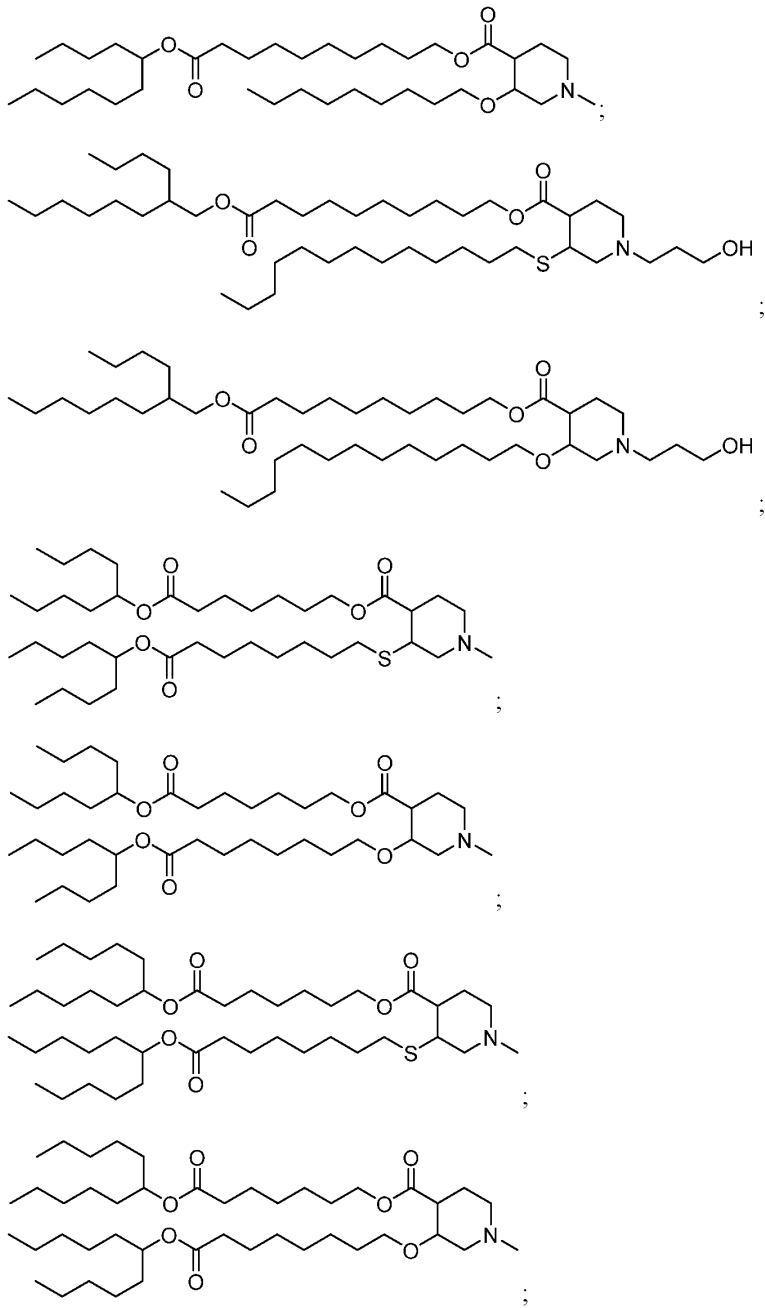


5



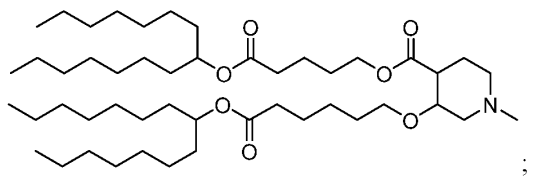
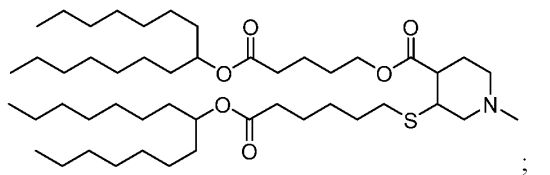
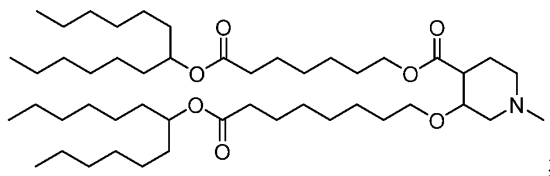
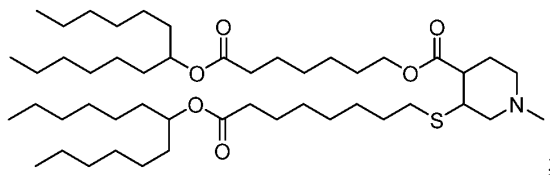
5

207

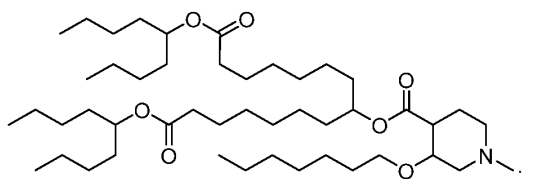
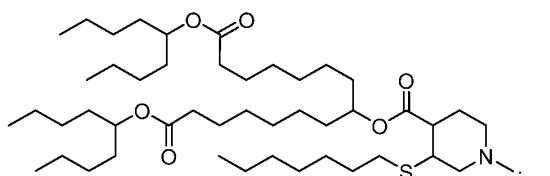


5

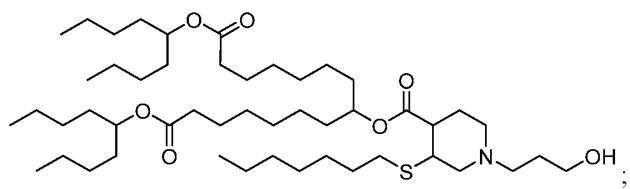
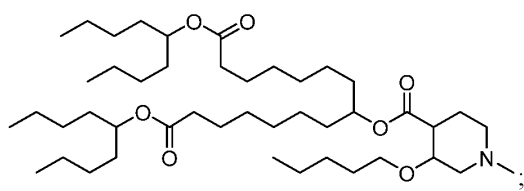
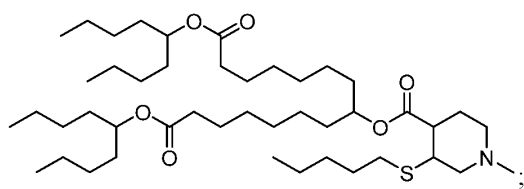
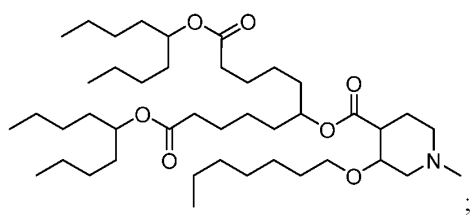
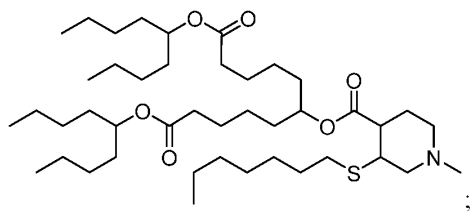
208



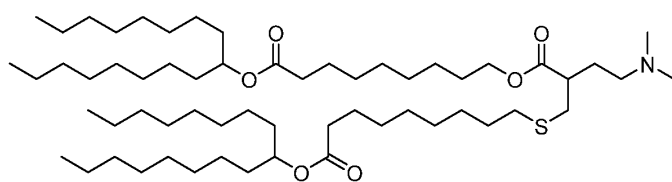
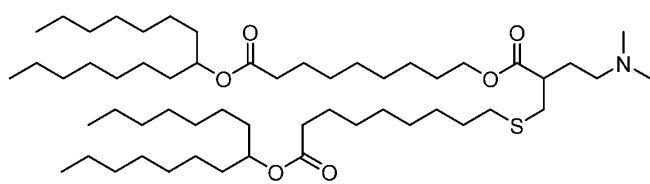
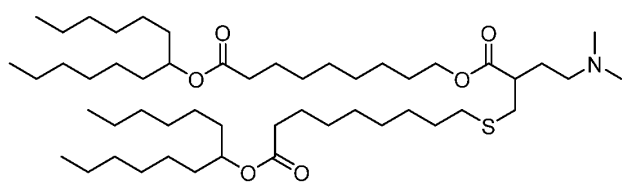
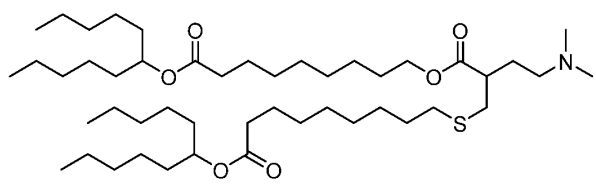
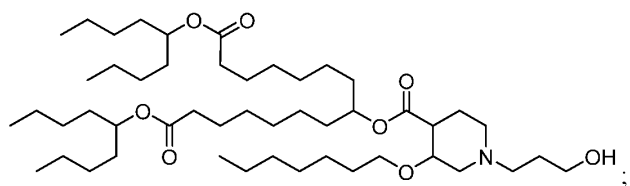
5



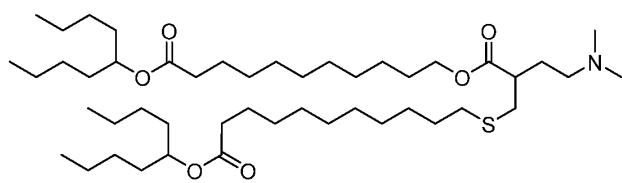
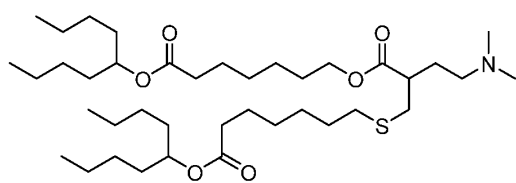
209



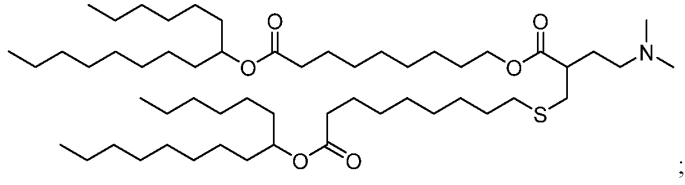
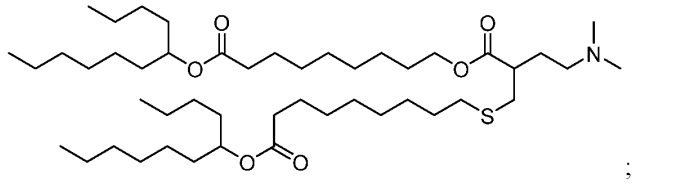
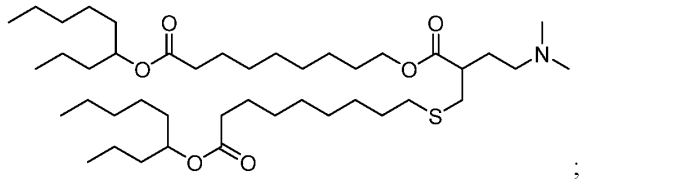
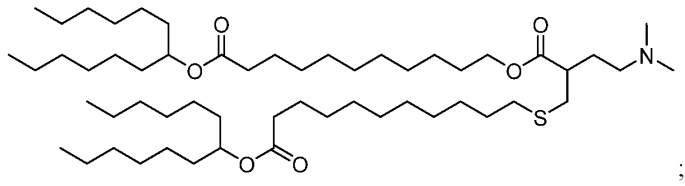
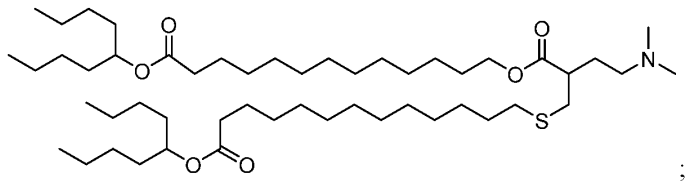
5



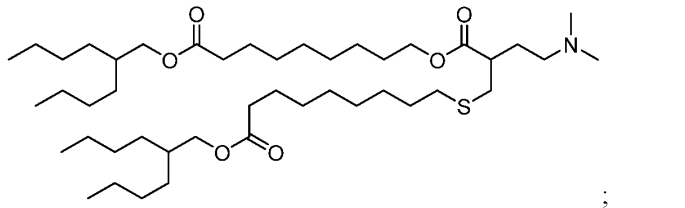
5



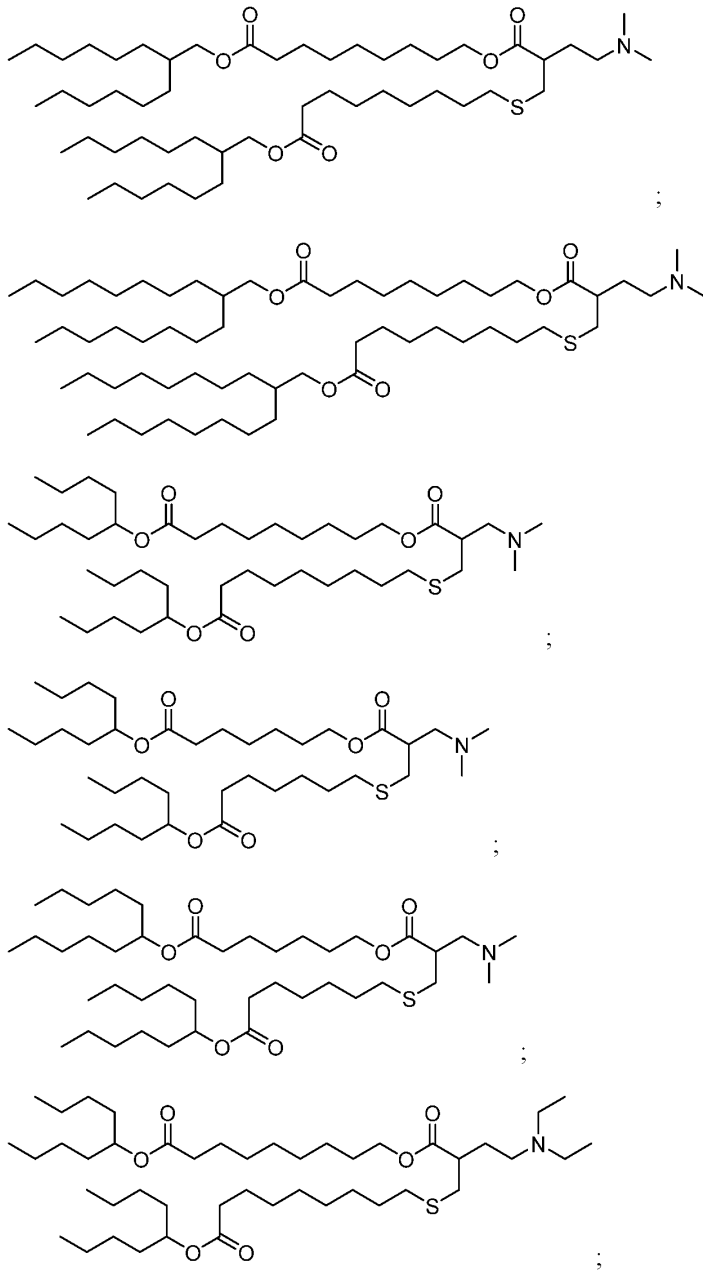
211



5

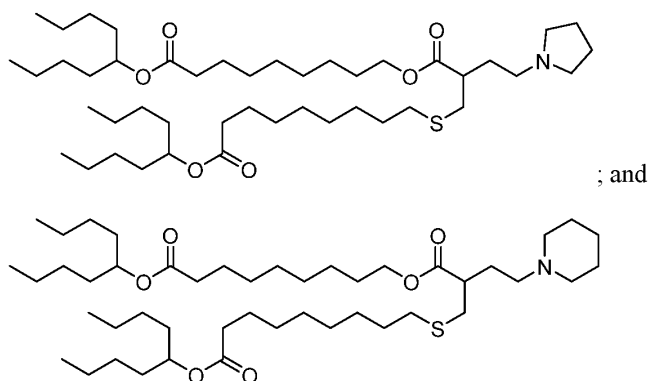


212



5

213



30. A lipid nanoparticle (LNP) comprising a lipid component comprising a
5 compound of any one of items 1 to 29.

31. The lipid nanoparticle of item 30, wherein the lipid component further comprises one or more of a neutral lipid, a structural lipid, and a PEGylated lipid.

32. The lipid nanoparticle of item 31, wherein the neutral lipid is selected from the group consisting of 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dioleoyl-
10 sn-glycero-3-phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-
15 glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OchemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-
20 glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-

dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), and sphingomyelin.

33. The lipid nanoparticle of item 31, wherein the structural lipid is selected from the group consisting of cholesterol, fecosterol, sitosterol, campesterol, stigmasterol, 5 brassicasterol, ergosterol, tomatidine, tomatine, ursolic acid and alpha-tocopherol.

34. The lipid nanoparticle of item 31, wherein the PEGylated lipid is selected from the group consisting of PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified dialkylamines, PEG-modified diacylglycerols, and PEG-modified dialkylglycerols, optionally PEG-c-
10 DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, and PEG-DSPE.

35. The lipid nanoparticle of item 31, wherein the lipid component comprises: about 25 mol % to about 60 mol % of a compound of claim 1; about 2 mol % to about 25 mol % neutral lipid; about 18.5 mol % to about 60 mol % structural lipid; and about 0.2 mol % to about 10 mol % of PEGylated lipid.

15 36. The lipid nanoparticle of any one of items 30 to 35, wherein the lipid nanoparticle further comprises a polynucleotide.

37. The lipid nanoparticle of item 36, wherein the polynucleotide is selected from the group consisting of: a messenger RNA (mRNA), a self-amplifying mRNA (sa-mRNA), a small interfering RNA (siRNA), a microRNA (miRNA), miRNA inhibitors
20 (antagomirs/antimirs), messenger-RNA-interfering complementary RNA (micRNA), short hairpin RNA (shRNA), multivalent RNA, dicer substrate RNA, an antisense oligonucleotide, plasmid DNA, DNA, and complementary DNA (cDNA).

38. The lipid nanoparticle of item 36 or item 37, wherein the polynucleotide is a ribonucleic acid (RNA).

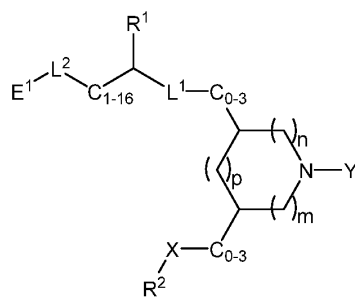
25 39. The lipid nanoparticle of any one of items 36 to 38, wherein the polynucleotide is a conventional or self-amplifying mRNA.

40. The lipid nanoparticle of any one of items 30 to 39, wherein the lipid nanoparticle has a diameter of from about 30 nm to about 160 nm.

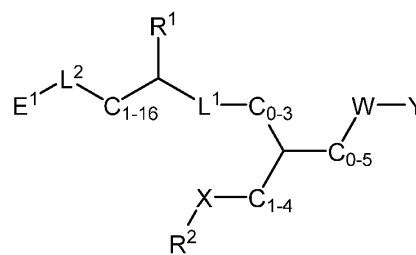
41. A pharmaceutical composition comprising a plurality of lipid nanoparticles of any one of items 30 to 40, and a pharmaceutically acceptable carrier.
42. A method of delivering a polynucleotide to a mammalian cell, including administering the lipid nanoparticle of any one of items 30 to 40, or the pharmaceutical composition of item 41, to a subject to thereby contact the cell with the lipid nanoparticle or pharmaceutical composition and deliver the polynucleotide to the cell.
43. The method of item 42 wherein the cell is a cell of a human subject.
44. A method of producing a polypeptide of interest in a mammalian cell, including the step of contacting the cell with a lipid nanoparticle of any one of items 30 to 40, or the pharmaceutical composition of item 41, wherein the lipid nanoparticle comprises a conventional mRNA or a self-amplifying mRNA encoding the polypeptide.
45. A method of treating a disease, disorder or condition in a subject in need of such treatment, comprising administering a lipid nanoparticle of any one of items 30 to 40, or the pharmaceutical composition of item 41, to the subject to thereby treat the disease, disorder or condition.
46. The method of item 45, wherein the disease, disorder or condition is selected from the group consisting of a rare disease, an infectious disease, cancer, a proliferative disease, a genetic disease, an autoimmune disease, diabetes, a neurodegenerative disease, a cardiovascular disease, a reno-vascular disease and a metabolic disease.
47. A vaccine comprising a lipid nanoparticle of any one of items 30 to 40, or the pharmaceutical composition of item 40, and a mRNA encoding a polypeptide.
48. The vaccine of item 47, wherein the vaccine is selected from a tumor vaccine, an influenza vaccine, and a SARS, including a SARS-CoV-2, vaccine.

CLAIMS

1. A compound of Formula I or a compound of Formula II:



Formula I;



Formula II;

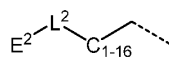
or a pharmaceutically acceptable salt, prodrug, or stereoisomer thereof, wherein:

X is selected from the group consisting of -S-, -O- and -C-;

E¹ is a linear or branched -C₁₋₃₀-alkyl;

R¹ is selected from the group consisting of -H and Formula IA;

R² is selected from the group consisting of linear or branched -C₁₋₃₀-alkyl and Formula IA;



Formula IA;

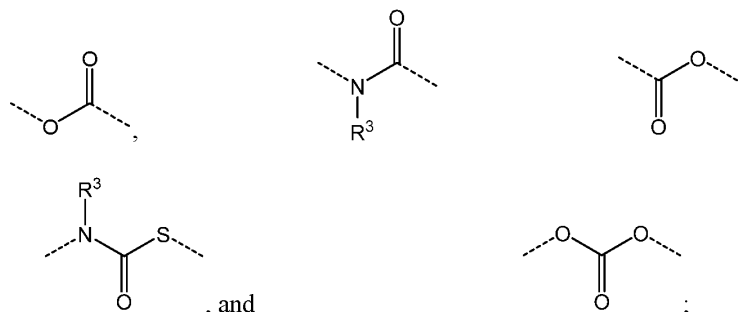
E², if present, is a linear or branched -C₁₋₃₀-alkyl;

m and n are each independently an integer from 0 to 3;

p is an integer from 0 to 2;

L¹ is selected from the group consisting of:

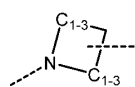
217



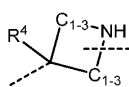
R^3 , if present, is selected from the group consisting of -H and linear or branched -C₁₋₈-alkyl;

L^2 is selected from -OC(O)- and -C(O)O-;

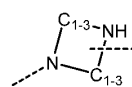
W is selected from the group consisting of Formula IIA, Formula IIB, Formula IIC, and Formula IID:



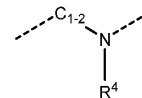
Formula IIA;



Formula IIB;



Formula IIC;



Formula IID;

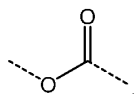
R^4 , if present, is selected from the group consisting of -H and linear or branched -C₁₋₅-alkyl;

Y is selected from the group consisting of -H, linear or branched -C₁₋₅-alkyl, and linear or branched -C₁₋₅-alkanol; and

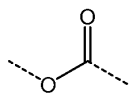
dashed lines represent a bond to an adjacent atom in the compound of Formula I or the compound of Formula II.

2. The compound of claim 1, wherein X is -S-.
3. The compound of claim 1, wherein E^1 is a linear -C₁₋₃₀-alkyl.
4. The compound of claim 1, wherein E^1 is a branched -C₁₋₃₀-alkyl.
5. The compound of claim 1, wherein R^1 is -H.
6. The compound of claim 1, wherein R^1 is Formula IA.

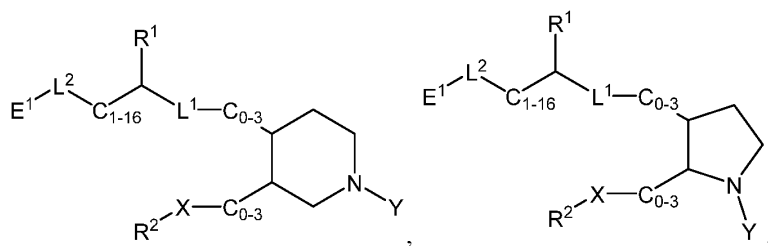
7. The compound of claim 1, wherein R^1 is Formula IA, and E^2 is a linear $-C_{1-30}$ -alkyl.
8. The compound of claim 1, wherein R^1 is formula IA, and E^2 is a branched $-C_{1-30}$ -alkyl.
9. The compound of claim 1, wherein L^1 is:



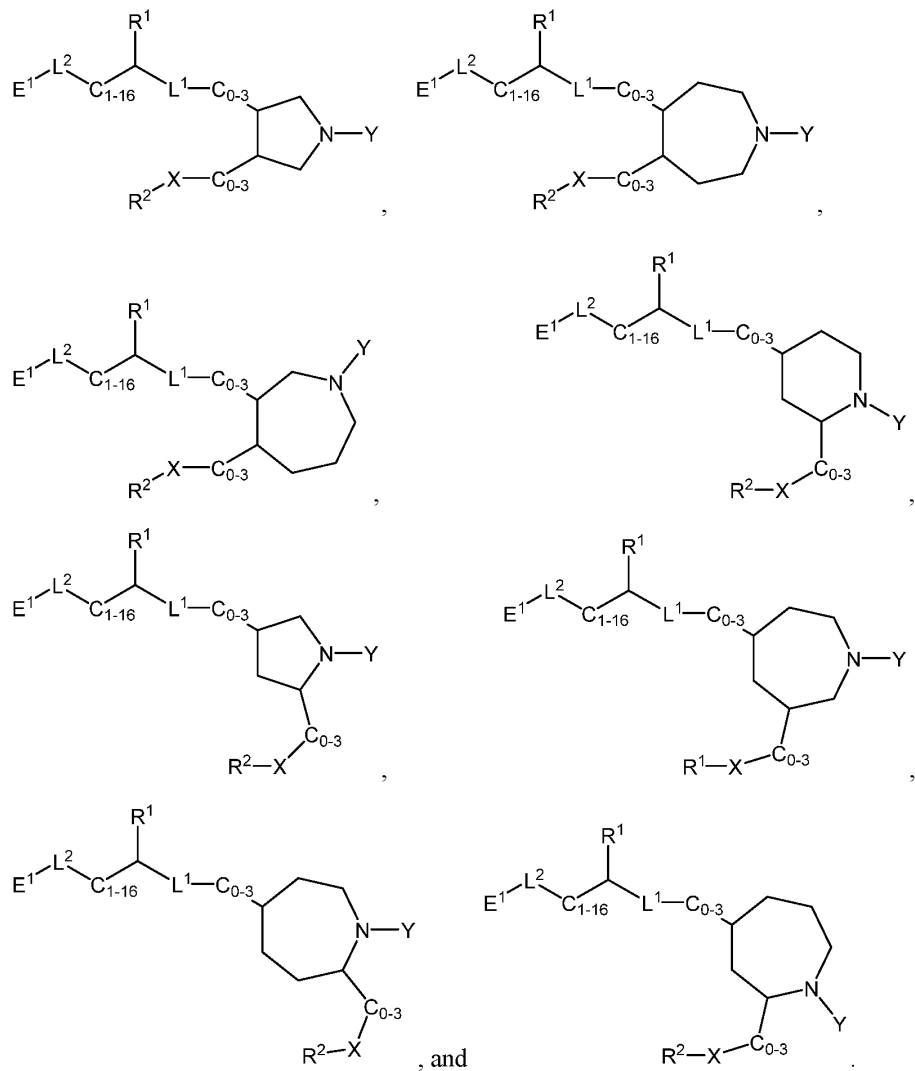
10. The compound of claim 1, wherein X is $-S-$, and L^1 is:



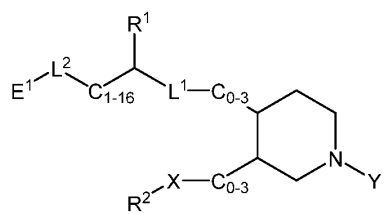
11. The compound of claim 1, wherein Y is a linear or branched $-C_{1-5}$ -alkyl.
12. The compound of claim 1, wherein Y is a linear or branched C_{1-5} -alkanol.
13. The compound of claim 1, wherein Y is a linear C_{1-5} -alkyl.
14. The compound of claim 1, wherein Y is a linear C_{1-5} -alkanol.
15. The compound of claim 1, wherein the compound is a compound of Formula I.
16. The compound of claim 1, wherein the compound of Formula I is selected from the group consisting of:



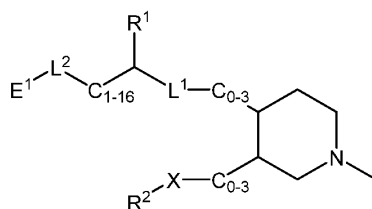
219



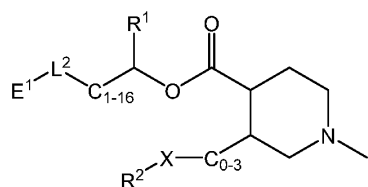
17. The compound of claim 1, wherein the compound of Formula I is:



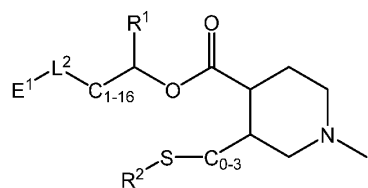
18. The compound of claim 1, wherein the compound of Formula I is:



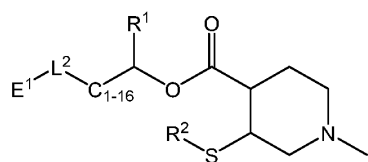
19. The compound of claim 1, wherein the compound of Formula I is:



20. The compound of claim 1, wherein the compound of Formula I is:



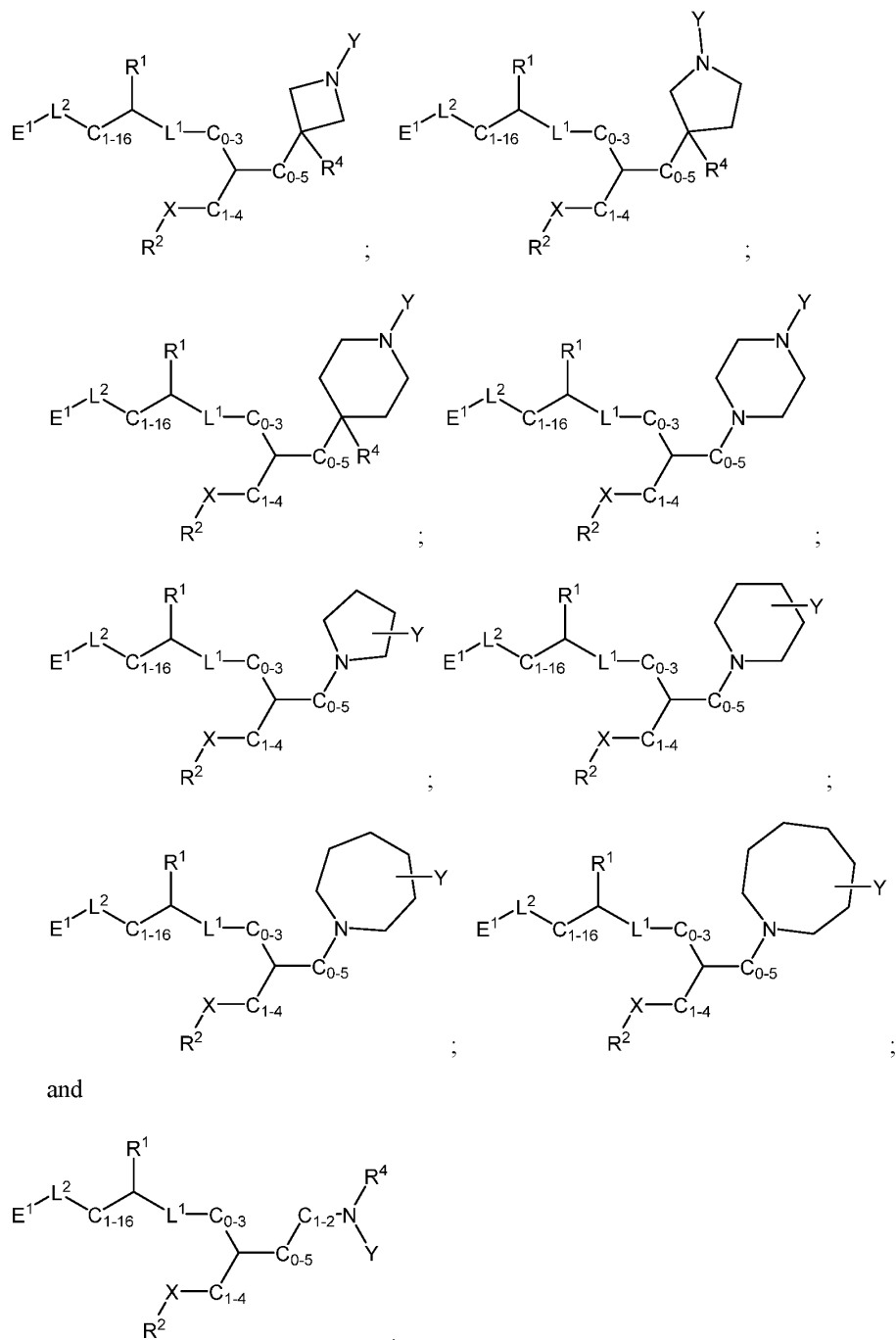
21. The compound of claim 1, wherein the compound of Formula I is:



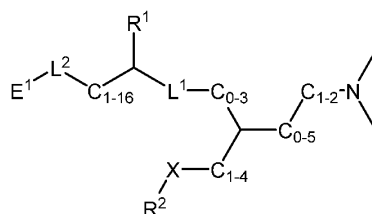
22. The compound of claim 1, wherein the compound is a compound of Formula II.

23. The compound of claim 1, wherein the compound of Formula II is selected from the group consisting of:

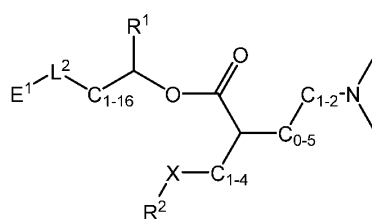
221



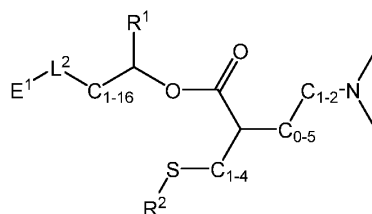
24. The compound of claim 1, wherein the compound of Formula II is:



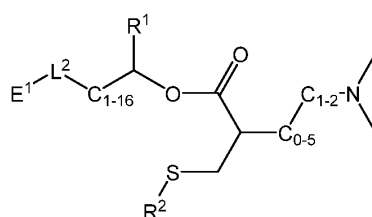
25. The compound of claim 1, wherein the compound of Formula II is:



26. The compound of claim 1, wherein the compound of Formula II is:

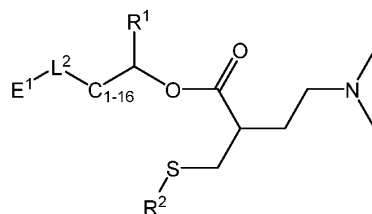


27. The compound of claim 1, wherein the compound of Formula II is:

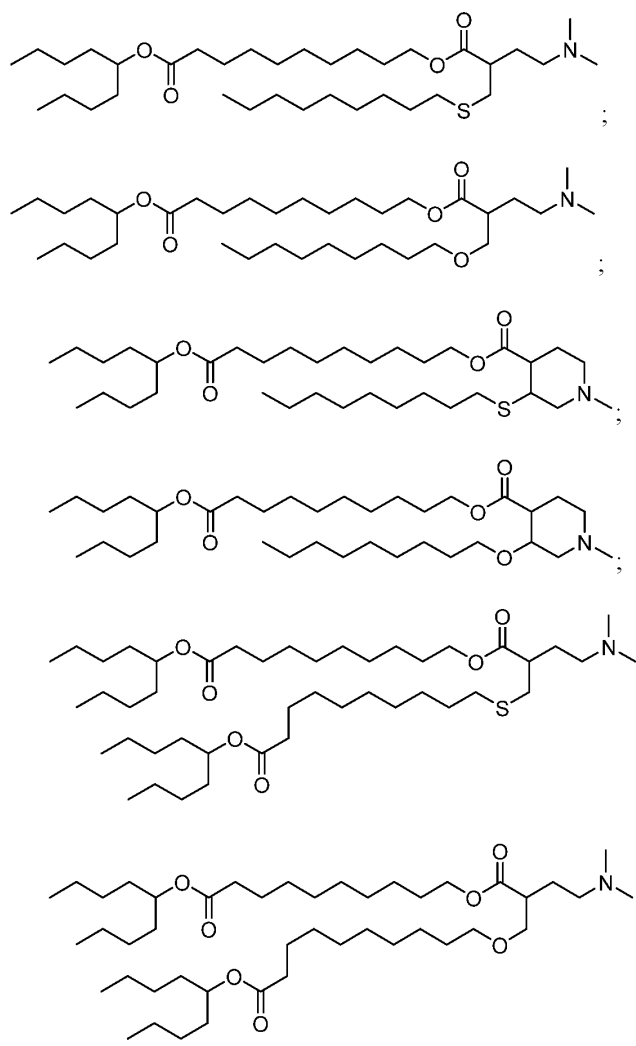


28. The compound of claim 1, wherein the compound of Formula II is:

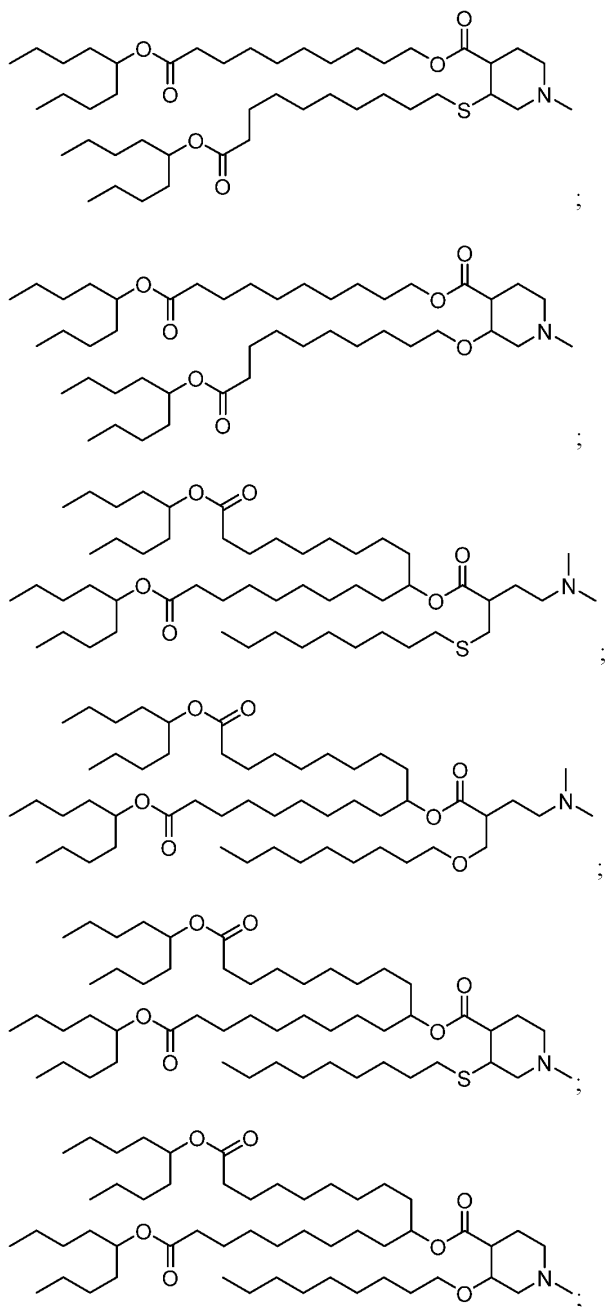
223



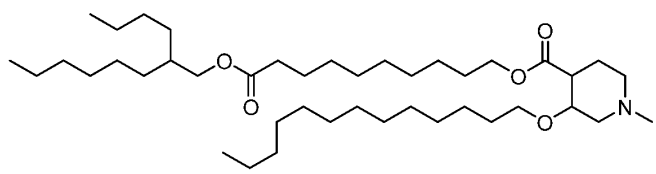
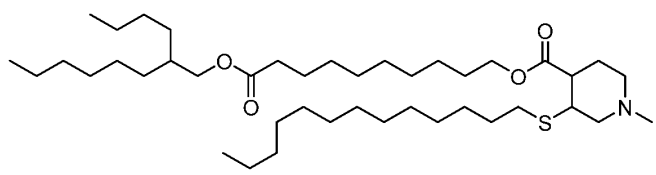
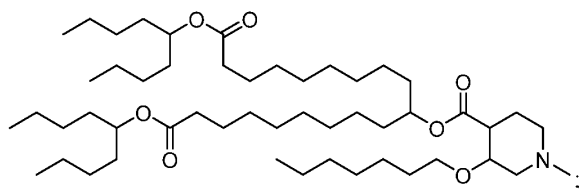
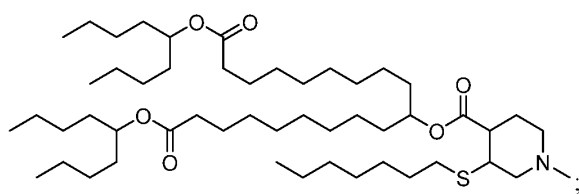
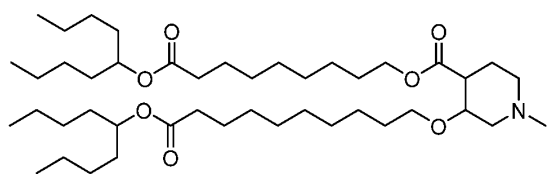
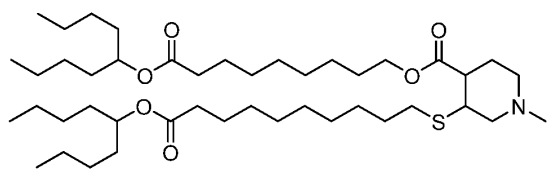
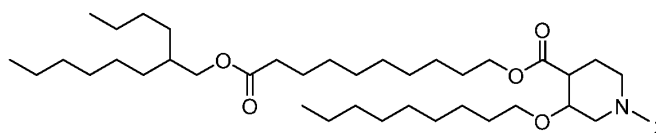
29. The compound of claim 1, wherein the compound is selected from the group consisting of:



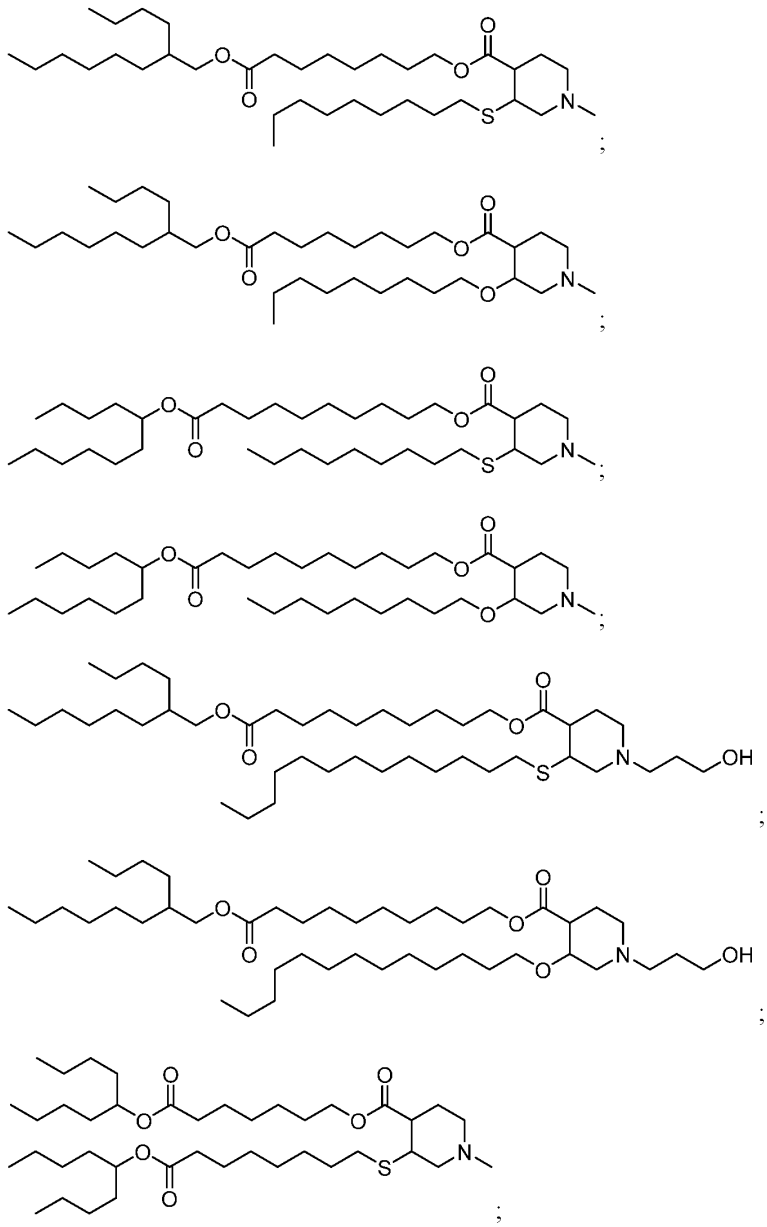
224



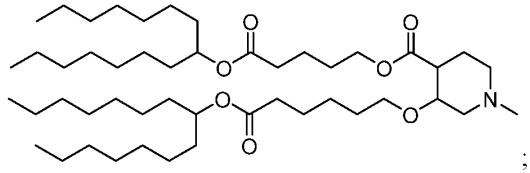
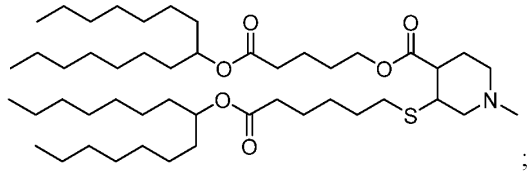
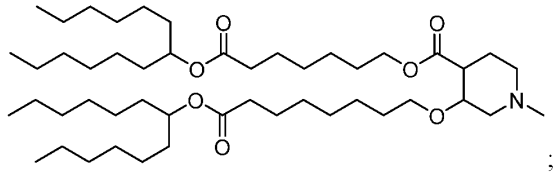
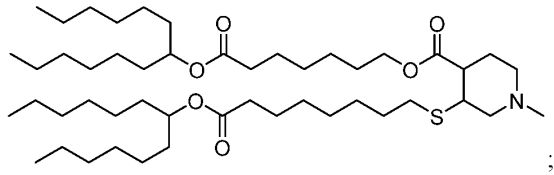
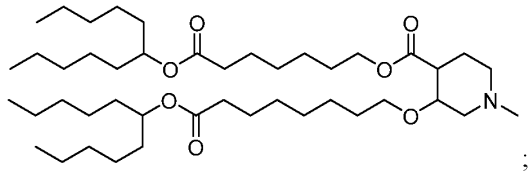
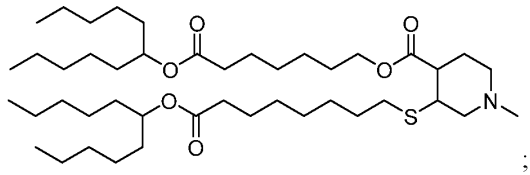
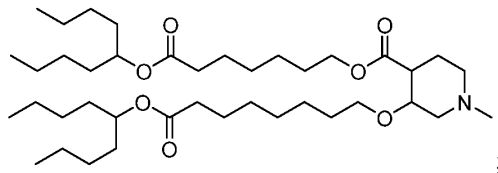
226

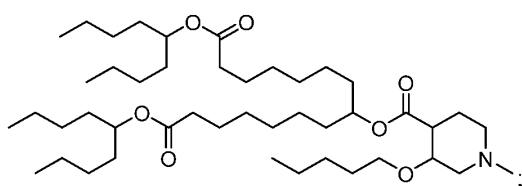
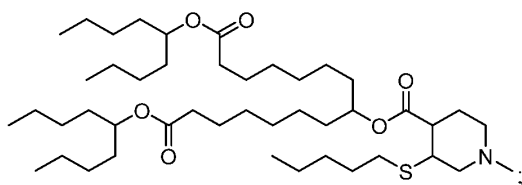
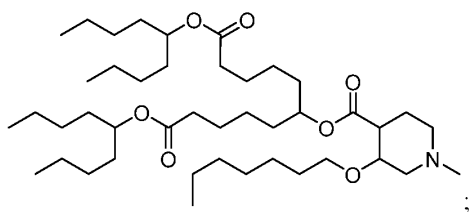
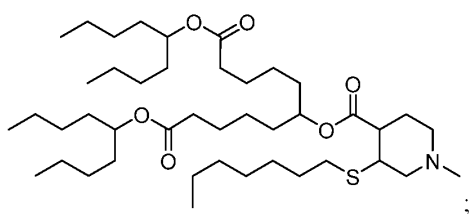
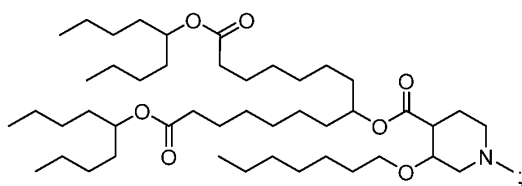
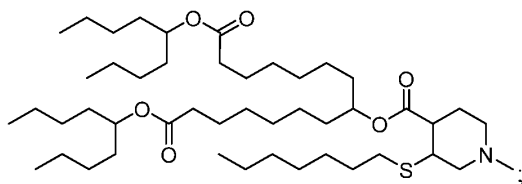


227

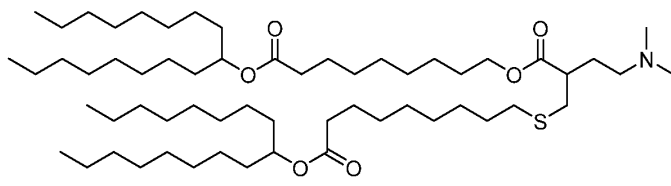
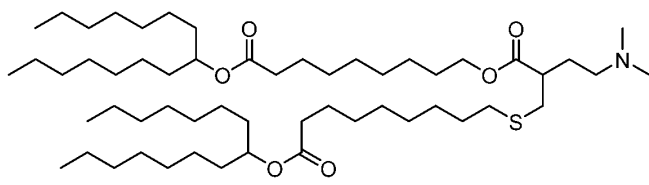
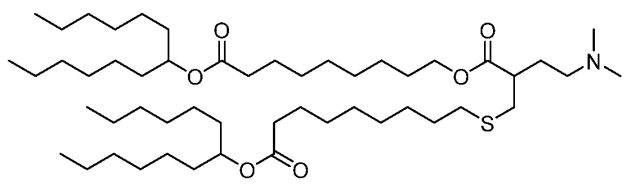
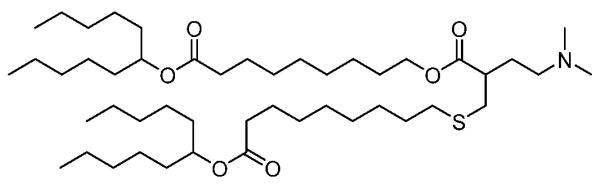
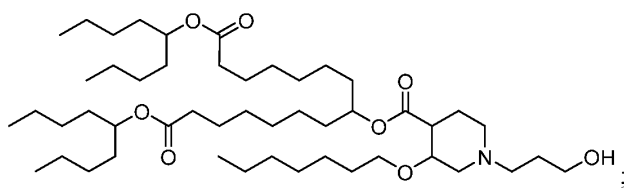
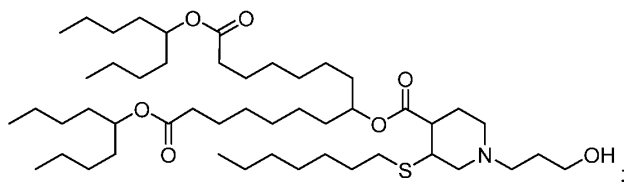


228

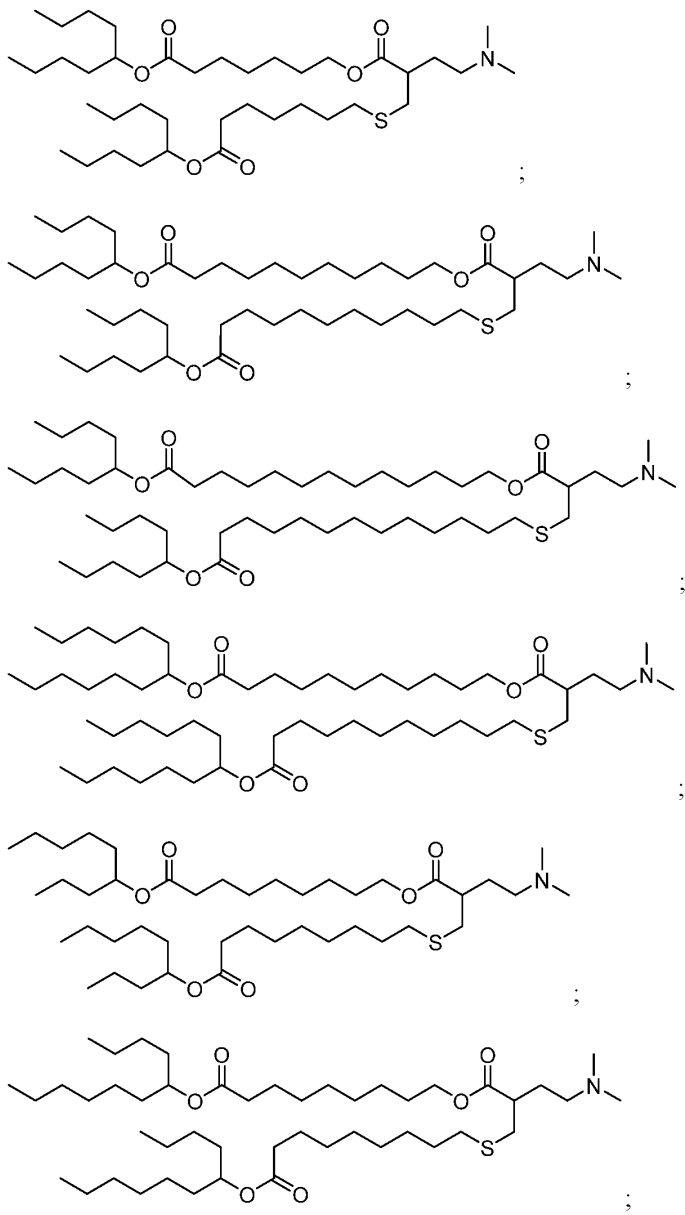




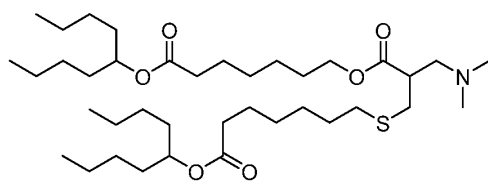
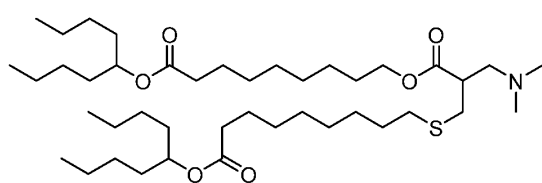
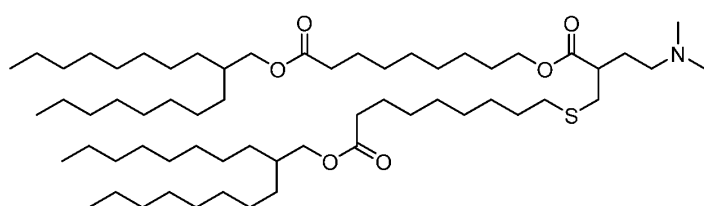
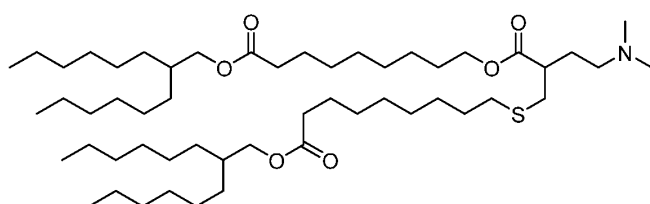
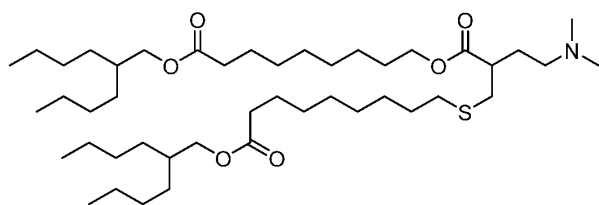
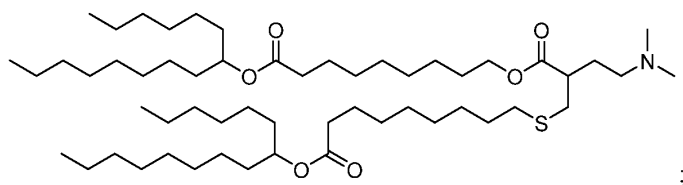
230



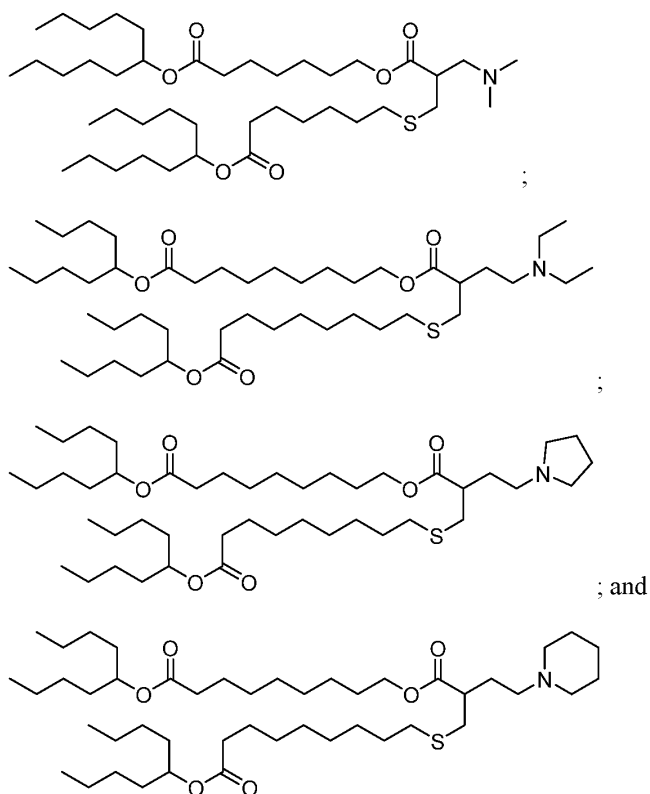
231



232



233



30. A lipid nanoparticle (LNP) comprising a lipid component comprising a compound of claim 1.

31. The lipid nanoparticle of claim 30, wherein the lipid component further comprises one or more of a neutral lipid, a structural lipid, and a PEGylated lipid.

32. The lipid nanoparticle of claim 31, wherein the neutral lipid is selected from the group consisting of 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-

glycero-3-phosphocholine (OchemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosaheptaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), and sphingomyelin.

33. The lipid nanoparticle of claim 31, wherein the structural lipid is selected from the group consisting of cholesterol, fecosterol, sitosterol, campesterol, stigmasterol, brassicasterol, ergosterol, tomatidine, tomatine, ursolic acid and alpha-tocopherol.

34. The lipid nanoparticle of claim 31, wherein the PEGylated lipid is selected from the group consisting of PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified dialkylamines, PEG-modified diacylglycerols, and PEG-modified dialkylglycerols, optionally PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, and PEG-DSPE.

35. The lipid nanoparticle of claim 31, wherein the lipid component comprises: about 25 mol % to about 60 mol % of a compound of claim 1; about 2 mol % to about 25 mol % neutral lipid; about 18.5 mol % to about 60 mol % structural lipid; and about 0.2 mol % to about 10 mol % of PEGylated lipid.

36. The lipid nanoparticle of claim 30, wherein the lipid nanoparticle further comprises a polynucleotide.

37. The lipid nanoparticle of claim 36, wherein the polynucleotide is selected from the group consisting of: a messenger RNA (mRNA), a self-amplifying mRNA (sa-mRNA), a small interfering RNA (siRNA), a microRNA (miRNA), miRNA inhibitors (antagomirs/antimirs), messenger-RNA-interfering complementary RNA (micRNA), short hairpin RNA (shRNA), multivalent RNA, dicer substrate RNA, an antisense oligonucleotide, plasmid DNA, DNA, and complementary DNA (cDNA).

38. The lipid nanoparticle of claim 36, wherein the polynucleotide is a ribonucleic acid (RNA).
39. The lipid nanoparticle of claim 36, wherein the polynucleotide is a conventional or self-amplifying mRNA.
40. The lipid nanoparticle of claim 30, wherein the lipid nanoparticle has a diameter of from about 30 nm to about 160 nm.
41. A pharmaceutical composition comprising a plurality of lipid nanoparticles of claim 30, and a pharmaceutically acceptable carrier.
42. A method of delivering a polynucleotide to a mammalian cell, including administering the lipid nanoparticle of claim 30, or the pharmaceutical composition of claim 41, to a subject to thereby contact the cell with the lipid nanoparticle or pharmaceutical composition and deliver the polynucleotide to the cell.
43. The method of claim 42 wherein the cell is a cell of a human subject.
44. A method of producing a polypeptide of interest in a mammalian cell, including the step of contacting the cell with a lipid nanoparticle of claim 30, or the pharmaceutical composition of claim 41, wherein the lipid nanoparticle comprises a conventional mRNA or a self-amplifying mRNA encoding the polypeptide.
45. A method of treating a disease, disorder or condition in a subject in need of such treatment, comprising administering a lipid nanoparticle of claim 30, or the pharmaceutical composition of claim 41, to the subject to thereby treat the disease, disorder or condition.
46. The method of claim 45, wherein the disease, disorder or condition is selected from the group consisting of a rare disease, an infectious disease, cancer, a proliferative disease, a genetic disease, an autoimmune disease, diabetes, a neurodegenerative disease, a cardiovascular disease, a reno-vascular disease and a metabolic disease.
47. A vaccine comprising a lipid nanoparticle of claim 30, or the pharmaceutical composition of claim 40, and a mRNA encoding a polypeptide.
48. The vaccine of claim 47, wherein the vaccine is selected from a tumor vaccine, an influenza vaccine, and a SARS, including a SARS-CoV-2, vaccine.