



(51) International Patent Classification:

A61K 9/127 (2006.01) A61K 48/00 (2006.01)

A61K 9/51 (2006.01) A61P 3/00 (2006.01)

A61K 31/7105 (2006.01) C12N 9/88 (2006.01)

A61K 38/51 (2006.01) C12N 15/88 (2006.01)

(21) International Application Number:

PCT/US2023/017573

(22) International Filing Date:

05 April 2023 (05.04.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/328,209 06 April 2022 (06.04.2022) US

63/417,229 18 October 2022 (18.10.2022) US

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

(72) Inventors: RICE, Lisa; c/o ModernaTX, Inc., 200 Technology Square, Cambridge, Massachusetts 02139 (US).

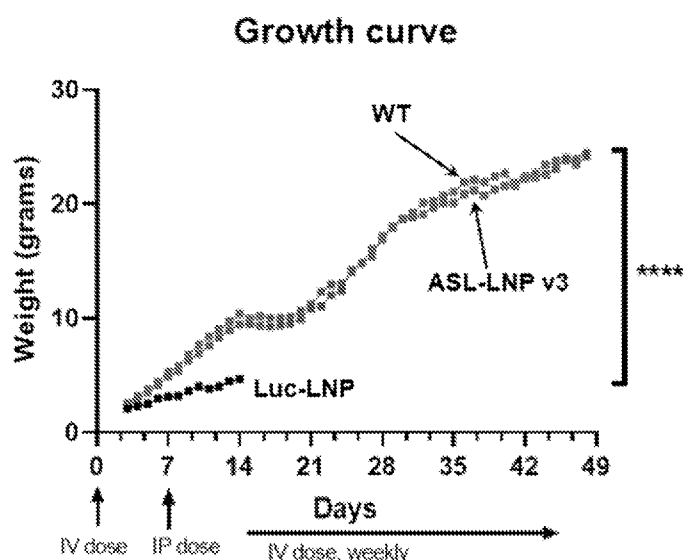
FRASSETTO, Andrea; c/o ModernaTX, Inc., 200 Technology Square, Cambridge, Massachusetts 02139 (US).
DOUSIS, Athanasios; 2 Carver Street, Unit 1, Somerville, Massachusetts 02143 (US).

(74) Agent: TALAPATRA, Sunit et al.; Foley & Lardner LLP, 3000 K. Street, N.W., Suite 600, Washington, District of Columbia 20007 (US).

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

(54) Title: LIPID NANOPARTICLES AND POLYNUCLEOTIDES ENCODING ARGININOSUCCINATE LYASE FOR THE TREATMENT OF ARGININOSUCCINIC ACIDURIA

FIG. 7A



(57) Abstract: This disclosure relates to mRNA therapy for the treatment of argininosuccinici aciduria. mRNAs for use in the present technology, when administered in vivo, encode argininosuccinic acid lyase (AST). mRNA therapies of the disclosure increase and/or restore deficient levels of AST expression and/or activity in subjects. mRNA therapies of the disclosure further decrease levels of toxic ammonia, ASA, or citrulline associated with deficient AST activity in subjects.



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

Published:

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

LIPID NANOPARTICLES AND POLYNUCLEOTIDES ENCODING ARGININOSUCCINATE LYASE FOR THE TREATMENT OF ARGININOSUCCINIC ACIDURIA

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Patent Application No. 63/328209, filed April 6, 2022, and U.S. Provisional Patent Application No. 63/417229, filed October 18, 2022, the entire contents of each of which is incorporated herein by reference.

BACKGROUND

Argininosuccinic aciduria (argininosuccinate lyase deficiency) is a rare autosomal recessive condition caused by mutations in the gene encoding argininosuccinate lyase (ASL). ASL converts argininosuccinic acid (ASA) into arginine and is integral to the liver-based urea cycle, which detoxifies neurotoxic ammonia. Ammonia, which is formed when proteins are broken down in the digestive process, is toxic if blood ammonia levels become too high. Thus, ASL defects can cause accumulation of ammonia in the blood (hyperammonaemia).

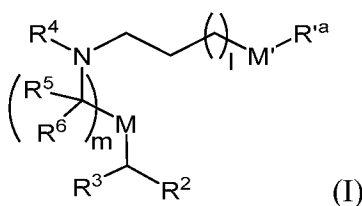
Argininosuccinic aciduria symptoms usually present themselves within the first few days of life and can include lethargy or unwillingness to eat, poorly controlled breathing rate or body temperature, and neurological symptoms including seizures, unusual body movements, or coma. Argininosuccinic aciduria may cause developmental delay and intellectual disability, chronic liver disease, neurodisability, high blood pressure (hypertension), skin lesions, and brittle hair. Milder forms of ASL deficiency may manifest in accumulation of plasma ammonia during periods of illness or stress, mild intellectual disability, or learning disabilities. Argininosuccinic aciduria is estimated to have a prevalence of about 1:70,000 to 1:218,000 newborns.

Argininosuccinic aciduria patients exhibit elevated levels of plasma ammonia, elevated plasma argininosuccinic acid, elevated plasma citrulline, and elevated urinary orotic acid. Standard of care aims to normalize ammoniaemia with protein-restricted diet, ammonia-scavenger drugs and, in severe cases, liver transplantation. However limited efficacy of medical therapy and poor quality of life indicate the need for novel therapies.

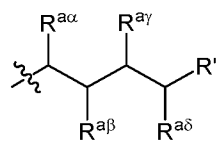
In view of the significant problems associated with existing argininosuccinic aciduria treatments, there is an unmet need for improved treatment for argininosuccinic aciduria.

SUMMARY

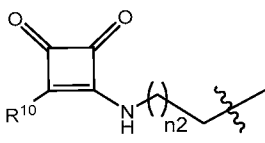
In one aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a lipid nanoparticle comprising: (a) an ionizable lipid of Formula (I):



or its N-oxide, or a salt or isomer thereof, wherein R'^a is R'^{branched}, wherein R'^{branched} is:

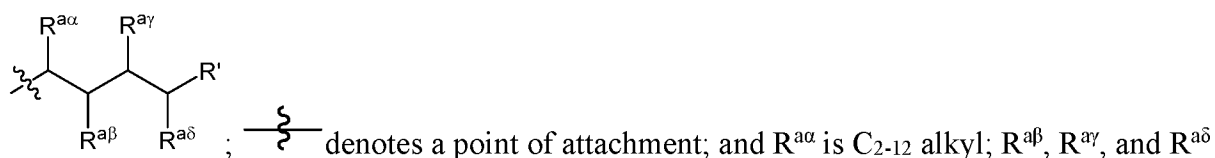


; wherein denotes a point of attachment; wherein R'^{alpha}, R'^{beta}, R'^{gamma}, and R'^{delta} are each independently selected from the group consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl; R² and R³ are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl; R⁴ is selected from the group consisting of -(CH₂)_nOH, wherein n is selected from the group



consisting of 1, 2, 3, 4, and 5, and , wherein denotes a point of attachment; wherein R¹⁰ is N(R)₂; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; wherein n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10; each R⁵ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H; each R⁶ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H; M and M' are each independently selected from the group consisting of -C(O)O- and -OC(O)-; R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl; l is selected from the group consisting of 1, 2, 3, 4, and 5; and m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13; and (b) an mRNA molecule comprising an open reading frame (ORF) encoding an ASL polypeptide having an amino acid sequence identical to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5.

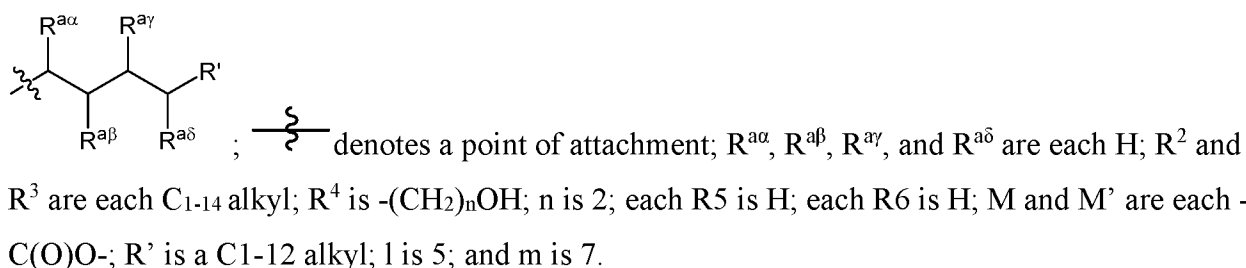
In some embodiments, in the compound of Formula (I): $R^{a\alpha}$ is $R^{\text{'branched}}$, $R^{\text{'branched}}$ is



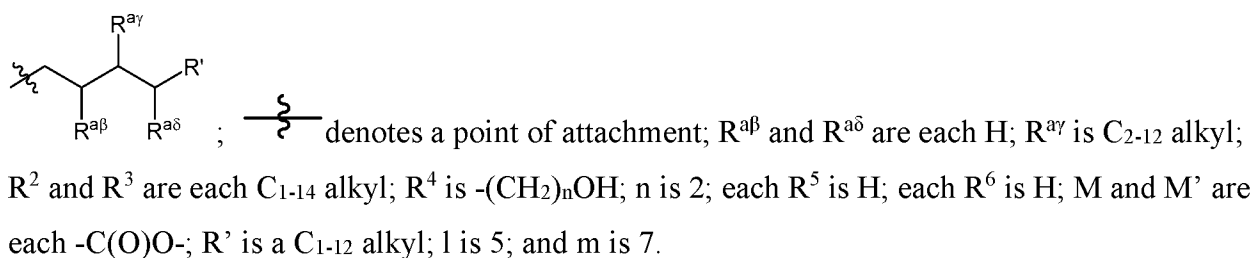
are each H; R^2 and R^3 are each C₁₋₁₄ alkyl; R^4 is ; R^{10} is NH(C₁₋₆ alkyl); n2 is 2; each R^5 is H; each R^6 is H; M and M' are each -C(O)O-; R' is a C₁₋₁₂ alkyl; l is 5; and m is

5 7.

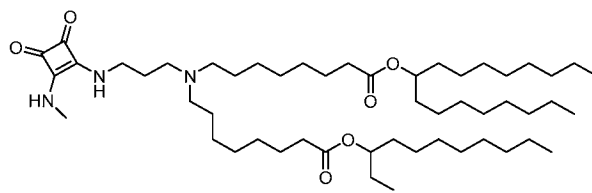
In some embodiments, in the compound of Formula (I): R^a is $R^{\text{'branched}}$; $R^{\text{'branched}}$ is



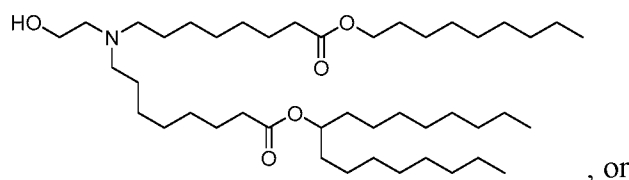
10 In some embodiments, in the compound of Formula (I): R^a is $R^{\text{'branched}}$; $R^{\text{'branched}}$ is

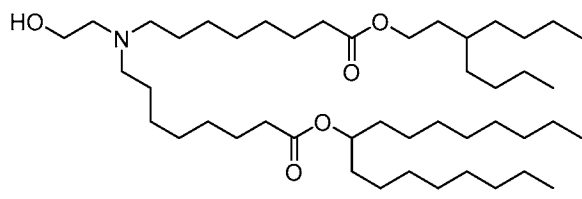


In some embodiments, the ionizable lipid is:



15





or N-oxides, salts, or isomers thereof.

In some embodiments, the lipid nanoparticle further comprises: a phospholipid; a structural lipid; and a PEG-lipid. In some embodiments, the lipid nanoparticle comprises: 40-50 mol% of the ionizable lipid, 30-45 mol% of the structural lipid, 5-15 mol% of the phospholipid, and 1-5 mol% of the PEG-lipid. In some embodiments, the lipid nanoparticle comprises: 45-50 mol.% of the ionizable amino lipid, 10-15 mol.% of the phospholipid, 35-40 mol.% of the structural lipid, and 1-3 mol.% of the PEG-modified lipid.

In some embodiments, the phospholipid comprises comprises 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-3-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-3-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-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), sphingomyelin, and mixtures thereof. In some embodiments, the phospholipid comprises DSPC.

In some embodiments, the structural lipid is selected from the group consisting of: cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, tomatine, ursolic acid, alpha-tocopherol, hopanoids, phytosterols, steroids, derivatives thereof, and mixtures thereof. In some embodiments, the structural lipid comprises cholesterol or a derivative thereof.

In some embodiments, the PEG-lipid is selected from: 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol (PEG-DMG), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)] (PEG-DSPE), PEG-disteryl glycerol (PEG-DSG), PEG-dipalmitoleyl, PEG-dioleyl, PEG-distearyl, PEG-diacylglycamide (PEG-DAG), PEG-
5 dipalmitoyl phosphatidylethanolamine (PEG-DPPE), PEG-1,2-dimyristyloxylpropyl-3-amine (PEG-c-DMA), or combinations thereof. In some embodiments, the PEG-lipid comprises PEG-DMG.

In some embodiments, mRNA molecule, when administered as a single dose to a human subject, is sufficient to: (i) reduce plasma ammonia levels in the human subject to a level of less
10 than or equal to about 500 μ M, less than or equal to about 400 μ M, less than or equal to about 300 μ M, less than or equal to about 200 μ M, less than or equal to about 100 μ M, or less than or equal to about 50 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or (ii) reduce plasma
15 citrulline levels in the human subject to a less than or equal to about 100 μ M, less than or equal to about 50 μ M, less than or equal to about 25 μ M, or less than or equal to about 10 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or (iii) reduce plasma argininosuccinic acid (ASA)
20 levels to less than or equal to about 150 μ M, less than or equal to about 100 μ M, less than or equal to about 50 μ M, or less than or equal to about 25 μ M, or less than or equal to about 10 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration.

In some embodiments, the mRNA molecule, when administered as a single dose to a human subject, is sufficient to: (iv) increase cellular ASL levels in the human subject by at least
25 about 1.1-fold, at least about 1.2-fold, at least about 1.5-fold, or at least about 2.0-fold, compared to the human subject's baseline cellular ASL levels for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120 hours, or at least one week post-administration; or (v) increase ASL activity in the human subject by at least about 1.5-fold, at least about 2.0-fold, or at least about 2.5-fold, compared to the human subject's baseline ASL
30 activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or (vi) reduce plasma orotate levels to less than or equal to about 150 μ mol/mmol creatinine, less than or equal to about 100 μ mol/mmol creatinine, less than or equal to about 50 μ mol/mmol creatinine, less than or equal to

about 25 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 10 $\mu\text{mol}/\text{mmol}$ creatinine, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration. In some embodiments, the mRNA molecule, when administered as a single dose to a human subject, is sufficient to: (i) increase liver glutathione levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; (ii) reduce liver xCT antiporter levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; (iii) reduce liver xCT antiporter activity in the subject in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline ASL activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; reduce total plasma homocysteine (HCyS) levels in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline total plasma HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or reduce total liver homocysteine (HCyS) levels in the subject at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline total

liver HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration.

In some embodiments, ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:1. In some embodiments, the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:2. In some embodiments, the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:3. In some embodiments, the ASL polypeptide encoded by mRNA molecule has an amino acid sequence identical to SEQ ID NO:4. In some embodiments, the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:5.

In some embodiments, the ORF is at least 65% identical to SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, or SEQ ID NO:33. In some embodiments, the ORF is at least 65% identical to SEQ ID NO:20.

In some embodiments, the mRNA molecule further comprises a 5'UTR having a nucleic acid sequence according to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the mRNA molecule further comprises a 3'UTR having a nucleic acid sequence according to SEQ ID NO:108, SEQ ID NO:111, SEQ ID NO:128, SEQ ID NO:137, SEQ ID NO:138, or SEQ ID NO:139.

In some embodiments, the mRNA molecule comprises a poly-A region, wherein the poly-A region about 100 nucleotides in length. In some embodiments, the mRNA molecule comprises at least one chemically modified nucleobase, sugar, backbone, or any combination thereof. In some embodiments, the at least one chemically modified nucleobase is selected from the group consisting of pseudouracil (ψ), N1-methylpseudouracil (m1 ψ), 1-ethylpseudouracil, 2-thiouracil (s2U), 4'-thiouracil, 5-methylcytosine, 5-methyluracil, 5-methoxyuracil, and any combination thereof.

In some embodiments, the mRNA molecule further comprises a 5' terminal cap. In some embodiments, the 5' terminal cap comprises a m7G-ppp-Gm-AG, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azidoguanosine, Cap2, Cap4, 5' methylG cap, or an analog thereof. In some embodiments, the 5' terminal cap comprises m7G-ppp-Gm-AG or Cap1.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a pharmaceutical composition comprising: the lipid nanoparticle according to any of the above embodiments; and a pharmaceutically acceptable carrier. In some embodiments, the pharmaceutical composition is for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, reducing plasma orotate levels, or reducing ureagenesis in a subject.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a pharmaceutical composition comprising: the lipid nanoparticle according to any of the above embodiments; and a pharmaceutically acceptable carrier. In some embodiments, the pharmaceutical composition is for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, increasing glutathione levels, increasing liver glutathione levels, reducing plasma orotate levels, reducing ureagenesis, reducing liver xCT antiporter levels, reducing liver xCT antiporter activity, reducing total plasma HcyS levels, or reducing total liver HcyS levels in a subject.

In another aspect, which may be combined with any other aspect or embodiment, the lipid nanoparticle, ASL polypeptide, or mRNA molecule according to any of the above embodiments is for use in the manufacture of a medicament or pharmaceutical composition for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, reducing plasma orotate levels, or reducing ureagenesis in a subject.

In another aspect, which may be combined with any other aspect or embodiment, the lipid nanoparticle, ASL polypeptide, or mRNA molecule according to any of the above embodiments is for use in the manufacture of a medicament or pharmaceutical composition for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, increasing glutathione levels, increasing liver glutathione levels, reducing plasma orotate levels, reducing ureagenesis,

reducing liver xCT antiporter levels, reducing liver xCT antiporter activity, reducing total plasma HCyS levels, or reducing total liver HCyS levels in a subject.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of treating or delaying the onset and/or progression of argininosuccinic aciduria in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing plasma ammonia levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing plasma argininosuccinic acid (ASA) levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing plasma citrulline levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of increasing ASL levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of increasing ASL activity in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing orotate levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing ureagenesis in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of

increasing glutathione levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of increasing liver glutathione levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing liver xCT antiporter levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing liver xCT antiporter activity in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing total plasma homocysteine (HCyS) levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a method of reducing total liver homocysteine (HCyS) levels in a human subject, comprising: administering to the human subject an effective amount of a pharmaceutical composition according to any of the above embodiments.

In some embodiments, the administration is a single administration. In some embodiments, the administration is a repeated administration. In some embodiments, the administration is about once per day, about once per week, about once per two weeks, or about once per month. In some embodiments, the pharmaceutical composition is administered
5 intravenously, intramuscularly, or subcutaneously.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to an ASL polypeptide having an amino acid sequence identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11. In some embodiments, the amino
10 acid sequence is identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5. In some embodiments, the amino acid sequence is identical to SEQ ID NO:2.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a pharmaceutical composition comprising: the ASL polypeptide according to any of the above embodiments; and a pharmaceutically acceptable carrier. In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a cell comprising the ASL polypeptide according to any of the above
5 embodiments.

In another aspect, which may be combined with any other aspect or embodiment, the present disclosure relates to a codon-optimized mRNA molecule encoding the ASL polypeptide according to any of the above embodiments. In some embodiments, the codon-optimized mRNA
10 molecule has a nucleic acid sequence having at least about 60% identity to the nucleic acid sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29. In some embodiments, the nucleic acid sequence is at least about 60% identical to the nucleic acid sequence of SEQ ID NO:20. In another aspect, which may be combined with any other aspect or
15 embodiment, the present disclosure relates to a cell comprising the mRNA molecule according to any of the above embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A is a plot showing expression of hASL protein variants encoded by mRNA constructs including different ORFs (expressed as fold change from eGFP) from in-cell
20 western at 96 hours. The construct including SEQ ID NO:306, which has an ORF according to SEQ ID NO:20, was trending to have the highest expression among the ORFs encoding hASL protein variants.

FIG. 1B is a graph showing the expression of hASL protein with different UTR variants (normalized to beta actin) detected at 24 hours after transfection. The construct
25 according to SEQ ID NO:300 showed the highest ASL expression among all UTR variants.

FIG. 2 shows the fumarate production in HEK293 cells transfected with eGFP (control) or ASL mRNA at 0-, 15-, 30-, and 60-minute time points. ASL mRNA transfected cells produced more fumarate than control cells. AS = 2mM Argininosuccinic Acid.

FIG. 3 shows a survival (Kaplan Meier) plot of the hypomorph mice administered
30 with hASL mRNA.

FIG. 4A shows the basal ASL levels in control (healthy), Patient 1, and Patient 2 fibroblasts measured by in-cell western. Patient samples had about 50% of ASL protein levels as observed in healthy controls.

FIG. 4B shows ASL levels after fibroblasts were administered luciferase mRNA in lipid nanoparticles (Luc-LNP), ASL mRNA (SEQ ID NO:334) in lipid nanoparticles (LNP-1A), naked ASL mRNA or PBS (vehicle) as measured by in-cell western. ASL mRNA administration increased ASL levels in all fibroblasts. (For each subject, the bars are (from left to right): Luc-LNP; ASL mRNA in LNP; naked ALS mRNA; PBS.)

FIG. 4C shows ASL activity after fibroblasts were administered luciferase mRNA in lipid nanoparticles (Luc-LNP), ASL mRNA (SEQ ID NO:334) in lipid nanoparticles (LNP-1A), naked ASL mRNA or PBS (vehicle) as measured by fumarate detection assay. ASL mRNA administration increased ASL activity levels in all fibroblasts. (For each subject, the bars are (from left to right): Luc-LNP; ASL mRNA in LNP; naked ALS mRNA.)

FIG. 5A shows growth curves for ASL-LNP v2 (SEQ ID NO:300 in LNP-3A), Luc-LNP administered mild phenotype (late death) ASA mice compared to their wild type littermates. ASL-LNP v2 (SEQ ID NO:300 in LNP-3A) significantly improved growth of the ASA mice. (n=3 for ASL-LNP, n=9 for Luc-LNP and n=4 for WT).

FIG. 5B shows a survival plot for ASL-LNP v2 (SEQ ID NO:300 in LNP-3A) and Luc-LNP administered mild phenotype (late death) ASA mice. ASL-LNP v2 (SEQ ID NO:300 in LNP-3A) administration significantly improved survival of the mild phenotype ASA mice.

FIG. 6A shows liver ASL protein levels as measured by western blot analysis (normalized to wild type ASL levels). Liver ASL levels were restored to near-normal levels 24 hours after ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) injection, however, liver ASL protein levels dropped to not significant (ns) levels as compared to Luc-LNP-treated mice livers at 72-hour and 7-day timepoints.

FIG. 6B shows the percentage of ASL-positive regions in liver histology sections. Histological analyses were performed as described in Example 1. ASL-positive regions in liver histology sections were significantly higher than control (Luc-LNP) 24 hours after ASL-

LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) injection, however, the difference was not significant (ns) at 72-hour and 7-day timepoints.

FIG. 6C shows ASL activity levels in livers. The ASL activity was restored to near-normal levels 24 hours after ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) injection, but was not significant at 72-hour and 7-day timepoints.

FIGS. 6D-6F show plasma ammonia, Argininosuccinic acid and citrulline levels, respectively. Beginning from the 24 hour timepoint and continuing through the 72-hour and 7-day timepoints, **(D)** plasma ammonia, **(E)** Argininosuccinic acid and **(F)** citrulline levels were significantly lower in ASL-LNP treated mice, as compared to Luc-LNP treated mice.

FIG. 6G shows plasma orotate (orotic acid) levels measured 24 hours post ASL-LNP or Luc-LNP administration. ASL-LNP-treated mice showed significantly low amounts of orotate (similar to WT levels) as compared to Luc-LNP-treated mice.

FIG. 7A shows growth curves for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A), Luc-LNP administered severe phenotype (early death) ASA mice compared to their wild type littermates. Administration of ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) completely rescued the growth of the severe phenotype mice, as compared to Luc-LNP control.

FIG. 7B shows survival plot for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A), Luc-LNP administered severe phenotype (early death) ASA mice. ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) administration significantly improved the survival of the severe phenotype ASA mice, all ASL-treated mice surviving up to 50 days.

FIG. 7C shows liver ASL protein levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Compared to Luc-LNP treated mice, the liver ASL protein levels were increased significantly to normal levels in ASL-LNP-treated mice.

FIG. 7D shows liver ASL activity levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Compared to Luc-LNP treated mice, the liver ASL activity levels were increased significantly to normal levels in ASL-LNP-treated mice.

FIG. 7E shows plasma ammonia levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Plasma ammonia levels were decreased to normal (WT) levels after ASL-LNP treatment.

FIG. 7F shows plasma ASA levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Plasma ASA levels were decreased to normal (WT) levels after ASL-LNP treatment.

FIG. 7G shows plasma citrulline levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Plasma citrulline levels were decreased to normal (WT) levels after ASL-LNP treatment.

FIG. 7H shows plasma ALT levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. Plasma ALT levels were increased to normal (WT) levels after ASL-LNP treatment.

FIG. 7I shows a longitudinal analysis of plasma ASA levels over 50 days. In ASL-LNP-treated mice, average plasma ASA levels were consistently low (as low as wild type levels, when there was a wild type measurement) throughout the observation period.

FIG. 7J shows a longitudinal analysis of plasma citrulline levels over 50 days. In ASL-LNP-treated mice, average plasma citrulline levels were consistently low (as low as wild type levels, when there was a wild type measurement) throughout the observation period.

FIG. 7K and FIG. 7L show average orotate levels in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. The average orotate levels were decreased to non-detectable levels, similar to wild-type mice.

FIG. 7M shows C¹³ ureagenesis assay results in WT, ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) and Luc-LNP-treated mice. C¹³ ureagenesis was restored completely to normal levels upon ASL-LNP treatment.

FIG. 8A shows growth curves for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) or Luc-LNP (in LNP-3A) administered ASA mice compared to their wild type littermates before and after weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight), with treatment beginning on day 21. ASL-LNP v3 (SEQ ID NO:300 (+ idT) in

LNP-3A) significantly improved growth of the ASA mice. (n=6 for ASL-LNP, n=6 for Luc-LNP and n=6 for WT).

FIG. 8B shows a survival plot for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A), Luc-LNP (in LNP-3A) administered ASA mice, with weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight) beginning on day 21. ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) administration significantly improved the survival of ASA mice, with all but one subject surviving until the study end point.

FIG. 9A shows growth of ASA mice at day 57 of ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) treatment, upon weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight) beginning on day 21.

FIG. 9B shows growth of wildtype mice at day 57.

FIG. 9C compares growth at day 27 of wildtype (left) and Luc-LNP (in LNP-3A) administered ASA mice (two right), along with one Luc-LNP administered ASA mouse at day 31.

FIG. 10A compares growth velocity curves for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) versus Luc-LNP (in LNP-3A) administered ASA mice after the first weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight), with treatment beginning on day 21.

FIG. 10B compares growth velocity curves for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) versus Luc-LNP (in LNP-3A) administered ASA mice after the second weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight), with treatment beginning on day 21.

FIG. 10C shows a growth velocity curve for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) administered ASA mice after the third weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight), with treatment beginning on day 21.

FIG. 10D shows a growth velocity curve for ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) administered ASA mice after the fourth weekly IV administration of ASL-LNP therapy (1 mg/kg bodyweight), with treatment beginning on day 21.

FIG. 11A shows results from preliminary kinetics studies of ASL levels (ASL/GAPDH) in ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) versus Luc-LNP (in LNP-3A) administered ASA neonatal mice after single intracerebroventricular (ICV) administration in the brain after 24 hours, 7 days, 2 weeks, and 3 weeks post-injection.

5 **FIG. 11B** shows results from preliminary kinetics studies of ASL activity (fold-difference in ASL activity) in ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) versus Luc-LNP (in LNP-3A) administered ASA neonatal mice after single ICV administration in the brain after 24 hours, 7 days, 2 weeks, and 3 weeks post-injection.

10 **FIG. 12** shows an image of a Western blot for protein expression of xCT in liver tissue from 2 week-old WT and *Asl^{Neo/Neo}* mice. Three biological replicates are shown. Actin was used as a loading control.

15 **FIG. 13A** shows representative PET/CT images of [¹⁸F]FSPG distribution (%ID/g) in individual WT and *Asl^{Neo/Neo}* mice intravenously administered [¹⁸F]FSPG to visualize uptake/retention of [¹⁸F]FSPG and xCT activity. Coronal plane images of the mice are shown, and axial plane images of the livers are shown.

FIG. 13B shows a quantification of the [¹⁸F]FSPG retention (%ID/g) in the liver (left) and skin (right) of WT and *Asl^{Neo/Neo}* mice intravenously administered [¹⁸F]FSPG.

20 **FIG. 14** shows a principal component analysis plot comparing treatment applied (untreated WT, *hASL* or *Luc* mRNA) and mouse genotype (WT or *Asl^{Neo/Neo}*) with percentage of variance associated with each axis.

25 **FIG. 15A** shows a volcano plot showing differential gene expression (DEG) analysis of *Luc* mRNA vs WT. Scatter plots show log transformed adjusted p-values (<0.05) on the y-axis against log2 fold change (>0.10) values on the x-axis. Blue (left) and red (right) dots represent genes that are significantly downregulated and upregulated, respectively, between groups. Grey dots represent genes that are not significantly altered.

FIG. 15B shows a volcano plot showing differential gene expression (DEG) analysis of *hASL* mRNA vs WT. Scatter plots show log transformed adjusted p-values (<0.05) on the y-axis against log2 fold change (>0.10) values on the x-axis. Blue (left) and red (right) dots

represent genes that are significantly downregulated and upregulated, respectively, between groups. Grey dots represent genes that are not significantly altered.

FIG. 15C shows a volcano plot showing differential gene expression (DEG) analysis of *hASL* mRNA vs *Luc* mRNA. Scatter plots show log transformed adjusted p-values (<0.05) on the y-axis against log2 fold change (>0.10) values on the x-axis. Blue (left) and red (right) dots represent genes that are significantly downregulated and upregulated, respectively, between groups. Grey dots represent genes that are not significantly altered.

FIG. 16 shows a pathway analysis highlighting genes of interest significantly altered in DEG analysis organised by their associated pathways when comparing the *Luc* mRNA vs WT and *hASL* mRNA vs *Luc* mRNA groups.

FIG. 17 shows a graph of total glutathione (GSH) levels from liver in WT, *Luc* mRNA-treated *Asl^{Neo/Neo}* mice, and *hASL* mRNA-treated *Asl^{Neo/Neo}* neonatal and adult mice. One-way ANOVA with Tukey's post-test, ns=not significant, *** $p<0.001$, **** $p<0.0001$, $n=4$.

FIG. 18A shows representative PET/CT images of [^{18}F]FSPG distribution (%ID/g) in individual WT mice, untreated *Asl^{Neo/Neo}* mice, and *hASL* mRNA-treated *Asl^{Neo/Neo}* mice intravenously administered [^{18}F]FSPG to visualize uptake/retention of [^{18}F]FSPG and xCT activity. Coronal plane images of the mice are shown, and axial plane images of the livers are shown.

FIG. 18B shows a quantification of the [^{18}F]FSPG retention (%ID/g) in the liver of WT mice, untreated *Asl^{Neo/Neo}* mice, and *hASL* mRNA-treated *Asl^{Neo/Neo}* mice intravenously administered [^{18}F]FSPG.

FIG. 19A shows a quantification of the [^{18}F]FSPG retention (%ID/g) in the skin of WT mice, untreated *Asl^{Neo/Neo}* mice, and *hASL* mRNA-treated *Asl^{Neo/Neo}* mice intravenously administered [^{18}F]FSPG.

FIG. 19B shows a quantification of the [^{18}F]FSPG retention (%ID/g) over time in the skin of untreated *Asl^{Neo/Neo}* mice and *hASL* mRNA-treated *Asl^{Neo/Neo}* mice intravenously administered [^{18}F]FSPG.

FIG. 20 shows a Western blot displaying xCT expression in the liver of WT mice, untreated *Asl*^{Neo/Neo} mice, and *hASL* mRNA-treated *Asl*^{Neo/Neo} mice.

FIG. 21 shows a quantification of total liver homocysteine (HCyS) levels in adult untreated *Asl*^{Neo/Neo} mice and *hASL* mRNA-treated *Asl*^{Neo/Neo} mice expressed as ratio of
5 observed levels in WT mice.

DETAILED DESCRIPTION

The present disclosure provides lipid nanoparticles (LNPs) comprising mRNA therapeutics for the treatment of argininosuccinic aciduria (“ASA”). Argininosuccinic aciduria (argininosuccinate lyase deficiency) is a rare autosomal recessive condition caused
10 by mutations in the gene encoding argininosuccinate lyase (ASL). ASL converts argininosuccinic acid into arginine and is integral to the liver-based urea cycle, which detoxifies neurotoxic ammonia. Ammonia, which is formed when proteins are broken down in the digestive process, is toxic if blood ammonia levels become too high. Thus, ASL defects can cause abnormal ammonia accumulation (hyperammonemia). LNPs comprising mRNA
15 therapeutics are particularly well-suited for the treatment of argininosuccinic aciduria, as the technology provides for the intracellular delivery of mRNA encoding ASL followed by *de novo* synthesis of functional ASL protein within target cells. After delivery of mRNA to the target cells, the desired ASL protein is expressed by the cells’ own translational machinery, and hence, fully functional ASL protein replaces the defective or missing protein.

One challenge associated with delivering nucleic acid-based therapeutics (*e.g.*, mRNA therapeutics) *in vivo* stems from the innate immune response, which can occur when the body’s immune system encounters foreign nucleic acids. Foreign mRNAs can activate the immune system *via* recognition through toll-like receptors (TLRs), in particular TLR7/8, which is activated by single-stranded RNA (ssRNA). In nonimmune cells, the recognition of
25 foreign mRNA can occur through the retinoic acid-inducible gene I (RIG-I). Immune recognition of foreign mRNAs can result in unwanted cytokine effects including interleukin-1β (IL-1β) production, tumor necrosis factor-α (TNF-α) distribution and a strong type I interferon (type I IFN) response. This disclosure features the incorporation of different modified nucleotides within therapeutic mRNAs to minimize the immune activation and
30 optimize the translation efficiency of mRNA to protein. Particular aspects feature a combination of nucleotide modification to reduce the innate immune response and sequence

optimization, in particular, within the open reading frame (ORF) of therapeutic mRNAs encoding ASL to enhance protein expression.

Certain embodiments of the mRNA therapeutic technology of the instant disclosure also feature delivery of mRNA encoding ASL *via* a lipid nanoparticle (LNP) delivery system.

5 Lipid nanoparticles (LNPs) are an ideal platform for the safe and effective delivery of mRNAs to target cells. LNPs have the unique ability to deliver nucleic acids by a mechanism involving cellular uptake, intracellular transport and endosomal release or endosomal escape. The present disclosure features ionizable amino lipid-based LNPs combined with mRNA encoding ASL which have improved properties when administered *in vivo*. Without being
10 bound in theory, it is believed that the ionizable amino lipid-based LNP formulations of the present disclosure have improved properties, for example, cellular uptake, intracellular transport and/or endosomal release or endosomal escape. LNPs administered by systemic route (*e.g.*, intravenous (IV) administration), for example, in a first administration, can accelerate the clearance of subsequently injected LNPs, for example, in further
15 administrations. This phenomenon is known as accelerated blood clearance (ABC) and is a key challenge, in particular, when replacing deficient enzymes (*e.g.*, ASL) in a therapeutic context. This is because repeat administration of mRNA therapeutics is in most instances essential to maintain necessary levels of enzyme in target tissues in subjects (*e.g.*, subjects suffering from argininosuccinic aciduria). Repeat dosing challenges can be addressed on
20 multiple levels. mRNA engineering and/or efficient delivery by LNPs can result in increased levels and or enhanced duration of protein (*e.g.*, ASL) being expressed following a first dose of administration, which in turn, can lengthen the time between first dose and subsequent dosing. It is known that the ABC phenomenon is, at least in part, transient in nature, with the immune responses underlying ABC resolving after sufficient time following systemic
25 administration. As such, increasing the duration of protein expression and/or activity following systemic delivery of an mRNA therapeutic of the disclosure in one aspect, combats the ABC phenomenon. Moreover, LNPs can be engineered to avoid immune sensing and/or recognition and can thus further avoid ABC upon subsequent or repeat dosing. An exemplary aspect of the disclosure features LNPs which have been engineered to have
30 reduced ABC.

1. Lipid Nanoparticle (LNP) Compositions

The present disclosure provides LNP compositions with advantageous properties. The lipid nanoparticle compositions described herein may be used for the delivery of therapeutic and/or prophylactic agents, *e.g.*, mRNAs, to mammalian cells or organs. For example, the lipid nanoparticles described herein have little or no immunogenicity. For example, the lipid compounds disclosed herein have a lower immunogenicity as compared to a reference lipid (*e.g.*, MC3, KC2, or DLinDMA). For example, a formulation comprising a lipid disclosed herein and a therapeutic or prophylactic agent, *e.g.*, mRNA, has an increased therapeutic index as compared to a corresponding formulation which comprises a reference lipid (*e.g.*, MC3, KC2, or DLinDMA) and the same therapeutic or prophylactic agent.

In some embodiments, the present application provides pharmaceutical compositions comprising:

- (a) a delivery agent comprising a lipid nanoparticle; and
- (b) a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide.

a. Lipid Nanoparticles

In some embodiments, polynucleotides of the present disclosure (*e.g.*, ASL mRNA) are included in a lipid nanoparticle (LNP). Lipid nanoparticles according to the present disclosure may comprise: (i) an ionizable lipid (*e.g.*, an ionizable amino lipid); (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-modified lipid. In some embodiments, lipid nanoparticles according to the present disclosure further comprise one or more polynucleotides of the present disclosure (*e.g.*, ASL mRNA).

The lipid nanoparticles according to the present disclosure can be generated using components, compositions, and methods as are generally known in the art, see for example PCT/US2016/052352; PCT/US2016/068300; PCT/US2017/037551; PCT/US2015/027400; PCT/US2016/047406; PCT/US2016000129; PCT/US2016/014280; PCT/US2016/014280; PCT/US2017/038426; PCT/US2014/027077; PCT/US2014/055394; PCT/US2016/52117; PCT/US2012/069610; PCT/US2017/027492; PCT/US2016/059575 and PCT/US2016/069491 all of which are incorporated by reference herein in their entirety.

In some embodiments, the lipid nanoparticle comprises an ionizable cationic lipid (e.g., an ionizable amino lipid) at a content of 20-60 mol.%, 25-60 mol.%, 30-60 mol.%, 35-60 mol.%, 40-60 mol.%, 45-60 mol.%, 20-55 mol.%, 25-55 mol.%, 30-55 mol.%, 35-55 mol.%, 40-55 mol.%, 45-55 mol.%, 20-50 mol.%, 25-50 mol.%, 30-50 mol.%, 35-50 mol.%, or 40-50 mol.%. For example, the lipid nanoparticle may comprise an ionizable cationic lipid (e.g., an ionizable amino lipid) at a content of 40-50 mol.%, 45-50 mol.%, 45-46 mol.%, 46-47 mol.%, 47-48 mol.%, 48-49 mol.%, or 49-50 mol.%, for example about 45 mol.%, about 45.5 mol.%, about 46 mol.%, about 46.5 mol.%, about 47 mol.%, about 47.5 mol.%, about 48 mol.%, about 48.5 mol.%, about 49 mol.%, or about 49.5 mol.% ionizable cationic lipid (e.g., an ionizable amino lipid).

In some embodiments, the lipid nanoparticle comprises a non-cationic helper lipid or phospholipid at a content of 5-25 mol.%. For example, the lipid nanoparticle may comprise a non-cationic helper lipid or phospholipid at a content of molar ratio of 5-25 mol.%, 5-20 mol.%, 5-15 mol.%, 10-25 mol.%, 10-20 mol.%, 10-15 mol.%, 5-6 mol.%, 6-7 mol.%, 7-8 mol.%, 8-9 mol.%, 9-10 mol.%, 10-11 mol.%, 11-12 mol.%, 12-13 mol.%, 13-14 mol.%, 14-15 mol.%, 10-14 mol.%, 10-13 mol.%, 10-12 mol.%, 10-11 mol.%, 9-15 mol.%, 9-14 mol.%, 9-13 mol.%, 9-12 mol.%, or 9-11 mol.% non-cationic lipid.

In some embodiments, the lipid nanoparticle comprises a sterol or other structural lipid at a content molar ratio of 25-55 mol.%, 25-50 mol.%, 25-45 mol.%, 25-40 mol.%, 25-35 mol.%, 30-55 mol.%, 30-50 mol.%, 30-45 mol.%, 30-40 mol.%, 30-35 mol.%, 35-55 mol.%, 35-50 mol.%, 35-45 mol.%, 35-40 mol.%, 25-30 mol.%, 30-35 mol.%, 25-28 mol.%, 28-30 mol.%, 30-33 mol.%, 35-38 mol.%, 38-40 mol.%, 36-40 mol.%, 37-40 mol.%, 38-40 mol.%, 38-39 mol.%, 36-40 mol.%, 37-40 mol.%, 36-39 mol.%, or 37-39 mol.%. For example, the lipid nanoparticle may comprise a sterol or other structural lipid at a content of about 30 mol.%, about 30.5 mol.%, about 31.0 mol.%, about 31.5 mol.%, about 32.0 mol.%, about 32.5 mol.%, about 33.0 mol.%, about 33.5 mol.%, about 34.0 mol.%, about 34.5 mol.%, about 35.0 mol.%, about 35.5 mol.%, about 36.0 mol.%, about 36.5 mol.%, about 37.0 mol.%, about 37.5 mol.%, about 38.0 mol.%, about 38.5 mol.%, about 39.0 mol.%, about 39.5 mol.%, about 40.0 mol.%, about 40.5 mol.%, about 41.0 mol.%, about 41.5 mol.%, about 42.0 mol.%, about 42.5 mol.%, about 43.0 mol.%, about 43.5 mol.%, about 44.0 mol.%, about 44.5 mol.%, or about 45.0 mol.%.

In some embodiments, the lipid nanoparticle comprises a PEG-modified lipid at a content of 0.5-15 mol.%, 1.0-15 mol.%, 1.5-15 mol.%, 2.0-15 mol.%, 2.5-15 mol.%, 3.0-15 mol.%, 3.5-15 mol.%, 4.0-15 mol.%, 4.5-15 mol.%, 5.0-15 mol.%, 10-15 mol.%, 0.5-10 mol.%, 0.5-5 mol.%, 0.5-4.5 mol.%, 0.5-4.0 mol.%, 0.5-3.5 mol.%, 0.5-3.0 mol.%, 0.5-2.5 mol.%, 0.5-2.0 mol.%, 0.5-1.5 mol.%, 0.5-1.0 mol.%, 1.0-10 mol.%, 1.0-5 mol.%, 1.0-4.5 mol.%, 1.0-4.0 mol.%, 1.0-3.5 mol.%, 1.0-3.0 mol.%, 1.0-2.5 mol.%, 1.0-2.0 mol.%, 1.0-1.5 mol.%, 1.5-5.0 mol.%, 1.5-4.5 mol.%, 1.5-4.0 mol.%, 1.5-3.5 mol.%, 1.5-3.0 mol.%, 1.5-2.5 mol.%, 1.5-2.0 mol.%, 2.0-5.0 mol.%, 2.0-4.5 mol.%, 2.0-4.0 mol.%, 2.0-3.5 mol.%, 2.0-3.0 mol.%, or 2.0-2.5 mol.%. For example, the lipid nanoparticle may comprise a PEG-modified lipid at a content of a about 0.5 mol.%, about 1.0 mol.%, about 1.5 mol.%, about 2.0 mol.%, about 2.5 mol.%, about 3.0 mol.%, about 3.5 mol.%, about 4.0 mol.%, about 4.5 mol.%, about 5.0 mol.%, about 6.0 mol.%, about 7.0 mol.%, about 8.0 mol.%, about 9.0 mol.%, about 10.0 mol.%, or about 15.0 mol.%.

In some embodiments, the lipid nanoparticle comprises: (i) 20 to 60 mol.% ionizable cationic lipid (*e.g.* ionizable amino lipid), (ii) 25 to 55 mol.% sterol or other structural lipid, (iii) 5 to 25 mol.% non-cationic lipid (*e.g.*, phospholipid), and (iv) 0.5 to 15 mol.% PEG-modified lipid.

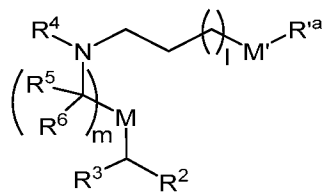
In some embodiments, the lipid nanoparticle comprises: (i) 40 to 50 mol.% ionizable cationic lipid (*e.g.* ionizable amino lipid), (ii) 30 to 45 mol.% sterol or other structural lipid, (iii) 5 to 15 mol.% non-cationic lipid (*e.g.*, phospholipid), and (iv) 1 to 5 mol.% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises: (i) 45 to 50 mol.% ionizable cationic lipid (*e.g.* ionizable amino lipid), (ii) 35 to 45 mol.% sterol or other structural lipid, (iii) 8 to 12 mol.% non-cationic lipid (*e.g.*, phospholipid), and (iv) 1.5 to 3.5 mol.% PEG-modified lipid.

In the following sections, “Compounds” numbered with an “I-” prefix (*e.g.*, “Compound I-1,” “Compound I-2,” “Compound I-3,” “Compound I-VI,” etc., indicate specific ionizable lipid compounds. Likewise, compounds numbered with a “P-” prefix (*e.g.*, “Compound P-I,” etc.) indicate a specific PEG-modified lipid compound.

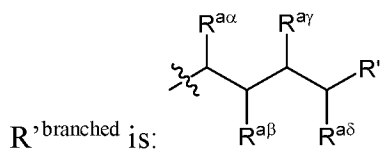
b. Ionizable amino lipids

In some embodiments, the lipid nanoparticle of the present disclosure comprises an ionizable cationic lipid (*e.g.*, an ionizable amino lipid) that is a compound of Formula (I):



(I) or its N-oxide, or a salt or isomer thereof,

5 wherein R'^a is R'^{branched} , wherein

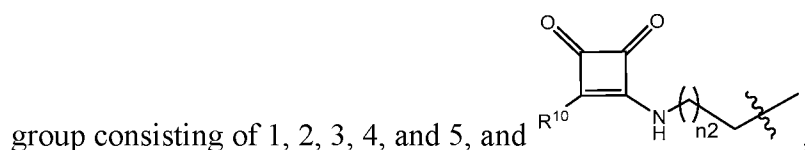


R'^{branched} is: ; wherein  denotes a point of attachment;

wherein $R^{a\alpha}$, $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each independently selected from the group consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl;

R^2 and R^3 are each independently selected from the group consisting of C₁₋₁₄ alkyl and
10 C₂₋₁₄ alkenyl;

R^4 is selected from the group consisting of $-(CH_2)_nOH$, wherein n is selected from the



group consisting of 1, 2, 3, 4, and 5, and

wherein  denotes a point of attachment; wherein

R^{10} is N(R)₂; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; and n₂ is selected from the group consisting of 1, 2, 3, 4,
15 5, 6, 7, 8, 9, and 10;

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

each R^6 is independently selected from the group consisting of C₁₋₃ alkyl,
20 C₂₋₃ alkenyl, and H;

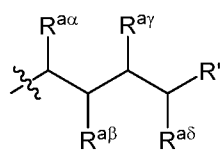
M and M' are each independently selected from the group consisting of -C(O)O- and -OC(O)-;

R^7 is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

l is selected from the group consisting of 1, 2, 3, 4, and 5; and

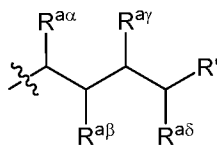
25 m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13.

In some embodiments, in Formula (I), $R^{a\alpha}$ is $R^{a\text{branched}}$, $R^{a\text{branched}}$ is



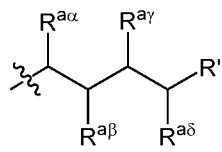
; denotes a point of attachment; $R^{a\alpha}$, $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each H; R^2 and R^3 are each C_{1-14} alkyl; R^4 is $-(CH_2)_nOH$; n is 2; each R^5 is H; each R^6 is H; M and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 5; and m is 7.

5 In some embodiments, in Formula (I), $R^{a\alpha}$ is $R^{a\text{branched}}$, $R^{a\text{branched}}$ is



; denotes a point of attachment; $R^{a\alpha}$, $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each H; R^2 and R^3 are each C_{1-14} alkyl; R^4 is $-(CH_2)_nOH$; n is 2; each R^5 is H; each R^6 is H; M and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 3; and m is 7.

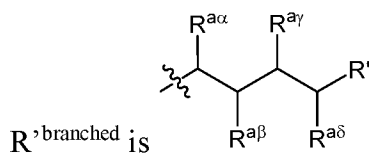
In some embodiments of the compounds of Formula (I), $R^{a\alpha}$ is $R^{a\text{branched}}$, $R^{a\text{branched}}$ is



10 ; denotes a point of attachment; $R^{a\alpha}$ is C_{2-12} alkyl; $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are

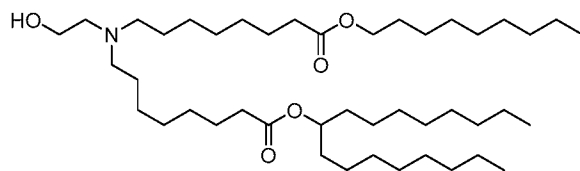
each H; R^2 and R^3 are each C_{1-14} alkyl; R^4 is ; R^{10} is $NH(C_{1-6}$ alkyl); n_2 is 2; R^5 is H; each R^6 is H; M and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 5; and m is 7.

In some embodiments of the compounds of Formula (I), $R^{a\alpha}$ is $R^{a\text{branched}}$,

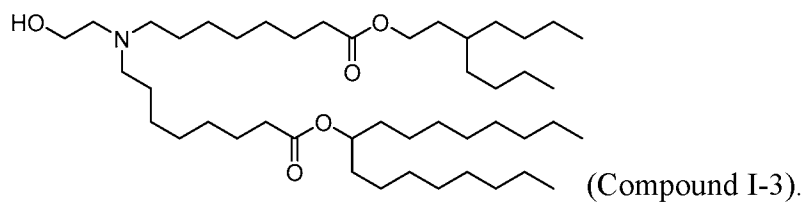
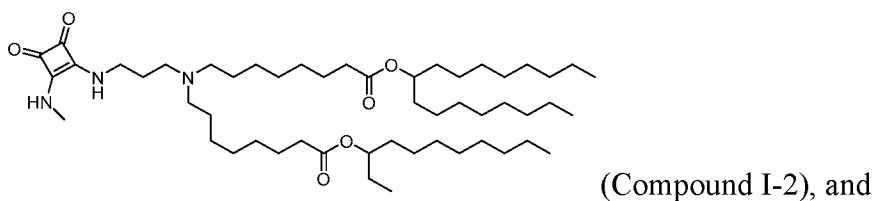


15 $R^{a\text{branched}}$ is $R^{a\text{branched}}$; denotes a point of attachment; $R^{a\alpha}$, $R^{a\beta}$, and $R^{a\delta}$ are each H; $R^{a\gamma}$ is C_{2-12} alkyl; R^2 and R^3 are each C_{1-14} alkyl; R^4 is $-(CH_2)_nOH$; n is 2; each R^5 is H; each R^6 is H; M and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 5; and m is 7.

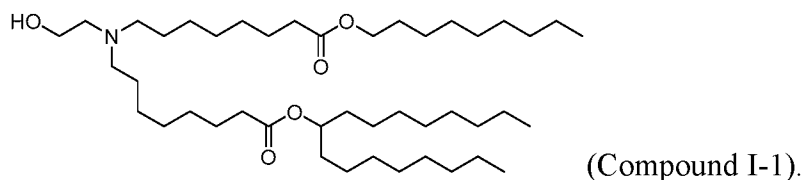
In some embodiments, the compound of Formula (I) is selected from:



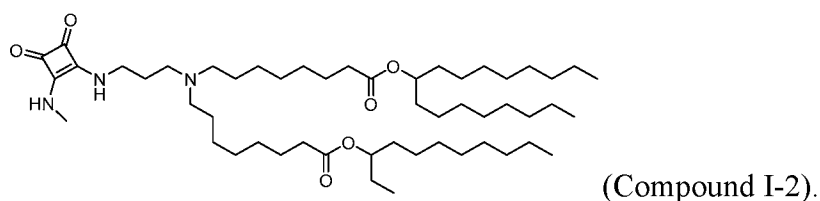
(Compound I-1),



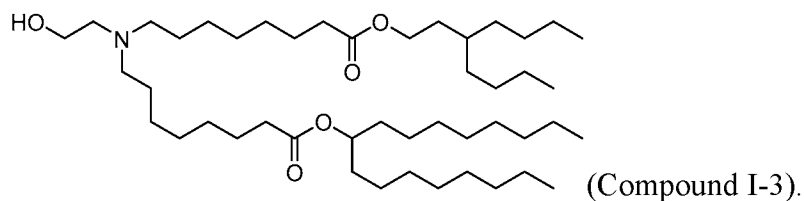
In some embodiments, the compound of Formula (I) is:



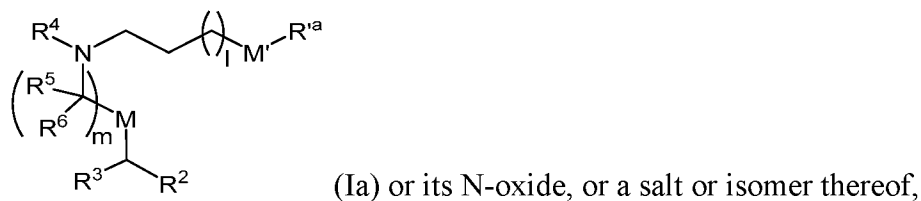
5 In some embodiments, the compound of Formula (I) is:



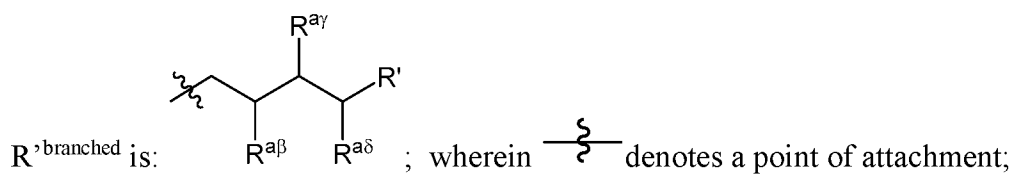
In some embodiments, the compound of Formula (I) is:



10 In some aspects, the disclosure relates to a compound of Formula (Ia):



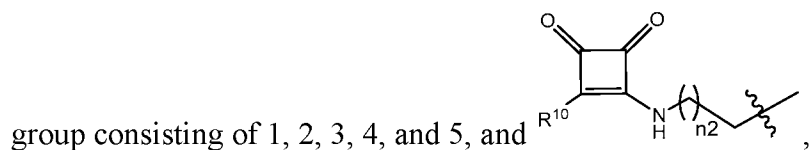
wherein R'^a is R'^{branched} , wherein



wherein $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each independently selected from the group consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl;

R² and R³ are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl;

5 R⁴ is selected from the group consisting of -(CH₂)_nOH wherein n is selected from the



wherein $\frac{\text{S}}{\text{Z}}$ denotes a point of attachment; wherein

R¹⁰ is N(R)₂; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; and n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

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

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

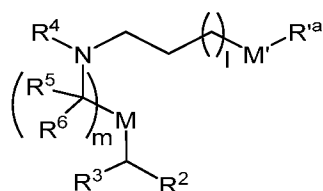
15 M and M' are each independently selected from the group consisting of -C(O)O- and -OC(O)-;

R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

1 is selected from the group consisting of 1, 2, 3, 4, and 5; and

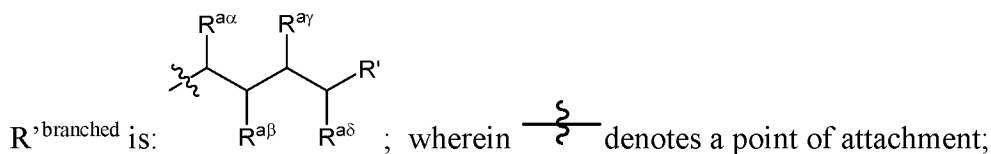
m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13.

20 In some aspects, the disclosure relates to a compound of Formula (Ib):



(Ib) or its N-oxide, or a salt or isomer thereof,

wherein R^a is R^{branched}; wherein



wherein $R^{a\alpha}$, $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each independently selected from the group

25 consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl;

R^2 and R^3 are each independently selected from the group consisting of C_{1-14} alkyl and C_{2-14} alkenyl;

R^4 is $-(CH_2)_nOH$, wherein n is selected from the group consisting of 1, 2, 3, 4, and 5;

each R^5 is independently selected from the group consisting of C_{1-3} alkyl,

5 C_{2-3} alkenyl, and H;

each R^6 is independently selected from the group consisting of C_{1-3} alkyl,

C_{2-3} alkenyl, and H;

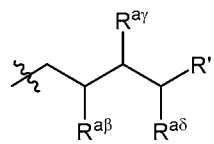
M and M' are each independently selected from the group consisting of $-C(O)O-$ and $-OC(O)-$;

10 R' is a C_{1-12} alkyl or C_{2-12} alkenyl;

l is selected from the group consisting of 1, 2, 3, 4, and 5; and

m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13.

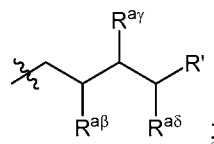
In some embodiments of Formula (I) or (Ib), R'^a is R'^{branched} , R'^{branched} is



;  denotes a point of attachment; $R^{aβ}$, $R^{aγ}$, and $R^{aδ}$ are each H; R^2 and

15 R^3 are each C_{1-14} alkyl; R^4 is $-(CH_2)_nOH$; n is 2; each R^5 is H; each R^6 is H; M and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 5; and m is 7.

In some embodiments of Formula (I) or (Ib), R'^a is R'^{branched} , R'^{branched} is

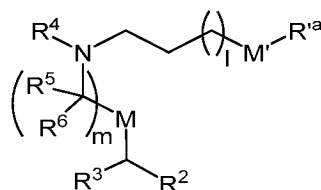


;  denotes a point of attachment; $R^{aβ}$ and $R^{aδ}$ are each H; $R^{aγ}$ is C_{2-12}

alkyl; R^2 and R^3 are each C_{1-14} alkyl; R^4 is $-(CH_2)_nOH$; n is 2; each R^5 is H; each R^6 is H; M

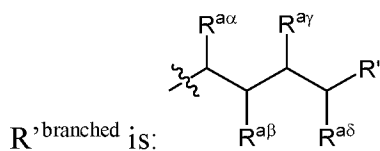
20 and M' are each $-C(O)O-$; R' is a C_{1-12} alkyl; l is 5; and m is 7.

In some embodiments, the disclosure relates to a compound of Formula (Ic):



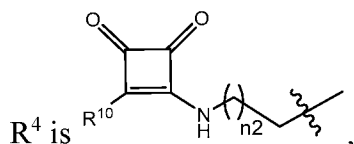
(Ic) or its N-oxide, or a salt or isomer thereof,

wherein R'^a is R'^{branched} , wherein



wherein $R^{a\alpha}$, $R^{a\beta}$, $R^{a\gamma}$, and $R^{a\delta}$ are each independently selected from the group consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl;

5 R² and R³ are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl;



wherein  denotes a point of attachment; wherein

R^{10} is $N(R)_2$; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

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

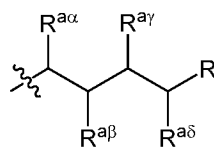
each R⁶ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

15 M and M' are each independently selected from the group consisting of -C(O)O- and
-OC(O)-;

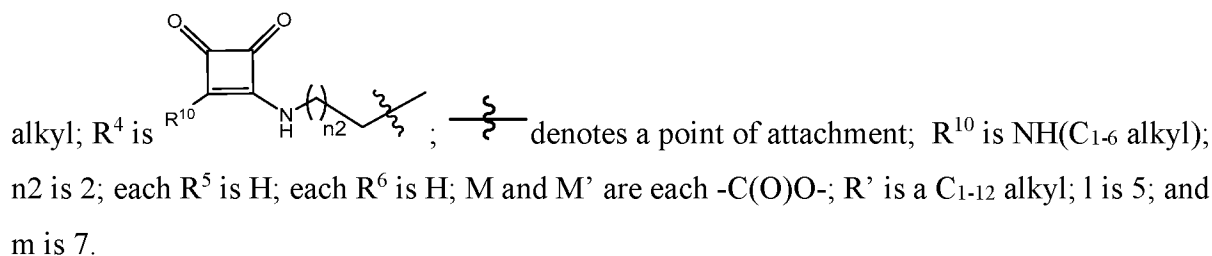
R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

1 is selected from the group consisting of 1, 2, 3, 4, and 5; and

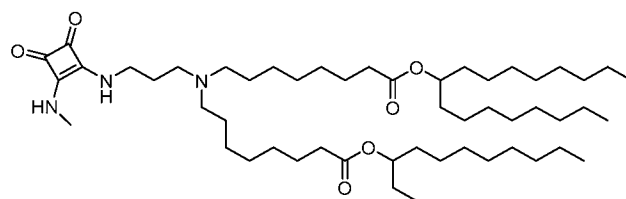
m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13.



20 In some embodiments, R^a is R^{branched}; R^{branched} is $\text{R}^{\alpha\beta} \text{---} \text{R}^{\alpha\delta}$;  denotes a point of attachment; R^β, R^γ, and R^δ are each H; R^α is C₂₋₁₂ alkyl; R² and R³ are each C₁₋₁₄

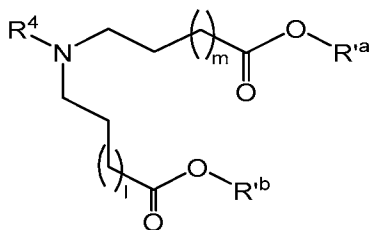


In some embodiments, the compound of Formula (Ic) is:



(Compound I-2).

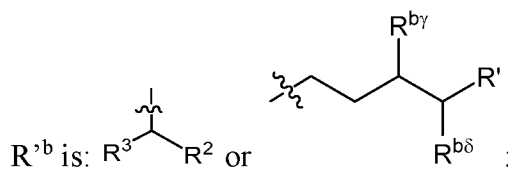
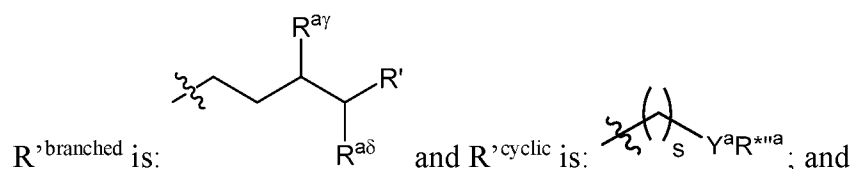
In some aspects, the disclosure relates to a compound of Formula (II):



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

5

wherein R^{a} is R^{branched} or R^{cyclic} ; wherein



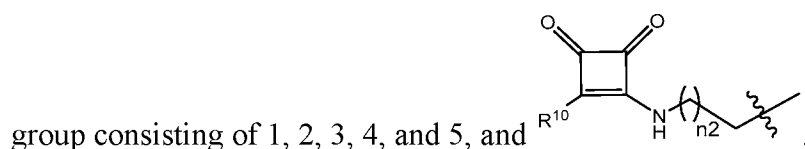
wherein denotes a point of attachment;

10 $R^{a\gamma}$ and $R^{a\delta}$ are each independently selected from the group consisting of H, C_{1-12} alkyl, and C_{2-12} alkenyl, wherein at least one of $R^{a\gamma}$ and $R^{a\delta}$ is selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

$R^{b\gamma}$ and $R^{b\delta}$ are each independently selected from the group consisting of H, C_{1-12} alkyl, and C_{2-12} alkenyl, wherein at least one of $R^{b\gamma}$ and $R^{b\delta}$ is selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

15 R^2 and R^3 are each independently selected from the group consisting of C_{1-14} alkyl and C_{2-14} alkenyl;

R^4 is selected from the group consisting of $-(CH_2)_nOH$ wherein n is selected from the



wherein denotes a point of attachment; wherein

R^{10} is $N(R)_2$; each R is independently selected from the group consisting of C_{1-6} alkyl, C_{2-3} alkenyl, and H; and n_2 is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

each R' independently is a C_{1-12} alkyl or C_{2-12} alkenyl;

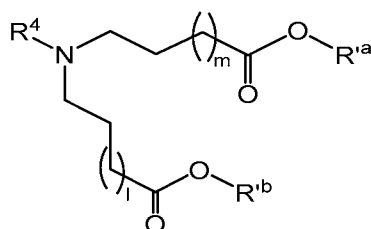
5 Y^a is a C_{3-6} carbocycle;

R^{*a} is selected from the group consisting of C_{1-15} alkyl and C_{2-15} alkenyl; and s is 2 or 3;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

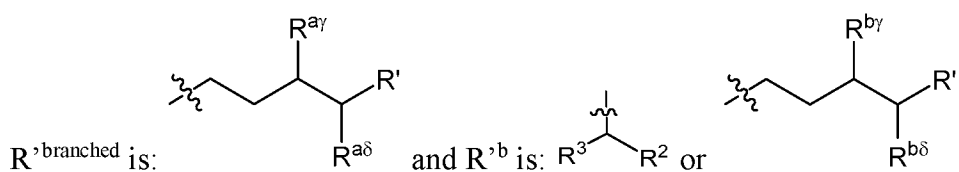
l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

10 In some aspects, the disclosure relates to a compound of Formula (II-a):



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

wherein R^{*a} is R^{*b} branched or R^{*b} cyclic; wherein



wherein $\text{---}\text{S}\text{---}$ denotes a point of attachment;

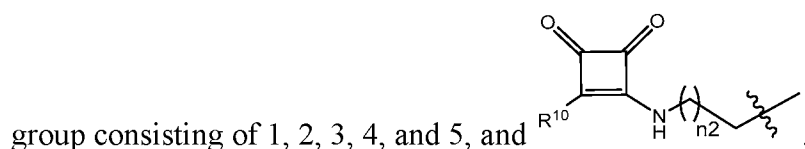
15 $R^{a\gamma}$ and $R^{a\delta}$ are each independently selected from the group consisting of H, C_{1-12} alkyl, and C_{2-12} alkenyl, wherein at least one of $R^{a\gamma}$ and $R^{a\delta}$ is selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

$R^{b\gamma}$ and $R^{b\delta}$ are each independently selected from the group consisting of H, C_{1-12} alkyl, and C_{2-12} alkenyl, wherein at least one of $R^{b\gamma}$ and $R^{b\delta}$ is selected from the group

20 consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

R^2 and R^3 are each independently selected from the group consisting of C_{1-14} alkyl and C_{2-14} alkenyl;

R^4 is selected from the group consisting of $-(CH_2)_nOH$ wherein n is selected from the



wherein  denotes a point of attachment; wherein

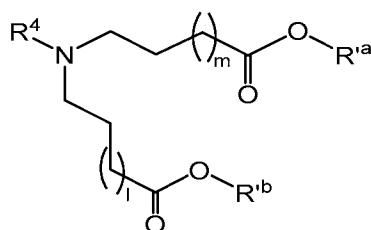
R^{10} is $N(R)_2$; each R is independently selected from the group consisting of C_{1-6} alkyl, C_{2-3} alkenyl, and H; and n_2 is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

5 each R' independently is a C_{1-12} alkyl or C_{2-12} alkenyl;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

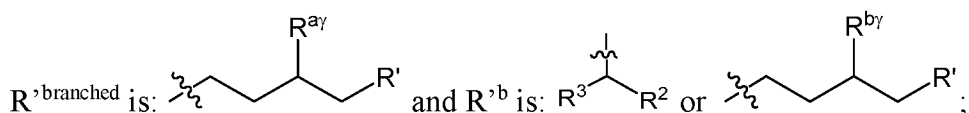
l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

In some aspects, the disclosure relates to a compound of Formula (II-b):



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

10 wherein R'^a is $R'^{branched}$ or R'^{cyclic} ; wherein

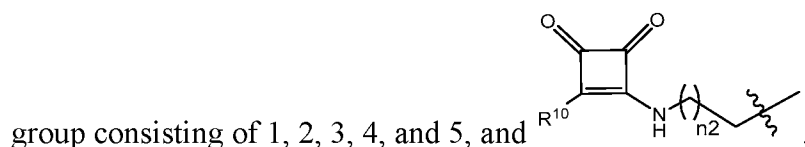


wherein  denotes a point of attachment;

R^{ay} and R^{by} are each independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

15 R^2 and R^3 are each independently selected from the group consisting of C_{1-14} alkyl and C_{2-14} alkenyl;

R^4 is selected from the group consisting of $-(CH_2)_nOH$ wherein n is selected from the



group consisting of 1, 2, 3, 4, and 5, and

wherein  denotes a point of attachment; wherein

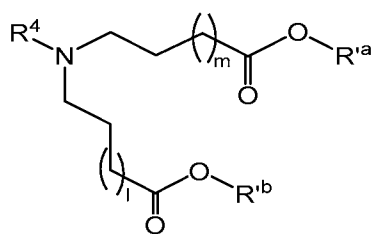
20 R^{10} is $N(R)_2$; each R is independently selected from the group consisting of C_{1-6} alkyl, C_{2-3} alkenyl, and H; and n_2 is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

each R' independently is a C_{1-12} alkyl or C_{2-12} alkenyl;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

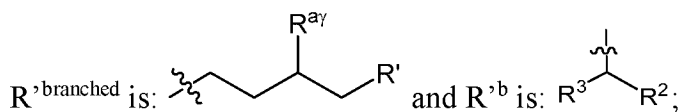
25 l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

In some aspects, the disclosure relates to a compound of Formula (II-c):



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

wherein R'a is R'branched or R'cyclic; wherein

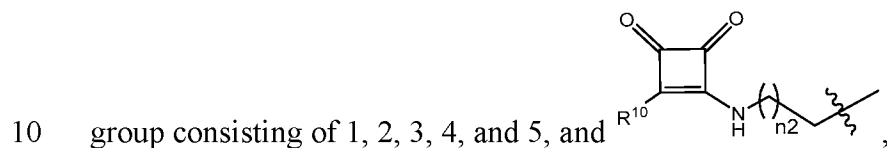


5 wherein  denotes a point of attachment;

wherein R'ay is selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

R² and R³ are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl;

R⁴ is selected from the group consisting of -(CH₂)_nOH wherein n is selected from the



wherein  denotes a point of attachment; wherein

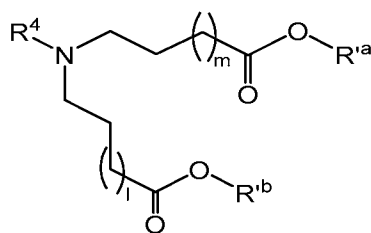
R¹⁰ is N(R)₂; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; and n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

15 R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

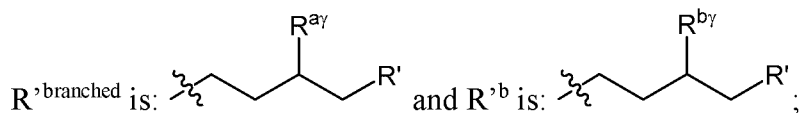
l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

In some aspects, the disclosure relates to a compound of Formula (II-d):



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

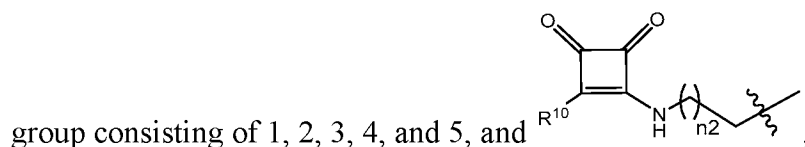
20 wherein R'a is R'branched or R'cyclic; wherein



wherein  denotes a point of attachment;

wherein R^{ay} and R^{by} are each independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

5 R^4 is selected from the group consisting of $-(\text{CH}_2)_n\text{OH}$ wherein n is selected from the



wherein  denotes a point of attachment; wherein

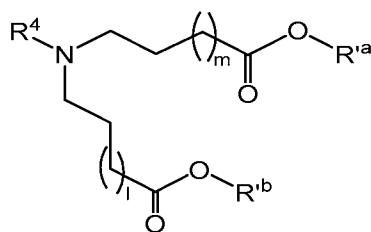
R^{10} is $\text{N}(\text{R})_2$; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H; and n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

each R' independently is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

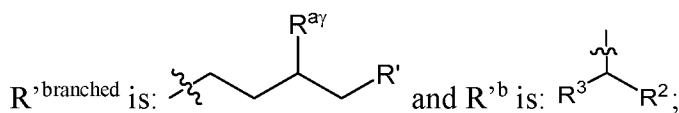
l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

In some aspects, the disclosure relates to a compound of Formula (II-e):



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

wherein R'^{a} is R'^{branched} or R'^{cyclic} ; wherein



wherein  denotes a point of attachment;

wherein R^{ay} is selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

20 R^2 and R^3 are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl;

R^4 is $-(\text{CH}_2)_n\text{OH}$ wherein n is selected from the group consisting of 1, 2, 3, 4, and 5;

R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

m is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9;

l is selected from 1, 2, 3, 4, 5, 6, 7, 8, and 9.

In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), m and l are each independently selected from 4, 5, and 6. In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), m and l are each 5.

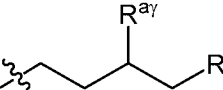
5 In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), each R' independently is a C₁₋₁₂ alkyl. In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), each R' independently is a C₂₋₅ alkyl.

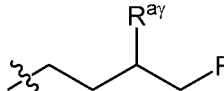
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

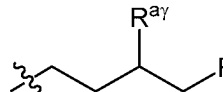
(II-e), R'^b is:  and R² and R³ are each independently a C₁₋₁₄ alkyl. In some
10 embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R'^b is:

 and R² and R³ are each independently a C₆₋₁₀ alkyl. In some embodiments of the
compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R'^b is:  and R² and R³
are each a C₈ alkyl.

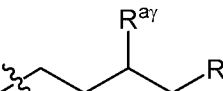
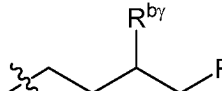
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

15 (II-e), R'^{branched} is:  and R'^b is: , R^{ay} is a C₁₋₁₂ alkyl and R² and R³
are each independently a C₆₋₁₀ alkyl. In some embodiments of the compound of Formula (II),

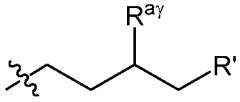
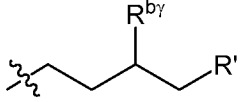
(II-a), (II-b), (II-c), (II-d), or (II-e), R'^{branched} is:  and R'^b is: , R^{ay} is
a C₂₋₆ alkyl and R² and R³ are each independently a C₆₋₁₀ alkyl. In some embodiments of the

compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R'^{branched} is: 
20 and R'^b is: , R^{ay} is a C₂₋₆ alkyl, and R² and R³ are each a C₈ alkyl.

In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

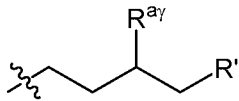
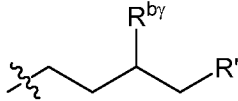
(II-e), R'^{branched} is: , R'^b is: , and R^{ay} and R^{by} are each a

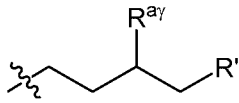
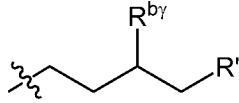
C₁₋₁₂ alkyl. In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-

d), or (II-e), R^{'branched} is: , R^{'b} is: , and R^{ay} and R^{by} are each a C₂₋₆ alkyl.

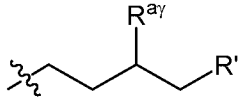
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), m and l are each independently selected from 4, 5, and 6 and each R' independently is a C₁₋₁₂ alkyl. In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), m and l are each 5 and each R' independently is a C₂₋₅ alkyl.

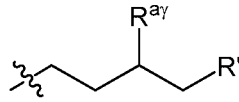
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

(II-e), R^{'branched} is: , R^{'b} is: , m and l are each independently selected from 4, 5, and 6, each R' independently is a C₁₋₁₂ alkyl, and R^{ay} and R^{by} are each a C₁₋₁₂ alkyl. In some embodiments of the compound of Formula (II), (II-a), (II-

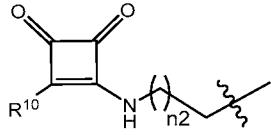
b), (II-c), (II-d), or (II-e), R^{'branched} is: , R^{'b} is: , m and l are each 5, each R' independently is a C₂₋₅ alkyl, and R^{ay} and R^{by} are each a C₂₋₆ alkyl.

In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

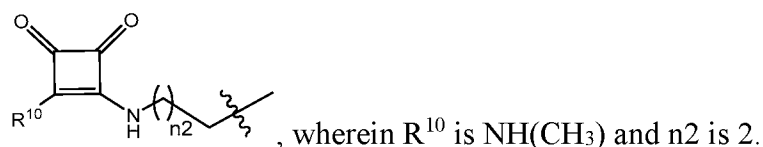
(II-e), R^{'branched} is: , and R^{'b} is: , m and l are each independently selected from 4, 5, and 6, R' is a C₁₋₁₂ alkyl, R^{ay} is a C₁₋₁₂ alkyl and R² and R³ are each independently a C₆₋₁₀ alkyl. In some embodiments of the compound of Formula (II), (II-a),

(II-b), (II-c), (II-d), or (II-e), R^{'branched} is: , and R^{'b} is: , m and l are each 5, R' is a C₂₋₅ alkyl, R^{ay} is a C₂₋₆ alkyl, and R² and R³ are each a C₈ alkyl.

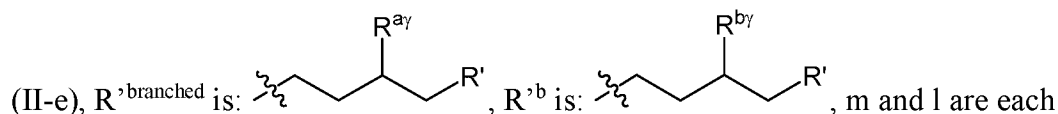
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

(II-e), R⁴ is , wherein R¹⁰ is NH(C₁₋₆ alkyl) and n₂ is 2. In some

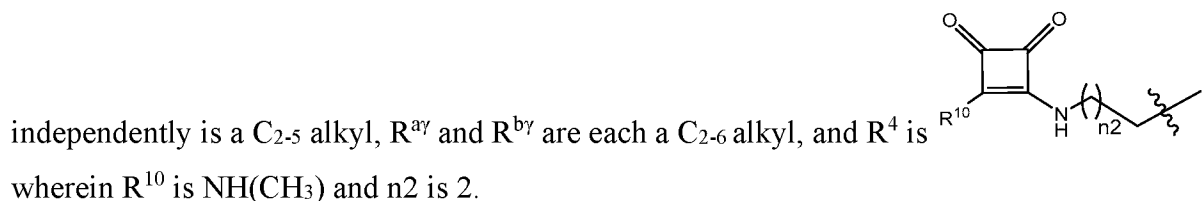
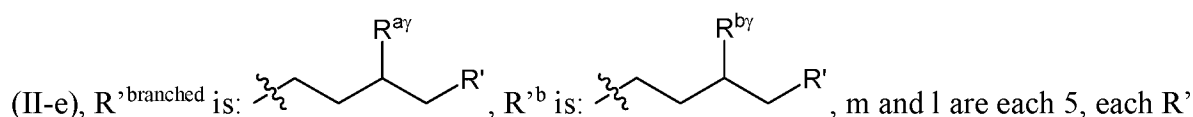
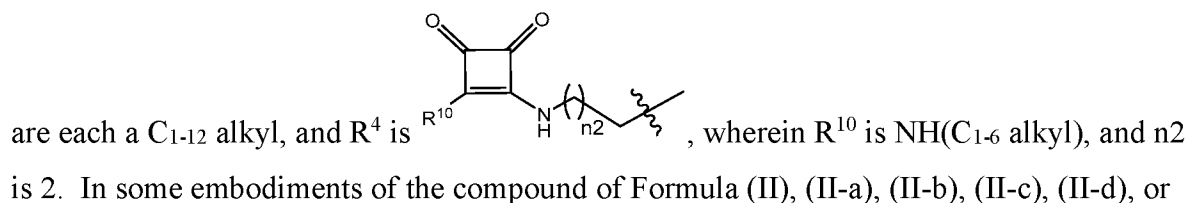
embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R^4 is



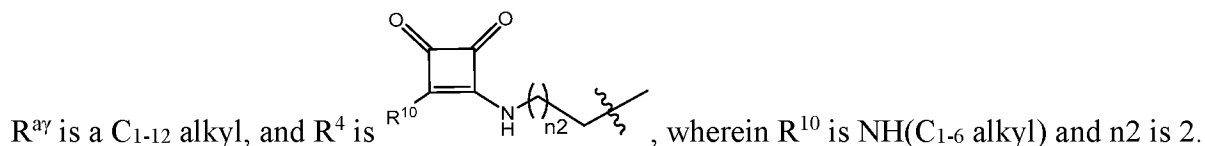
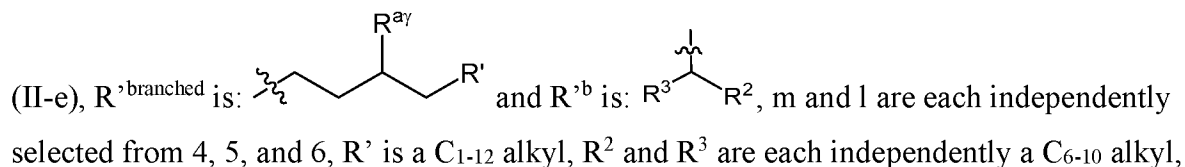
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or



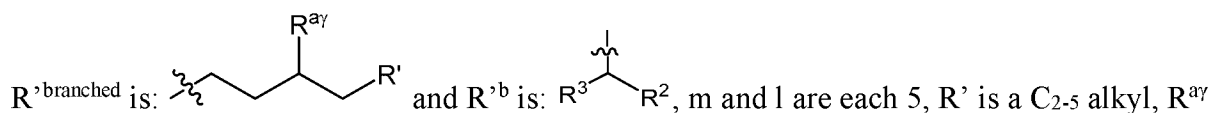
5 independently selected from 4, 5, and 6, each R' independently is a C_{1-12} alkyl, R^{ay} and R^{by}

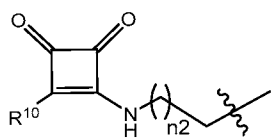


In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or



15 In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e),

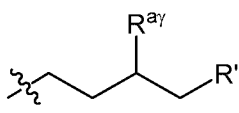
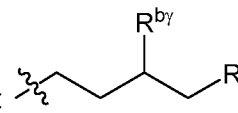


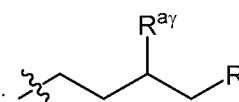
is a C₂₋₆ alkyl, R² and R³ are each a C₈ alkyl, and R⁴ is , wherein R¹⁰ is NH(CH₃) and n₂ is 2.

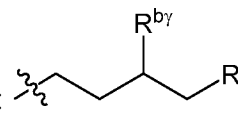
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R⁴ is -(CH₂)_nOH and n is 2, 3, or 4. In some embodiments of the compound of

5 Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R⁴ is -(CH₂)_nOH and n is 2.

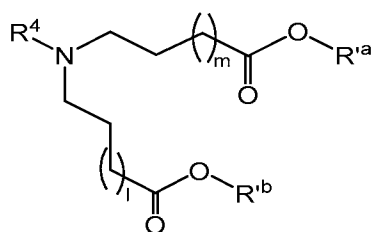
In some embodiments of the compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or

(II-e), R^{'branched} is: , R^{'b} is: , m and l are each independently selected from 4, 5, and 6, each R' independently is a C₁₋₁₂ alkyl, R^{ay} and R^{by} are each a C₁₋₁₂ alkyl, R⁴ is -(CH₂)_nOH, and n is 2, 3, or 4. In some embodiments of the

10 compound of Formula (II), (II-a), (II-b), (II-c), (II-d), or (II-e), R^{'branched} is: 

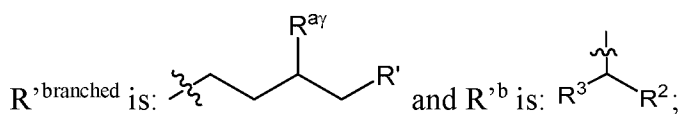
, R^{'b} is: , m and l are each 5, each R' independently is a C₂₋₅ alkyl, R^{ay} and R^{by} are each a C₂₋₆ alkyl, R⁴ is -(CH₂)_nOH, and n is 2.

In some aspects, the disclosure relates to a compound of Formula (II-f):



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

15 wherein R^{'a} is R^{'branched} or R^{'cyclic}; wherein



wherein  denotes a point of attachment;

R^{ay} is a C₁₋₁₂ alkyl;

R² and R³ are each independently a C₁₋₁₄ alkyl;

20 R⁴ is -(CH₂)_nOH wherein n is selected from the group consisting of 1, 2, 3, 4, and 5;

R' is a C₁₋₁₂ alkyl;

m is selected from 4, 5, and 6; and

l is selected from 4, 5, and 6.

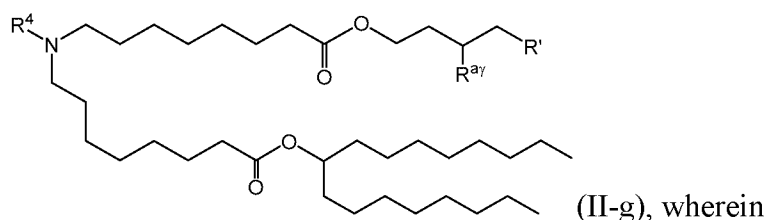
In some embodiments of the compound of Formula (II-f), m and l are each 5, and n is

2, 3, or 4.

In some embodiments of the compound of Formula (II-f) R' is a C₂₋₅ alkyl, R^{ay} is a C₂₋₆ alkyl, and R² and R³ are each a C₆₋₁₀ alkyl.

In some embodiments of the compound of Formula (II-f), m and l are each 5, n is 2, 3, or 4, R' is a C₂₋₅ alkyl, R^{ay} is a C₂₋₆ alkyl, and R² and R³ are each a C₆₋₁₀ alkyl.

In some aspects, the disclosure relates to a compound of Formula (II-g):

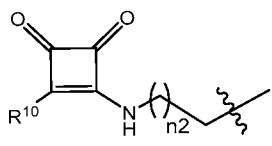


R^{ay} is a C₂₋₆ alkyl;

R' is a C₂₋₅ alkyl; and

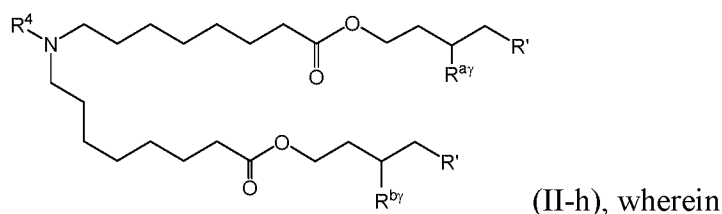
R⁴ is selected from the group consisting of -(CH₂)_nOH wherein n is selected from the

group consisting of 3, 4, and 5, and



wherein  denotes a point of attachment, R¹⁰ is NH(C₁₋₆ alkyl), and n₂ is selected from the group consisting of 1, 2, and 3.

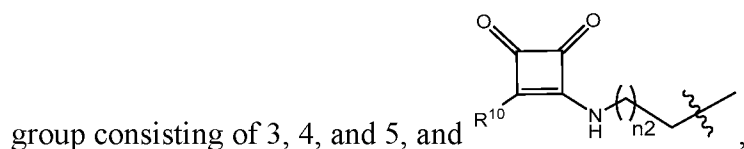
In some aspects, the disclosure relates to a compound of Formula (II-h):



R^{ay} and R^{by} are each independently a C₂₋₆ alkyl;

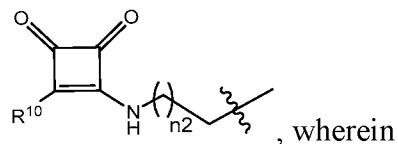
each R' independently is a C₂₋₅ alkyl; and

R^4 is selected from the group consisting of $-(CH_2)_nOH$ wherein n is selected from the



wherein  denotes a point of attachment, R^{10} is $NH(C_{1-6} \text{ alkyl})$, and n_2 is selected from the group consisting of 1, 2, and 3.

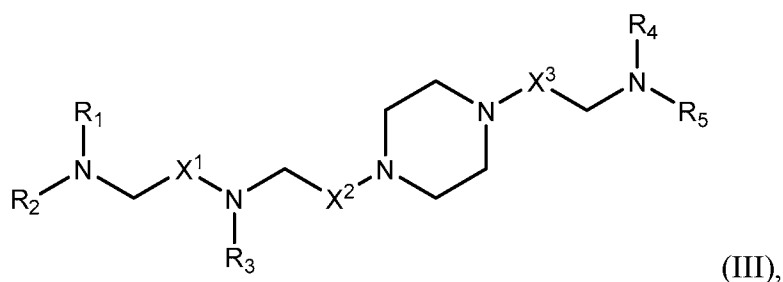
5 In some embodiments of the compound of Formula (II-g) or (II-h), R^4 is



R^{10} is $NH(CH_3)$ and n_2 is 2.

In some embodiments of the compound of Formula (II-g) or (II-h), R^4 is $-(CH_2)_2OH$.

In some aspects, the disclosure relates to a compound having the Formula (III):



10 or a salt or isomer thereof, wherein

R_1 , R_2 , R_3 , R_4 , and R_5 are independently selected from the group consisting of C_{5-20} alkyl, C_{5-20} alkenyl, $-R''MR'$, $-R^*YR''$, $-YR''$, and $-R^*OR''$;

each M is independently selected from the group consisting

15 of $-C(O)O-$, $-OC(O)-$, $-OC(O)O-$, $-C(O)N(R')$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$,

$-CH(OH)-$, $-P(O)(OR')O-$, $-S(O)_2-$, an aryl group, and a heteroaryl group;

X^1 , X^2 , and X^3 are independently selected from the group consisting of a bond, $-CH_2-$, $-(CH_2)_2-$, $-CHR-$, $-CHY-$, $-C(O)-$, $-C(O)O-$, $-OC(O)-$, $-C(O)-CH_2-$, $-CH_2-C(O)-$, $-C(O)O-CH_2-$, $-OC(O)-CH_2-$, $-CH_2-C(O)O-$, $-CH_2-OC(O)-$, $-CH(OH)-$, $-C(S)-$, and $-CH(SH)-$;

each Y is independently a C_{3-6} carbocycle;

each R^* is independently selected from the group consisting of C_{1-12} alkyl and C_{2-12} alkenyl;

each R is independently selected from the group consisting of C₁₋₃ alkyl and a C₃₋₆ carbocycle;

each R' is independently selected from the group consisting of C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, and H; and

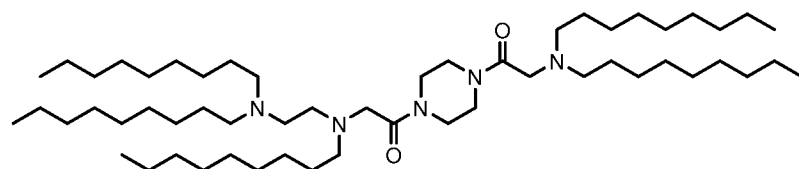
5 each R'' is independently selected from the group consisting of C₃₋₁₂ alkyl and C₃₋₁₂ alkenyl, and wherein:

i) at least one of X¹, X², and X³ is not -CH₂-; and/or

ii) at least one of R₁, R₂, R₃, R₄, and R₅ is -R''MR'.

10 In some embodiments, R₁, R₂, R₃, R₄, and R₅ are each C₅₋₂₀ alkyl; X¹ is -CH₂-; and X² and X³ are each -C(O)-.

In some embodiments, the compound of Formula (III) is:



(Compound I-VI), or a salt or

isomer thereof.

c. Phospholipids

15 The lipid composition of the lipid nanoparticle composition disclosed herein can comprise one or more phospholipids, for example, one or more saturated or (poly)unsaturated phospholipids or a combination thereof. In general, phospholipids comprise a phospholipid moiety and one or more fatty acid moieties.

20 A phospholipid moiety can be selected, for example, from the non-limiting group consisting of phosphatidyl choline, phosphatidyl ethanolamine, phosphatidyl glycerol, phosphatidyl serine, phosphatidic acid, 2-lysophosphatidyl choline, and a sphingomyelin.

A fatty acid moiety can be selected, for example, from the non-limiting group consisting of lauric acid, myristic acid, myristoleic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, alpha-linolenic acid, erucic acid, phytanoic acid, 25 arachidic acid, arachidonic acid, eicosapentaenoic acid, behenic acid, docosapentaenoic acid, and docosahexaenoic acid.

Particular phospholipids can facilitate fusion to a membrane. For example, a cationic phospholipid can interact with one or more negatively charged phospholipids of a membrane

(*e.g.*, a cellular or intracellular membrane). Fusion of a phospholipid to a membrane can allow one or more elements (*e.g.*, a therapeutic agent) of a lipid-containing composition (*e.g.*, LNPs) to pass through the membrane permitting, *e.g.*, delivery of the one or more elements to a target tissue.

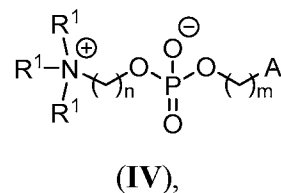
5 Non-natural phospholipid species including natural species with modifications and substitutions including branching, oxidation, cyclization, and alkynes are also contemplated. For example, a phospholipid can be functionalized with or cross-linked to one or more alkynes (*e.g.*, an alkenyl group in which one or more double bonds is replaced with a triple bond). Under appropriate reaction conditions, an alkyne group can undergo a copper-
10 catalyzed cycloaddition upon exposure to an azide. Such reactions can be useful in functionalizing a lipid bilayer of a nanoparticle composition to facilitate membrane permeation or cellular recognition or in conjugating a nanoparticle composition to a useful component such as a targeting or imaging moiety (*e.g.*, a dye).

Phospholipids include, but are not limited to, glycerophospholipids such as
15 phosphatidylcholines, phosphatidylethanolamines, phosphatidylserines, phosphatidylinositols, phosphatidyl glycerols, and phosphatidic acids. Phospholipids also include phosphosphingolipid, such as sphingomyelin.

In some embodiments, a phospholipid of the present disclosure comprises 1,2-
distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dioleoyl-sn-glycero-3-
20 phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-3-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-
25 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-
30 phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-

phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and mixtures thereof.

In certain embodiments, a phospholipid useful or potentially useful in the present disclosure is an analog or variant of DSPC. In certain embodiments, a phospholipid useful or potentially useful in the present disclosure is a compound of Formula (IV):

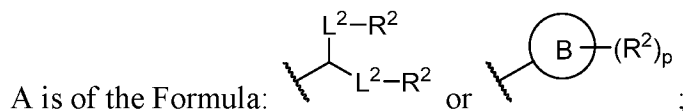


or a salt thereof, wherein:

each R^1 is independently optionally substituted alkyl; or optionally two R^1 are joined together with the intervening atoms to form optionally substituted monocyclic carbocyclyl or optionally substituted monocyclic heterocyclyl; or optionally three R^1 are joined together with the intervening atoms to form optionally substituted bicyclic carbocyclyl or optionally substitute bicyclic heterocyclyl;

n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

m is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;



each instance of L^2 is independently a bond or optionally substituted C_{1-6} alkylene, wherein one methylene unit of the optionally substituted C_{1-6} alkylene is optionally replaced with O, $\text{N}(\text{R}^N)$, S, C(O), $\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})$, $\text{C}(\text{O})\text{O}$, $\text{OC}(\text{O})$, $\text{OC}(\text{O})\text{O}$, $\text{OC}(\text{O})\text{N}(\text{R}^N)$, - $\text{NR}^N\text{C}(\text{O})\text{O}$, or $\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)$;

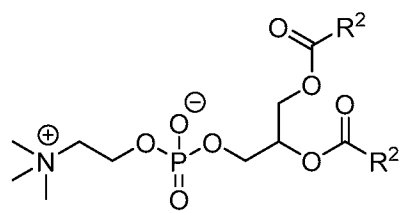
each instance of R^2 is independently optionally substituted C_{1-30} alkyl, optionally substituted C_{1-30} alkenyl, or optionally substituted C_{1-30} alkynyl; optionally wherein one or more methylene units of R^2 are independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, $\text{N}(\text{R}^N)$, O, S, C(O), $\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})$, - $\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{C}(\text{O})\text{O}$, $\text{OC}(\text{O})$, $\text{OC}(\text{O})\text{O}$, $\text{OC}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})\text{O}$, $\text{C}(\text{O})\text{S}$, $\text{SC}(\text{O})$, - $\text{C}(=\text{NR}^N)$, $\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(=\text{NR}^N)$, $\text{NR}^N\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)$, C(S), $\text{C}(\text{S})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{S})$, $\text{NR}^N\text{C}(\text{S})\text{N}(\text{R}^N)$, S(O), OS(O), S(O)O, OS(O)O, OS(O)₂, S(O)₂O, OS(O)₂O, $\text{N}(\text{R}^N)\text{S}(\text{O})$, - $\text{S}(\text{O})\text{N}(\text{R}^N)$, $\text{N}(\text{R}^N)\text{S}(\text{O})\text{N}(\text{R}^N)$, OS(O)N(R^N), $\text{N}(\text{R}^N)\text{S}(\text{O})\text{O}$, S(O)₂, $\text{N}(\text{R}^N)\text{S}(\text{O})_2$, S(O)₂N(R^N), $\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{N}(\text{R}^N)$, OS(O)₂N(R^N), or $\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{O}$;

each instance of R^N is independently hydrogen, optionally substituted alkyl, or a nitrogen protecting group;

Ring B is optionally substituted carbocyclyl, optionally substituted heterocyclyl, optionally substituted aryl, or optionally substituted heteroaryl; and

5 p is 1 or 2;

provided that the compound is not of the Formula:

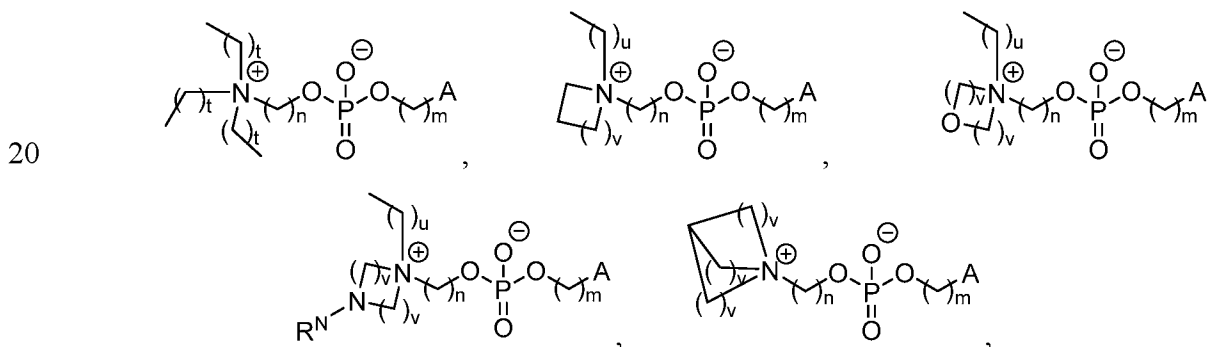


wherein each instance of R^2 is independently unsubstituted alkyl, unsubstituted alkenyl, or unsubstituted alkynyl.

10 In some embodiments, the phospholipids may be one or more of the phospholipids described in U.S. Application No. 62/520,530.

i. Phospholipid Head Modifications

In certain embodiments, a phospholipid useful or potentially useful in the present disclosure comprises a modified phospholipid head (*e.g.*, a modified choline group). In certain embodiments, a phospholipid with a modified head is DSPC, or analog thereof, with a modified quaternary amine. For example, in embodiments of Formula (IV), at least one of R^1 is not methyl. In certain embodiments, at least one of R^1 is not hydrogen or methyl. In certain embodiments, the compound of Formula (IV) is of one of the following Formulae:



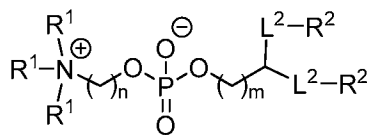
or a salt thereof, wherein:

each t is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

each u is independently 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10; and

25 each v is independently 1, 2, or 3.

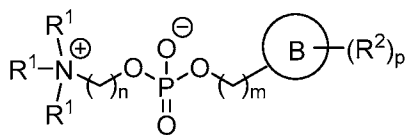
In certain embodiments, a compound of Formula (IV) is of Formula (IV-a):



(IV-a),

or a salt thereof.

5 In certain embodiments, a phospholipid useful or potentially useful in the present disclosure comprises a cyclic moiety in place of the glyceride moiety. In certain embodiments, a phospholipid useful in the present disclosure is DSPC, or analog thereof, with a cyclic moiety in place of the glyceride moiety. In certain embodiments, the compound of Formula (IV) is of Formula (IV-b):



(IV-b),

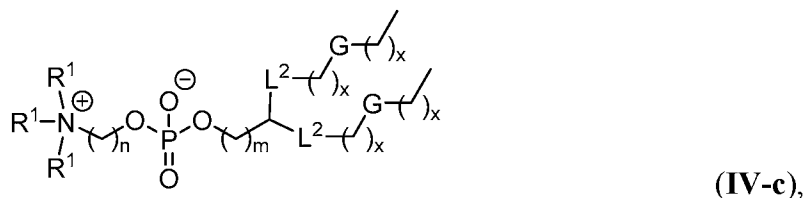
or a salt thereof.

ii. Phospholipid Tail Modifications

15 In certain embodiments, a phospholipid useful or potentially useful in the present disclosure comprises a modified tail. In certain embodiments, a phospholipid useful or potentially useful in the present disclosure is DSPC, or analog thereof, with a modified tail. As described herein, a “modified tail” may be a tail with shorter or longer aliphatic chains, aliphatic chains with branching introduced, aliphatic chains with substituents introduced, aliphatic chains wherein one or more methylenes are replaced by cyclic or heteroatom groups, or any combination thereof. For example, in certain embodiments, the compound of (IV) is of Formula (IV-a), or a salt thereof, wherein at least one instance of R² is each instance of R² is optionally substituted C₁₋₃₀ alkyl, wherein one or more methylene units of R² are independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, N(R^N),

O, S, C(O), C(O)N(R^N), NR^NC(O), NR^NC(O)N(R^N), C(O)O, OC(O), OC(O)O, OC(O)N(R^N), NR^NC(O)O, C(O)S, SC(O), C(=NR^N), C(=NR^N)N(R^N), NR^NC(=NR^N), NR^NC(=NR^N)N(R^N), C(S), C(S)N(R^N), NR^NC(S), NR^NC(S)N(R^N), S(O), OS(O), S(O)O, OS(O)O, OS(O)₂, - S(O)₂O, OS(O)₂O, N(R^N)S(O), S(O)N(R^N), N(R^N)S(O)N(R^N), OS(O)N(R^N), N(R^N)S(O)O, S(O)₂, N(R^N)S(O)₂, S(O)₂N(R^N), N(R^N)S(O)₂N(R^N), OS(O)₂N(R^N), or N(R^N)S(O)₂O.

In certain embodiments, the compound of Formula (IV) is of Formula (IV-c):

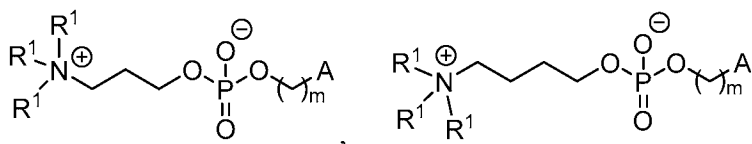


or a salt thereof, wherein:

each x is independently an integer between 0-30, inclusive; and

each instance is G is independently selected from the group consisting of optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, N(R^N), O, S, C(O), C(O)N(R^N), NR^NC(O), - NR^NC(O)N(R^N), C(O)O, OC(O), OC(O)O, OC(O)N(R^N), NR^NC(O)O, C(O)S, SC(O), - C(=NR^N), C(=NR^N)N(R^N), NR^NC(=NR^N), NR^NC(=NR^N)N(R^N), C(S), C(S)N(R^N), NR^NC(S), NR^NC(S)N(R^N), S(O), OS(O), S(O)O, OS(O)O, OS(O)₂, S(O)₂O, OS(O)₂O, N(R^N)S(O), - S(O)N(R^N), N(R^N)S(O)N(R^N), OS(O)N(R^N), N(R^N)S(O)O, S(O)₂, N(R^N)S(O)₂, S(O)₂N(R^N), N(R^N)S(O)₂N(R^N), OS(O)₂N(R^N), or N(R^N)S(O)₂O. Each possibility represents a separate embodiment of the present disclosure.

In certain embodiments, a phospholipid useful or potentially useful in the present disclosure comprises a modified phosphocholine moiety, wherein the alkyl chain linking the quaternary amine to the phosphoryl group is not ethylene (*e.g.*, n is not 2). Therefore, in certain embodiments, a phospholipid useful or potentially useful in the present disclosure is a compound of Formula (IV), wherein n is 1, 3, 4, 5, 6, 7, 8, 9, or 10. For example, in certain embodiments, a compound of Formula (IV) is of one of the following Formulae:



or a salt thereof.

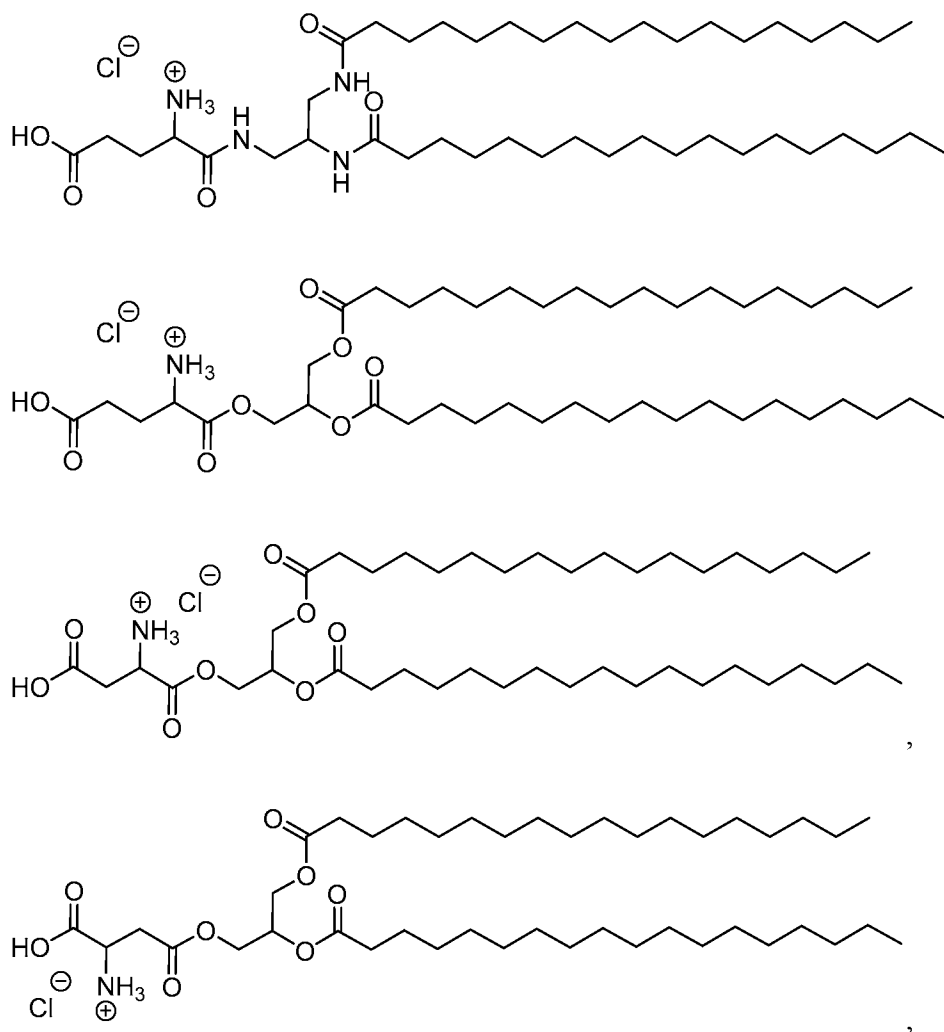
iii. Alternative Lipids

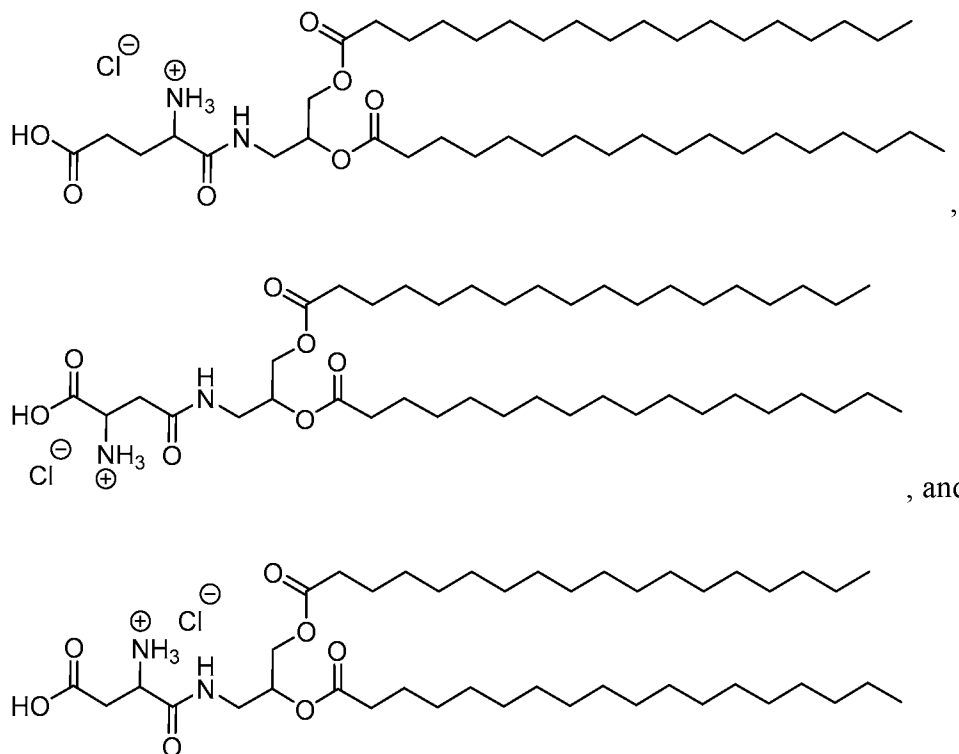
In certain embodiments, a phospholipid useful or potentially useful in the present disclosure comprises a modified phosphocholine moiety, wherein the alkyl chain linking the quaternary amine to the phosphoryl group is not ethylene (*e.g.*, *n* is not 2). Therefore, in certain embodiments, a phospholipid useful.

In certain embodiments, an alternative lipid is used in place of a phospholipid of the present disclosure.

In certain embodiments, an alternative lipid of the present disclosure is oleic acid.

In certain embodiments, the alternative lipid is one of the following:





5

d. Structural Lipids

The lipid composition of a pharmaceutical composition disclosed herein can comprise one or more structural lipids. As used herein, the term "structural lipid" refers to sterols and also to lipids containing sterol moieties.

Incorporation of structural lipids in the lipid nanoparticle may help mitigate aggregation of other lipids in the particle. Structural lipids can be selected from the group including but not limited to, cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, tomatine, ursolic acid, alpha-tocopherol, hopanoids, phytosterols, steroids, and mixtures thereof. In some embodiments, the structural lipid is a sterol. As defined herein, "sterols" are a subgroup of steroids consisting of steroid alcohols.

In certain embodiments, the structural lipid is a steroid. In certain embodiments, the structural lipid is cholesterol. In certain embodiments, the structural lipid is an analog of cholesterol. In certain embodiments, the structural lipid is alpha-tocopherol.

In some embodiments, the structural lipids may be one or more of the structural lipids described in U.S. Application No. 62/520,530.

20

e. Polyethylene Glycol (PEG)-Lipids

The lipid composition of a pharmaceutical composition disclosed herein can comprise one or more a polyethylene glycol (PEG) lipid.

As used herein, the term “PEG-lipid” refers to polyethylene glycol (PEG)-modified lipids. Non-limiting examples of PEG-lipids include PEG-modified phosphatidylethanolamine and phosphatidic acid, PEG-ceramide conjugates (*e.g.*, PEG-CerC14 or PEG-CerC20), PEG-modified dialkylamines and PEG-modified 1,2-diacyloxypropan-3-amines. Such lipids are also referred to as PEGylated lipids. For example, a PEG lipid can be PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE lipid.

In some embodiments, the PEG-lipid includes, but not limited to 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol (PEG-DMG), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)] (PEG-DSPE), PEG-disteryl glycerol (PEG-DSG), PEG-dipalmetoleyl, PEG-dioleoyl, PEG-distearoyl, PEG-diacylglycamide (PEG-DAG), PEG-dipalmitoyl phosphatidylethanolamine (PEG-DPPE), or PEG-1,2-dimyristyloxylpropyl-3-amine (PEG-c-DMA).

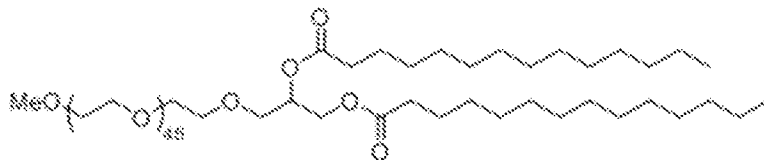
In some embodiments, the PEG-lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof.

In some embodiments, the lipid moiety of the PEG-lipids includes those having lengths of from about C₁₄ to about C₂₂, preferably from about C₁₄ to about C₁₆. In some embodiments, a PEG moiety, for example an mPEG-NH₂, has a size of about 1000, 2000, 5000, 10,000, 15,000 or 20,000 daltons. In some embodiments, the PEG-lipid is PEG_{2k}-DMG.

In some embodiments, the lipid nanoparticles described herein can comprise a PEG lipid which is a non-diffusible PEG. Non-limiting examples of non-diffusible PEGs include PEG-DSG and PEG-DSPE.

In general, some of the other lipid components (*e.g.*, PEG lipids) of various Formulae, described herein may be synthesized as described International Patent Application No. PCT/US2016/000129, filed December 10, 2016, entitled “Compositions and Methods for Delivery of Therapeutic Agents,” which is incorporated by reference in its entirety.

In some embodiments the PEG-modified lipids are a modified form of PEG DMG. PEG-DMG has the following structure:



In certain embodiments, a PEG lipid useful in the present disclosure is a compound of Formula (V). Provided herein are compounds of Formula (V):



or salts thereof, wherein:

R^3 is $-\text{OR}^O$;

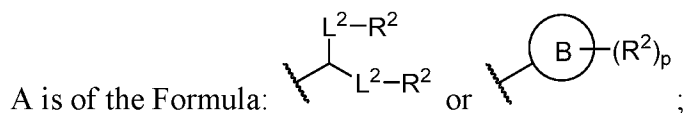
R^O is hydrogen, optionally substituted alkyl, or an oxygen protecting group;

5 r is an integer between 1 and 100, inclusive;

L^1 is optionally substituted C_{1-10} alkylene, wherein at least one methylene of the optionally substituted C_{1-10} alkylene is independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, O, $\text{N}(\text{R}^N)$, S, $\text{C}(\text{O})$, $\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})$, $\text{C}(\text{O})\text{O}$, -
 10 $\text{OC}(\text{O})$, $\text{OC}(\text{O})\text{O}$, $\text{OC}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})\text{O}$, or $\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)$;

D is a moiety obtained by click chemistry or a moiety cleavable under physiological conditions;

m is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;



15 each instance of L^2 is independently a bond or optionally substituted C_{1-6} alkylene, wherein one methylene unit of the optionally substituted C_{1-6} alkylene is optionally replaced with O, $\text{N}(\text{R}^N)$, S, $\text{C}(\text{O})$, $\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})$, $\text{C}(\text{O})\text{O}$, $\text{OC}(\text{O})$, $\text{OC}(\text{O})\text{O}$, $\text{OC}(\text{O})\text{N}(\text{R}^N)$, -
 $\text{NR}^N\text{C}(\text{O})\text{O}$, or $\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)$;

each instance of R^2 is independently optionally substituted C_{1-30} alkyl, optionally
 20 substituted C_{1-30} alkenyl, or optionally substituted C_{1-30} alkynyl; optionally wherein one or more methylene units of R^2 are independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, $\text{N}(\text{R}^N)$, O, S, $\text{C}(\text{O})$, $\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})$, -
 $\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)$, $\text{C}(\text{O})\text{O}$, $\text{OC}(\text{O})$, $\text{OC}(\text{O})\text{O}$, $\text{OC}(\text{O})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{O})\text{O}$, $\text{C}(\text{O})\text{S}$, $\text{SC}(\text{O})$, -
 25 $\text{C}(=\text{NR}^N)$, $\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(=\text{NR}^N)$, $\text{NR}^N\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)$, $\text{C}(\text{S})$, $\text{C}(\text{S})\text{N}(\text{R}^N)$, $\text{NR}^N\text{C}(\text{S})$, $\text{NR}^N\text{C}(\text{S})\text{N}(\text{R}^N)$, $\text{S}(\text{O})$, $\text{OS}(\text{O})$, $\text{S}(\text{O})\text{O}$, $\text{OS}(\text{O})\text{O}$, $\text{OS}(\text{O})_2$, $\text{S}(\text{O})_2\text{O}$, $\text{OS}(\text{O})_2\text{O}$, $\text{N}(\text{R}^N)\text{S}(\text{O})$, -
 $\text{S}(\text{O})\text{N}(\text{R}^N)$, $\text{N}(\text{R}^N)\text{S}(\text{O})\text{N}(\text{R}^N)$, $\text{OS}(\text{O})\text{N}(\text{R}^N)$, $\text{N}(\text{R}^N)\text{S}(\text{O})\text{O}$, $\text{S}(\text{O})_2$, $\text{N}(\text{R}^N)\text{S}(\text{O})_2$, $\text{S}(\text{O})_2\text{N}(\text{R}^N)$, $\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{N}(\text{R}^N)$, $\text{OS}(\text{O})_2\text{N}(\text{R}^N)$, or $\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{O}$;

each instance of R^N is independently hydrogen, optionally substituted alkyl, or a
 30 nitrogen protecting group;

Ring B is optionally substituted carbocyclyl, optionally substituted heterocyclyl, optionally substituted aryl, or optionally substituted heteroaryl; and

p is 1 or 2.

In certain embodiments, the compound of Formula (V) is a PEG-OH lipid (*i.e.*, R³ is –OR^O, and R^O is hydrogen). In certain embodiments, the compound of Formula (V) is of Formula (V-OH):



or a salt thereof.

In certain embodiments, a PEG lipid useful in the present disclosure is a PEGylated fatty acid. In certain embodiments, a PEG lipid useful in the present disclosure is a compound of Formula (VI). Provided herein are compounds of Formula (VI):



or a salts thereof, wherein:

R³ is –OR^O;

R^O is hydrogen, optionally substituted alkyl or an oxygen protecting group;

r is an integer between 1 and 100, inclusive;

R⁵ is optionally substituted C₁₀₋₄₀ alkyl, optionally substituted C₁₀₋₄₀ alkenyl, or optionally substituted C₁₀₋₄₀ alkynyl; and optionally one or more methylene groups of R⁵ are replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, N(R^N), O, S, C(O), -C(O)N(R^N), NR^NC(O), NR^NC(O)N(R^N), C(O)O, OC(O), OC(O)O, OC(O)N(R^N), -NR^NC(O)O, C(O)S, SC(O), C(=NR^N), C(=NR^N)N(R^N), NR^NC(=NR^N), NR^NC(=NR^N)N(R^N), C(S), C(S)N(R^N), NR^NC(S), NR^NC(S)N(R^N), S(O), OS(O), S(O)O, OS(O)O, OS(O)₂, -S(O)₂O, OS(O)₂O, N(R^N)S(O), S(O)N(R^N), N(R^N)S(O)N(R^N), OS(O)N(R^N), N(R^N)S(O)O, S(O)₂, N(R^N)S(O)₂, S(O)₂N(R^N), N(R^N)S(O)₂N(R^N), OS(O)₂N(R^N), or N(R^N)S(O)₂O; and

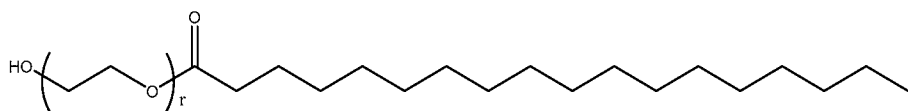
each instance of R^N is independently hydrogen, optionally substituted alkyl, or a nitrogen protecting group.

In certain embodiments, the compound of Formula (VI) is of Formula (VI-OH):



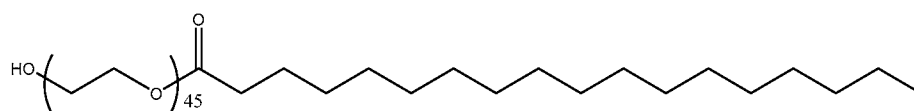
or a salt thereof. In some embodiments, r is 45.

In yet other embodiments the compound of Formula (VI) is:



or a salt thereof. In one embodiment, r is 40-50.

5 In some embodiments, the compound of Formula (VI) is



(Compound P-I).

In some aspects, the lipid composition of the pharmaceutical compositions disclosed herein does not comprise a PEG-lipid.

10 In some embodiments, the PEG-lipids may be one or more of the PEG lipids described in U.S. Application No. 62/520,530.

In some embodiments, a PEG lipid of the present disclosure comprises a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof. In some embodiments, the PEG-modified lipid is
15 PEG-DMG, PEG-c-DOMG (also referred to as PEG-DOMG), PEG-DSG and/or PEG-DPG.

In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of any of Formula I, II or III, a phospholipid comprising DSPC, a structural lipid, and a PEG lipid comprising PEG-DMG.

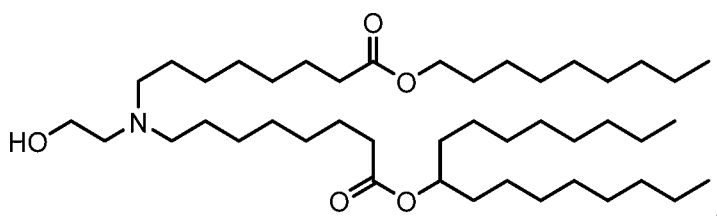
20 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of any of Formula I, II or III, a phospholipid comprising DSPC, a structural lipid, and a PEG lipid comprising a compound having Formula VI.

In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of Formula I, II or III, a phospholipid comprising a compound having Formula IV, a structural lipid, and the PEG lipid comprising a compound having Formula V or VI.

5 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of Formula I, II or III, a phospholipid comprising a compound having Formula IV, a structural lipid, and the PEG lipid comprising a compound having Formula V or VI.

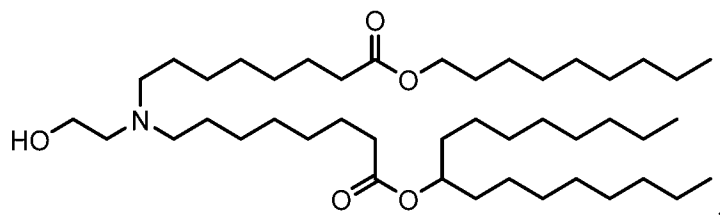
In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of Formula I, II or III, a phospholipid having Formula IV, a structural lipid, and a PEG lipid comprising a compound having Formula VI.

10 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of



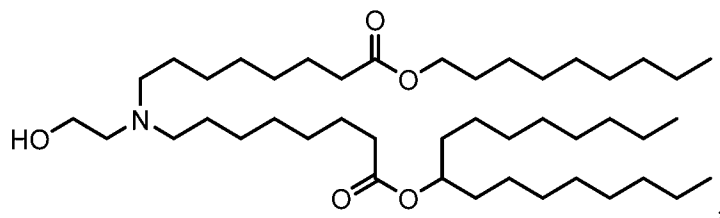
and a PEG lipid comprising Formula VI.

15 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of



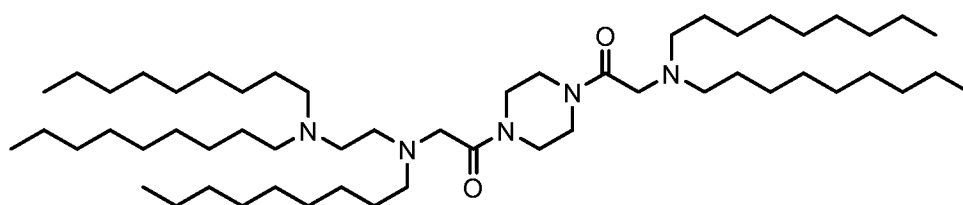
and an alternative lipid comprising oleic acid.

In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of



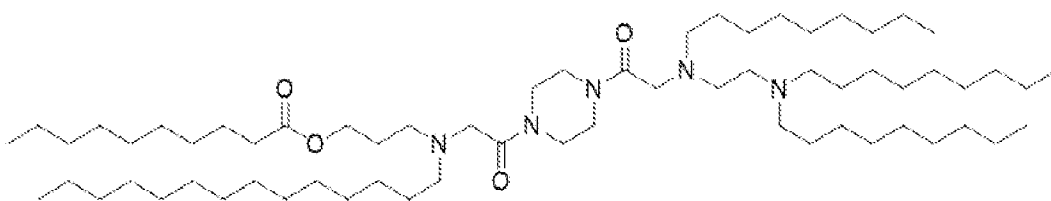
an alternative lipid comprising oleic acid, a structural lipid comprising cholesterol, and a PEG lipid comprising a compound having Formula VI.

5 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of



a phospholipid comprising DOPE, a structural lipid comprising cholesterol, and a PEG lipid comprising a compound having Formula VI.

10 In some embodiments, a LNP of the present disclosure comprises an ionizable cationic lipid of



a phospholipid comprising DOPE, a structural lipid comprising cholesterol, and a PEG lipid comprising a compound having Formula VI.

15 In some embodiments, a LNP of the present disclosure comprises an N:P ratio of from about 2:1 to about 30:1.

In some embodiments, a LNP of the present disclosure comprises an N:P ratio of about 6:1.

In some embodiments, a LNP of the present disclosure comprises an N:P ratio of about 3:1.

In some embodiments, a LNP of the present disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of from about 10:1 to about 100:1.

5 In some embodiments, a LNP of the present disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of about 20:1.

In some embodiments, a LNP of the present disclosure comprises a wt/wt ratio of the ionizable cationic lipid component to the RNA of about 10:1.

10 In some embodiments, a LNP of the present disclosure has a mean diameter from about 50nm to about 150nm.

In some embodiments, a LNP of the present disclosure has a mean diameter from about 70nm to about 120nm.

15 As used herein, the term "alkyl", "alkyl group", or "alkylene" means a linear or branched, saturated hydrocarbon including one or more carbon atoms (*e.g.*, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more carbon atoms), which is optionally substituted. The notation "C₁₋₁₄ alkyl" means an optionally substituted linear or branched, saturated hydrocarbon including 1-14 carbon atoms. Unless otherwise specified, an alkyl group described herein refers to both unsubstituted and substituted alkyl groups.

20 As used herein, the term "alkenyl", "alkenyl group", or "alkenylene" means a linear or branched hydrocarbon including two or more carbon atoms (*e.g.*, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more carbon atoms) and at least one double bond, which is optionally substituted. The notation "C₂₋₁₄ alkenyl" means an optionally substituted linear or branched
25 hydrocarbon including 2-14 carbon atoms and at least one carbon-carbon double bond. An alkenyl group may include one, two, three, four, or more carbon-carbon double bonds. For example, C₁₈ alkenyl may include one or more double bonds. A C₁₈ alkenyl group including two double bonds may be a linoleyl group. Unless otherwise specified, an alkenyl group described herein refers to both unsubstituted and substituted alkenyl groups.

As used herein, the term "alkynyl", "alkynyl group", or "alkynylene" means a linear or branched hydrocarbon including two or more carbon atoms (*e.g.*, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or more carbon atoms) and at least one carbon-carbon triple bond, which is optionally substituted. The notation "C₂₋₁₄ alkynyl" means an optionally substituted linear or branched hydrocarbon including 2-14 carbon atoms and at least one carbon-carbon triple bond. An alkynyl group may include one, two, three, four, or more carbon-carbon triple bonds. For example, C₁₈ alkynyl may include one or more carbon-carbon triple bonds. Unless otherwise specified, an alkynyl group described herein refers to both unsubstituted and substituted alkynyl groups.

As used herein, the term "carbocycle" or "carbocyclic group" means an optionally substituted mono- or multi-cyclic system including one or more rings of carbon atoms. Rings may be three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty membered rings. The notation "C₃₋₆ carbocycle" means a carbocycle including a single ring having 3-6 carbon atoms. Carbocycles may include one or more carbon-carbon double or triple bonds and may be non-aromatic or aromatic (*e.g.*, cycloalkyl or aryl groups). Examples of carbocycles include cyclopropyl, cyclopentyl, cyclohexyl, phenyl, naphthyl, and 1,2 dihydronaphthyl groups. The term "cycloalkyl" as used herein means a non-aromatic carbocycle and may or may not include any double or triple bond. Unless otherwise specified, carbocycles described herein refers to both unsubstituted and substituted carbocycle groups, *i.e.*, optionally substituted carbocycles.

As used herein, the term "heterocycle" or "heterocyclic group" means an optionally substituted mono- or multi-cyclic system including one or more rings, where at least one ring includes at least one heteroatom. Heteroatoms may be, for example, nitrogen, oxygen, or sulfur atoms. Rings may be three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, or fourteen membered rings. Heterocycles may include one or more double or triple bonds and may be non-aromatic or aromatic (*e.g.*, heterocycloalkyl or heteroaryl groups). Examples of heterocycles include imidazolyl, imidazolidinyl, oxazolyl, oxazolidinyl, thiazolyl, thiazolidinyl, pyrazolidinyl, pyrazolyl, isoxazolidinyl, isoxazolyl, isothiazolidinyl, isothiazolyl, morpholinyl, pyrrolyl, pyrrolidinyl, furyl, tetrahydrofuryl, thiophenyl, pyridinyl, piperidinyl, quinolyl, and isoquinolyl groups. The term "heterocycloalkyl" as used herein

means a non-aromatic heterocycle and may or may not include any double or triple bond. Unless otherwise specified, heterocycles described herein refers to both unsubstituted and substituted heterocycle groups, *i.e.*, optionally substituted heterocycles.

As used herein, the term "heteroalkyl", "heteroalkenyl", or "heteroalkynyl", refers respectively to an alkyl, alkenyl, alkynyl group, as defined herein, which further comprises one or more (*e.g.*, 1, 2, 3, or 4) heteroatoms (*e.g.*, oxygen, sulfur, nitrogen, boron, silicon, phosphorus) wherein the one or more heteroatoms is inserted between adjacent carbon atoms within the parent carbon chain and/or one or more heteroatoms is inserted between a carbon atom and the parent molecule, *i.e.*, between the point of attachment. Unless otherwise specified, heteroalkyls, heteroalkenyls, or heteroalkynyls described herein refers to both unsubstituted and substituted heteroalkyls, heteroalkenyls, or heteroalkynyls, *i.e.*, optionally substituted heteroalkyls, heteroalkenyls, or heteroalkynyls.

As used herein, a "biodegradable group" is a group that may facilitate faster metabolism of a lipid in a mammalian entity. A biodegradable group may be selected from the group consisting of, but is not limited to, -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, an aryl group, and a heteroaryl group. As used herein, an "aryl group" is an optionally substituted carbocyclic group including one or more aromatic rings. Examples of aryl groups include phenyl and naphthyl groups. As used herein, a "heteroaryl group" is an optionally substituted heterocyclic group including one or more aromatic rings. Examples of heteroaryl groups include pyrrolyl, furyl, thiophenyl, imidazolyl, oxazolyl, and thiazolyl. Both aryl and heteroaryl groups may be optionally substituted. For example, M and M' can be selected from the non-limiting group consisting of optionally substituted phenyl, oxazole, and thiazole. In the Formulas herein, M and M' can be independently selected from the list of biodegradable groups above. Unless otherwise specified, aryl or heteroaryl groups described herein refers to both unsubstituted and substituted groups, *i.e.*, optionally substituted aryl or heteroaryl groups.

Alkyl, alkenyl, and cyclyl (*e.g.*, carbocyclyl and heterocyclyl) groups may be optionally substituted unless otherwise specified. Optional substituents may be selected from the group consisting of, but are not limited to, a halogen atom (*e.g.*, a chloride, bromide, fluoride, or iodide group), a carboxylic acid (*e.g.*, C(O)OH), an alcohol (*e.g.*, a hydroxyl, OH), an ester (*e.g.*, C(O)OR OC(O)R), an aldehyde (*e.g.*, C(O)H), a carbonyl (*e.g.*, C(O)R,

alternatively represented by C=O), an acyl halide (*e.g.*, C(O)X, in which X is a halide selected from bromide, fluoride, chloride, and iodide), a carbonate (*e.g.*, OC(O)OR), an alkoxy (*e.g.*, OR), an acetal (*e.g.*, C(OR)₂R^{'''}), in which each OR are alkoxy groups that can be the same or different and R^{'''} is an alkyl or alkenyl group), a phosphate (*e.g.*, P(O)₄³⁻), a thiol (*e.g.*, SH), a sulfoxide (*e.g.*, S(O)R), a sulfinic acid (*e.g.*, S(O)OH), a sulfonic acid (*e.g.*, S(O)₂OH), a thial (*e.g.*, C(S)H), a sulfate (*e.g.*, S(O)₄²⁻), a sulfonyl (*e.g.*, S(O)₂), an amide (*e.g.*, C(O)NR₂, or N(R)C(O)R), an azido (*e.g.*, N₃), a nitro (*e.g.*, NO₂), a cyano (*e.g.*, CN), an isocyano (*e.g.*, NC), an acyloxy (*e.g.*, OC(O)R), an amino (*e.g.*, NR₂, NRH, or NH₂), a carbamoyl (*e.g.*, OC(O)NR₂, OC(O)NRH, or OC(O)NH₂), a sulfonamide (*e.g.*, S(O)₂NR₂, S(O)₂NRH, S(O)₂NH₂, N(R)S(O)₂R, N(H)S(O)₂R, N(R)S(O)₂H, or N(H)S(O)₂H), an alkyl group, an alkenyl group, and a cyclyl (*e.g.*, carbocyclyl or heterocyclyl) group. In any of the preceding, R is an alkyl or alkenyl group, as defined herein. In some embodiments, the substituent groups themselves may be further substituted with, for example, one, two, three, four, five, or six substituents as defined herein. For example, a C₁₋₆ alkyl group may be further substituted with one, two, three, four, five, or six substituents as described herein.

Compounds of the disclosure that contain nitrogens can be converted to N-oxides by treatment with an oxidizing agent (*e.g.*, 3-chloroperoxybenzoic acid (mCPBA) and/or hydrogen peroxides) to afford other compounds of the disclosure. Thus, all shown and claimed nitrogen-containing compounds are considered, when allowed by valency and structure, to include both the compound as shown and its N-oxide derivative (which can be designated as N→O or N⁺-O⁻). Furthermore, in other instances, the nitrogens in the compounds of the disclosure can be converted to N-hydroxy or N-alkoxy compounds. For example, N-hydroxy compounds can be prepared by oxidation of the parent amine by an oxidizing agent such as mCPBA. All shown and claimed nitrogen-containing compounds are also considered, when allowed by valency and structure, to cover both the compound as shown and its N-hydroxy (*i.e.*, N-OH) and N-alkoxy (*i.e.*, N-OR, wherein R is substituted or unsubstituted C₁-C₆ alkyl, C₁-C₆ alkenyl, C₁-C₆ alkynyl, 3-14-membered carbocycle or 3-14-membered heterocycle) derivatives.

(vi) Other Lipid Composition Components

The lipid composition of a pharmaceutical composition disclosed herein can include one or more components in addition to those described above. For example, the lipid

composition can include one or more permeability enhancer molecules, carbohydrates, polymers, surface altering agents (*e.g.*, surfactants), or other components. For example, a permeability enhancer molecule can be a molecule described by U.S. Patent Application Publication No. 2005/0222064. Carbohydrates can include simple sugars (*e.g.*, glucose) and polysaccharides (*e.g.*, glycogen and derivatives and analogs thereof).

A polymer can be included in and/or used to encapsulate or partially encapsulate a pharmaceutical composition disclosed herein (*e.g.*, a pharmaceutical composition in lipid nanoparticle form). A polymer can be biodegradable and/or biocompatible. A polymer can be selected from, but is not limited to, polyamines, polyethers, polyamides, polyesters, polycarbamates, polyureas, polycarbonates, polystyrenes, polyimides, polysulfones, polyurethanes, polyacetylenes, polyethylenes, polyethyleneimines, polyisocyanates, polyacrylates, polymethacrylates, polyacrylonitriles, and polyarylates.

The ratio between the lipid composition and the polynucleotide range can be from about 10:1 to about 60:1 (wt/wt).

In some embodiments, the ratio between the lipid composition and the polynucleotide can be about 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, 50:1, 51:1, 52:1, 53:1, 54:1, 55:1, 56:1, 57:1, 58:1, 59:1 or 60:1 (wt/wt). In some embodiments, the wt/wt ratio of the lipid composition to the polynucleotide encoding a therapeutic agent is about 20:1 or about 15:1.

In some embodiments, the pharmaceutical composition disclosed herein can contain more than one polypeptides. For example, a pharmaceutical composition disclosed herein can contain two or more polynucleotides (*e.g.*, RNA, *e.g.*, mRNA).

In some embodiments, the lipid nanoparticles described herein can comprise polynucleotides (*e.g.*, mRNA) in a lipid:polynucleotide weight ratio of 5:1, 10:1, 15:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, 55:1, 60:1 or 70:1, or a range or any of these ratios such as, but not limited to, 5:1 to about 10:1, from about 5:1 to about 15:1, from about 5:1 to about 20:1, from about 5:1 to about 25:1, from about 5:1 to about 30:1, from about 5:1 to about 35:1, from about 5:1 to about 40:1, from about 5:1 to about 45:1, from about 5:1 to about 50:1, from about 5:1 to about 55:1, from about 5:1 to about 60:1, from about 5:1 to about 70:1, from about 10:1 to about 15:1, from about 10:1 to about 20:1, from about 10:1 to about

25:1, from about 10:1 to about 30:1, from about 10:1 to about 35:1, from about 10:1 to about 40:1, from about 10:1 to about 45:1, from about 10:1 to about 50:1, from about 10:1 to about 55:1, from about 10:1 to about 60:1, from about 10:1 to about 70:1, from about 15:1 to about 20:1, from about 15:1 to about 25:1, from about 15:1 to about 30:1, from about 15:1 to about 35:1, from about 15:1 to about 40:1, from about 15:1 to about 45:1, from about 15:1 to about 50:1, from about 15:1 to about 55:1, from about 15:1 to about 60:1 or from about 15:1 to about 70:1.

In some embodiments, the lipid nanoparticles described herein can comprise the polynucleotide in a concentration from approximately 0.1 mg/ml to 2 mg/ml such as, but not limited to, 0.1 mg/ml, 0.2 mg/ml, 0.3 mg/ml, 0.4 mg/ml, 0.5 mg/ml, 0.6 mg/ml, 0.7 mg/ml, 0.8 mg/ml, 0.9 mg/ml, 1.0 mg/ml, 1.1 mg/ml, 1.2 mg/ml, 1.3 mg/ml, 1.4 mg/ml, 1.5 mg/ml, 1.6 mg/ml, 1.7 mg/ml, 1.8 mg/ml, 1.9 mg/ml, 2.0 mg/ml or greater than 2.0 mg/ml.

f. Nanoparticle Compositions

In some embodiments, the pharmaceutical compositions disclosed herein are Formulated as lipid nanoparticles (LNP). Accordingly, the present disclosure also provides nanoparticle compositions comprising (i) a lipid composition comprising a delivery agent such as compound as described herein, and (ii) a polynucleotide encoding an ASL polypeptide. In such nanoparticle composition, the lipid composition disclosed herein can encapsulate the polynucleotide encoding an ASL polypeptide.

Nanoparticle compositions are typically sized on the order of micrometers or smaller and can include a lipid bilayer. Nanoparticle compositions encompass lipid nanoparticles (LNPs), liposomes (*e.g.*, lipid vesicles), and lipoplexes. For example, a nanoparticle composition can be a liposome having a lipid bilayer with a diameter of 500 nm or less.

Nanoparticle compositions include, for example, lipid nanoparticles (LNPs), liposomes, and lipoplexes. In some embodiments, nanoparticle compositions are vesicles including one or more lipid bilayers. In certain embodiments, a nanoparticle composition includes two or more concentric bilayers separated by aqueous compartments. Lipid bilayers can be functionalized and/or crosslinked to one another. Lipid bilayers can include one or more ligands, proteins, or channels.

In some embodiments, a lipid nanoparticle comprises an ionizable amino lipid, a structural lipid, a phospholipid, and mRNA. In some embodiments, the LNP comprises an ionizable amino lipid, a PEG-modified lipid, a sterol and a structural lipid. In some
 5 structural lipid; about 30-45% sterol; and about 1-5% PEG-modified lipid.

In some embodiments, the lipid nanoparticle comprises 47-49 mol.% ionizable cationic lipid (*e.g.* ionizable amino lipid, *e.g.*, Compound I-1, Compound I-2, or Compound I-3), 10-12 mol.% non-cationic lipid (*e.g.*, phospholipid, *e.g.*, DSPC), 38-40 mol.% sterol (*e.g.*, cholesterol) or other structural lipid, and 1-3 mol.% PEG-modified lipid (*e.g.*, PEG-
 10 DMG or Compound P-I).

For instance, in some embodiments, the lipid nanoparticle (“LNP-1”) may comprise the following components at the following molar ratios:

- (i) 45-50 mol.% Compound I-1
- (ii) 35-45 mol.% sterol (*e.g.*, cholesterol);
- 15 (iii) 8-12 mol.% phospholipid (*e.g.*, DSPC or DOPE); and
- (iv) 1.5-3.5 mol.% PEG-lipid (*e.g.*, Compound P-I or PEG-DMG).

For instance, in some embodiments, the lipid nanoparticle (“LNP-1A”) may comprise the following components at the following molar ratios:

- (i) 45-50 mol.% Compound I-1
- 20 (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% PEG-DMG.

For instance, in some embodiments, the lipid nanoparticle (“LNP-1B”) may comprise the following components at the following molar ratios:

- 25 (i) 45-50 mol.% Compound I-1
- (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% Compound P-I.

In some embodiments, the lipid nanoparticle (“LNP-2”) may comprise the following:

- 30 (i) 45-50 mol.% Compound I-2;
- (ii) 35-45 mol.% sterol (*e.g.*, Cholesterol);

- (iii) 8-12 mol.% phospholipid (*e.g.*, DSPC or DOPE); and
- (iv) 1.5-3.5 mol.% PEG-lipid (*e.g.*, Compound P-I or PEG-DMG).

In some embodiments, the lipid nanoparticle (“LNP-2A”) may comprise the following:

- (i) 45-50 mol.% Compound I-2;
- 5 (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% PEG-DMG.

For instance, in some embodiments, the lipid nanoparticle (“LNP-2B”) may comprise the following components at the following molar ratios:

- 10 (i) 45-50 mol.% Compound I-2;
- (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% Compound P-I.

In some embodiments, the lipid nanoparticle (“LNP-3”) may comprise the following:

- 15 (i) 45-50 mol.% Compound I-3;
- (ii) 35-45 mol.% sterol (*e.g.*, Cholesterol);
- (iii) 8-12 mol.% phospholipid (*e.g.*, DSPC or DOPE); and
- (iv) 1.5-3.5 mol.% PEG-lipid (*e.g.*, Compound P-I or PEG-DMG).

In some embodiments, the lipid nanoparticle (“LNP-3A”) may comprise the following:

- 20 (i) 45-50 mol.% Compound I-3;
- (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% PEG-DMG.

In some embodiments, the lipid nanoparticle (“LNP-3B”) may comprise the following:

- 25 (i) 45-50 mol.% Compound I-3;
- (ii) 35-45 mol.% Cholesterol;
- (iii) 8-12 mol.% DSPC; and
- (iv) 1.5-3.5 mol.% Compound P-I.

In some embodiments, the LNP has a polydispersity value of less than 0.4. In some
30 embodiments, the LNP has a net neutral charge at a neutral pH. In some embodiments, the

LNP has a mean diameter of 50-150 nm. In some embodiments, the LNP has a mean diameter of 80-100 nm.

As generally defined herein, the term “lipid” refers to a small molecule that has hydrophobic or amphiphilic properties. Lipids may be naturally occurring or synthetic.

5 Examples of classes of lipids include, but are not limited to, fats, waxes, sterol-containing metabolites, vitamins, fatty acids, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, and polyketides, and prenol lipids. In some instances, the amphiphilic properties of some lipids leads them to form liposomes, vesicles, or membranes in aqueous media.

10 In some embodiments, a lipid nanoparticle (LNP) may comprise an ionizable amino lipid. As used herein, the term “ionizable amino lipid” has its ordinary meaning in the art and may refer to a lipid comprising one or more charged moieties. In some embodiments, an ionizable amino lipid may be positively charged or negatively charged. An ionizable amino lipid may be positively charged, in which case it can be referred to as “cationic lipid”. In
15 certain embodiments, an ionizable amino lipid molecule may comprise an amine group, and can be referred to as an ionizable amino lipid. As used herein, a “charged moiety” is a chemical moiety that carries a formal electronic charge, *e.g.*, monovalent (+1, or -1), divalent (+2, or -2), trivalent (+3, or -3), etc. The charged moiety may be anionic (*i.e.*, negatively charged) or cationic (*i.e.*, positively charged). Examples of positively-charged moieties
20 include amine groups (*e.g.*, primary, secondary, and/or tertiary amines), ammonium groups, pyridinium group, guanidine groups, and imidazolium groups. In a particular embodiment, the charged moieties comprise amine groups. Examples of negatively- charged groups or precursors thereof, include carboxylate groups, sulfonate groups, sulfate groups, phosphonate groups, phosphate groups, hydroxyl groups, and the like. The charge of the charged moiety
25 may vary, in some cases, with the environmental conditions, for example, changes in pH may alter the charge of the moiety, and/or cause the moiety to become charged or uncharged. In general, the charge density of the molecule may be selected as desired.

It should be understood that the terms “charged” or “charged moiety” does not refer to a “partial negative charge” or “partial positive charge” on a molecule. The terms “partial
30 negative charge” and “partial positive charge” are given its ordinary meaning in the art. A “partial negative charge” may result when a functional group comprises a bond that becomes polarized such that electron density is pulled toward one atom of the bond, creating a partial

negative charge on the atom. Those of ordinary skill in the art will, in general, recognize bonds that can become polarized in this way.

The ionizable amino lipid is sometimes referred to in the art as an “ionizable cationic lipid”. In some embodiments, the ionizable amino lipid may have a positively charged
5 hydrophilic head and a hydrophobic tail that are connected via a linker structure.

In addition to these, an ionizable amino lipid may also be a lipid including a cyclic amine group.

In some embodiments, the ionizable amino lipid may be selected from, but not limited to, an ionizable amino lipid described in International Publication Nos. WO2013086354 and
10 WO2013116126; the contents of each of which are herein incorporated by reference in their entirety.

In yet another embodiment, the ionizable amino lipid may be selected from, but not limited to, Formula CLI-CLXXXXII of US Patent No. 7,404,969; each of which is herein incorporated by reference in their entirety.

15 In some embodiments, the lipid may be a cleavable lipid such as those described in International Publication No. WO2012170889, herein incorporated by reference in its entirety. In some embodiments, the lipid may be synthesized by methods known in the art and/or as described in International Publication Nos. WO2013086354; the contents of each of which are herein incorporated by reference in their entirety.

20 Nanoparticle compositions can be characterized by a variety of methods. For example, microscopy (*e.g.*, transmission electron microscopy or scanning electron microscopy) can be used to examine the morphology and size distribution of a nanoparticle composition. Dynamic light scattering or potentiometry (*e.g.*, potentiometric titrations) can be used to measure zeta potentials. Dynamic light scattering can also be utilized to determine
25 particle sizes. Instruments such as the Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK) can also be used to measure multiple characteristics of a nanoparticle composition, such as particle size, polydispersity index, and zeta potential.

The size of the nanoparticles can help counter biological reactions such as, but not limited to, inflammation, or can increase the biological effect of the polynucleotide.

As used herein, “size” or “mean size” in the context of nanoparticle compositions refers to the mean diameter of a nanoparticle composition.

In some embodiments, the polynucleotide encoding an ASL polypeptide are Formulated in lipid nanoparticles having a diameter from about 10 to about 100 nm such as, but not limited to, about 10 to about 20 nm, about 10 to about 30 nm, about 10 to about 40 nm, about 10 to about 50 nm, about 10 to about 60 nm, about 10 to about 70 nm, about 10 to about 80 nm, about 10 to about 90 nm, about 20 to about 30 nm, about 20 to about 40 nm, about 20 to about 50 nm, about 20 to about 60 nm, about 20 to about 70 nm, about 20 to about 80 nm, about 20 to about 90 nm, about 20 to about 100 nm, about 30 to about 40 nm, about 30 to about 50 nm, about 30 to about 60 nm, about 30 to about 70 nm, about 30 to about 80 nm, about 30 to about 90 nm, about 30 to about 100 nm, about 40 to about 50 nm, about 40 to about 60 nm, about 40 to about 70 nm, about 40 to about 80 nm, about 40 to about 90 nm, about 40 to about 100 nm, about 50 to about 60 nm, about 50 to about 70 nm, about 50 to about 80 nm, about 50 to about 90 nm, about 50 to about 100 nm, about 60 to about 70 nm, about 60 to about 80 nm, about 60 to about 90 nm, about 60 to about 100 nm, about 70 to about 80 nm, about 70 to about 90 nm, about 70 to about 100 nm, about 80 to about 90 nm, about 80 to about 100 nm and/or about 90 to about 100 nm.

In some embodiments, the nanoparticles have a diameter from about 10 to 500 nm. In some embodiments, the nanoparticle has a diameter greater than 100 nm, greater than 150 nm, greater than 200 nm, greater than 250 nm, greater than 300 nm, greater than 350 nm, greater than 400 nm, greater than 450 nm, greater than 500 nm, greater than 550 nm, greater than 600 nm, greater than 650 nm, greater than 700 nm, greater than 750 nm, greater than 800 nm, greater than 850 nm, greater than 900 nm, greater than 950 nm or greater than 1000 nm.

In some embodiments, the largest dimension of a nanoparticle composition is 1 μm or shorter (*e.g.*, 1 μm , 900 nm, 800 nm, 700 nm, 600 nm, 500 nm, 400 nm, 300 nm, 200 nm, 175 nm, 150 nm, 125 nm, 100 nm, 75 nm, 50 nm, or shorter).

A nanoparticle composition can be relatively homogenous. A polydispersity index can be used to indicate the homogeneity of a nanoparticle composition, *e.g.*, the particle size distribution of the nanoparticle composition. A small (*e.g.*, less than 0.3) polydispersity index generally indicates a narrow particle size distribution. A nanoparticle composition can have a polydispersity index from about 0 to about 0.25, such as 0.01, 0.02, 0.03, 0.04, 0.05,

0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.20, 0.21, 0.22, 0.23, 0.24, or 0.25. In some embodiments, the polydispersity index of a nanoparticle composition disclosed herein can be from about 0.10 to about 0.20.

The zeta potential of a nanoparticle composition can be used to indicate the electrokinetic potential of the composition. For example, the zeta potential can describe the surface charge of a nanoparticle composition. Nanoparticle compositions with relatively low charges, positive or negative, are generally desirable, as more highly charged species can interact undesirably with cells, tissues, and other elements in the body. In some embodiments, the zeta potential of a nanoparticle composition disclosed herein can be from about -10 mV to about +20 mV, from about -10 mV to about +15 mV, from about -10 mV to about +10 mV, from about -10 mV to about +5 mV, from about -10 mV to about 0 mV, from about -10 mV to about -5 mV, from about -5 mV to about +20 mV, from about -5 mV to about +15 mV, from about -5 mV to about +10 mV, from about -5 mV to about +5 mV, from about -5 mV to about 0 mV, from about 0 mV to about +20 mV, from about 0 mV to about +15 mV, from about 0 mV to about +10 mV, from about 0 mV to about +5 mV, from about +5 mV to about +20 mV, from about +5 mV to about +15 mV, or from about +5 mV to about +10 mV.

In some embodiments, the zeta potential of the lipid nanoparticles can be from about 0 mV to about 100 mV, from about 0 mV to about 90 mV, from about 0 mV to about 80 mV, from about 0 mV to about 70 mV, from about 0 mV to about 60 mV, from about 0 mV to about 50 mV, from about 0 mV to about 40 mV, from about 0 mV to about 30 mV, from about 0 mV to about 20 mV, from about 0 mV to about 10 mV, from about 10 mV to about 100 mV, from about 10 mV to about 90 mV, from about 10 mV to about 80 mV, from about 10 mV to about 70 mV, from about 10 mV to about 60 mV, from about 10 mV to about 50 mV, from about 10 mV to about 40 mV, from about 10 mV to about 30 mV, from about 10 mV to about 20 mV, from about 20 mV to about 100 mV, from about 20 mV to about 90 mV, from about 20 mV to about 80 mV, from about 20 mV to about 70 mV, from about 20 mV to about 60 mV, from about 20 mV to about 50 mV, from about 20 mV to about 40 mV, from about 20 mV to about 30 mV, from about 30 mV to about 100 mV, from about 30 mV to about 90 mV, from about 30 mV to about 80 mV, from about 30 mV to about 70 mV, from about 30 mV to about 60 mV, from about 30 mV to about 50 mV, from about 30 mV to about 40 mV, from about 40 mV to about 100 mV, from about 40 mV to about 90 mV, from about

40 mV to about 80 mV, from about 40 mV to about 70 mV, from about 40 mV to about 60 mV, and from about 40 mV to about 50 mV. In some embodiments, the zeta potential of the lipid nanoparticles can be from about 10 mV to about 50 mV, from about 15 mV to about 45 mV, from about 20 mV to about 40 mV, and from about 25 mV to about 35 mV. In some
5 embodiments, the zeta potential of the lipid nanoparticles can be about 10 mV, about 20 mV, about 30 mV, about 40 mV, about 50 mV, about 60 mV, about 70 mV, about 80 mV, about 90 mV, and about 100 mV.

The term “encapsulation efficiency” of a polynucleotide describes the amount of the polynucleotide that is encapsulated by or otherwise associated with a nanoparticle
10 composition after preparation, relative to the initial amount provided. As used herein, “encapsulation” can refer to complete, substantial, or partial enclosure, confinement, surrounding, or encasement.

Encapsulation efficiency is desirably high (*e.g.*, close to 100%). The encapsulation efficiency can be measured, for example, by comparing the amount of the polynucleotide in a
15 solution containing the nanoparticle composition before and after breaking up the nanoparticle composition with one or more organic solvents or detergents.

Fluorescence can be used to measure the amount of free polynucleotide in a solution. For the nanoparticle compositions described herein, the encapsulation efficiency of a polynucleotide can be at least 50%, for example 50%, 55%, 60%, 65%, 70%, 75%, 80%,
20 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the encapsulation efficiency can be at least 80%. In certain embodiments, the encapsulation efficiency can be at least 90%.

The amount of a polynucleotide present in a pharmaceutical composition disclosed herein can depend on multiple factors such as the size of the polynucleotide, desired target
25 and/or application, or other properties of the nanoparticle composition as well as on the properties of the polynucleotide.

For example, the amount of an mRNA useful in a nanoparticle composition can depend on the size (expressed as length, or molecular mass), sequence, and other characteristics of the mRNA. The relative amounts of a polynucleotide in a nanoparticle
30 composition can also vary.

The relative amounts of the lipid composition and the polynucleotide present in a lipid nanoparticle composition of the present disclosure can be optimized according to considerations of efficacy and tolerability. For compositions including an mRNA as a polynucleotide, the N:P ratio can serve as a useful metric.

5 As the N:P ratio of a nanoparticle composition controls both expression and tolerability, nanoparticle compositions with low N:P ratios and strong expression are desirable. N:P ratios vary according to the ratio of lipids to RNA in a nanoparticle composition.

10 In general, a lower N:P ratio is preferred. The one or more RNA, lipids, and amounts thereof can be selected to provide an N:P ratio from about 2:1 to about 30:1, such as 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 12:1, 14:1, 16:1, 18:1, 20:1, 22:1, 24:1, 26:1, 28:1, or 30:1. In certain embodiments, the N:P ratio can be from about 2:1 to about 8:1. In other embodiments, the N:P ratio is from about 5:1 to about 8:1. In certain embodiments, the N:P ratio is between 5:1 and 6:1. In one specific aspect, the N:P ratio is about 5.67:1.

15 In addition to providing nanoparticle compositions, the present disclosure also provides methods of producing lipid nanoparticles comprising encapsulating a polynucleotide. Such method comprises using any of the pharmaceutical compositions disclosed herein and producing lipid nanoparticles in accordance with methods of production of lipid nanoparticles known in the art. *See, e.g.*, Wang et al. (2015) "Delivery of
20 oligonucleotides with lipid nanoparticles" Adv. Drug Deliv. Rev. 87:68-80; Silva et al. (2015) "Delivery Systems for Biopharmaceuticals. Part I: Nanoparticles and Microparticles" Curr. Pharm. Technol. 16: 940-954; Naseri et al. (2015) "Solid Lipid Nanoparticles and Nanostructured Lipid Carriers: Structure, Preparation and Application" Adv. Pharm. Bull. 5:305-13; Silva et al. (2015) "Lipid nanoparticles for the delivery of biopharmaceuticals"
25 Curr. Pharm. Biotechnol. 16:291-302, and references cited therein.

In some embodiments, the LNP formulations described herein can additionally comprise a permeability enhancer molecule. Non-limiting permeability enhancer molecules are described in U.S. Pub. No. US20050222064, herein incorporated by reference in its entirety.

30 The LNP formulations can further contain a phosphate conjugate. The phosphate conjugate can increase *in vivo* circulation times and/or increase the targeted delivery of the

nanoparticle. Phosphate conjugates can be made by the methods described in, *e.g.*, Intl. Pub. No. WO2013033438 or U.S. Pub. No. US20130196948. The LNP formulation can also contain a polymer conjugate (*e.g.*, a water soluble conjugate) as described in, *e.g.*, U.S. Pub. Nos. US20130059360, US20130196948, and US20130072709. Each of the references is
5 herein incorporated by reference in its entirety.

The LNP formulations can comprise a conjugate to enhance the delivery of nanoparticles of the present disclosure in a subject. Further, the conjugate can inhibit phagocytic clearance of the nanoparticles in a subject. In some embodiments, the conjugate can be a "self" peptide designed from the human membrane protein CD47 (*e.g.*, the "self" particles described by Rodriguez et al, *Science* 2013 339, 971-975, herein incorporated by
10 reference in its entirety). As shown by Rodriguez et al., the self peptides delayed macrophage-mediated clearance of nanoparticles which enhanced delivery of the nanoparticles.

The LNP formulations can comprise a carbohydrate carrier. As a non-limiting
15 example, the carbohydrate carrier can include, but is not limited to, an anhydride-modified phytoglycogen or glycogen-type material, phytoglycogen octenyl succinate, phytoglycogen beta-dextrin, anhydride-modified phytoglycogen beta-dextrin (*e.g.*, Intl. Pub. No. WO2012109121, herein incorporated by reference in its entirety).

The LNP formulations can be coated with a surfactant or polymer to improve the
20 delivery of the particle. In some embodiments, the LNP can be coated with a hydrophilic coating such as, but not limited to, PEG coatings and/or coatings that have a neutral surface charge as described in U.S. Pub. No. US20130183244, herein incorporated by reference in its entirety.

The LNP formulations can be engineered to alter the surface properties of particles so
25 that the lipid nanoparticles can penetrate the mucosal barrier as described in U.S. Pat. No. 8,241,670 or Intl. Pub. No. WO2013110028, each of which is herein incorporated by reference in its entirety.

The LNP engineered to penetrate mucus can comprise a polymeric material (*i.e.*, a polymeric core) and/or a polymer-vitamin conjugate and/or a tri-block co-polymer. The
30 polymeric material can include, but is not limited to, polyamines, polyethers, polyamides, polyesters, polycarbamates, polyureas, polycarbonates, poly(styrenes), polyimides,

polysulfones, polyurethanes, polyacetylenes, polyethylenes, polyethyleneimines, polyisocyanates, polyacrylates, polymethacrylates, polyacrylonitriles, and polyarylates.

LNP engineered to penetrate mucus can also include surface altering agents such as, but not limited to, polynucleotides, anionic proteins (*e.g.*, bovine serum albumin), surfactants (*e.g.*, cationic surfactants such as for example dimethyldioctadecyl-ammonium bromide), sugars or sugar derivatives (*e.g.*, cyclodextrin), nucleic acids, polymers (*e.g.*, heparin, polyethylene glycol and poloxamer), mucolytic agents (*e.g.*, N-acetylcysteine, mugwort, bromelain, papain, clerodendrum, acetylcysteine, bromhexine, carbocysteine, eprazinone, mesna, ambroxol, sobrerol, domiodol, letosteine, stepronin, tiopronin, gelsolin, thymosin β 4 dornase alfa, neltenexine, erdosteine) and various DNases including rhDNase.

In some embodiments, the mucus penetrating LNP can be a hypotonic formulation comprising a mucosal penetration enhancing coating. The formulation can be hypotonic for the epithelium to which it is being delivered. Non-limiting examples of hypotonic formulations can be found in, *e.g.*, Intl. Pub. No. WO2013110028, herein incorporated by reference in its entirety.

In some embodiments, the polynucleotide described herein is Formulated as a lipoplex, such as, without limitation, the ATUPLEXTM system, the DACC system, the DBTC system and other siRNA-lipoplex technology from Silence Therapeutics (London, United Kingdom), STEMFECTTM from STEMGENT® (Cambridge, MA), and polyethylenimine (PEI) or protamine-based targeted and non-targeted delivery of nucleic acids (Aleku et al. Cancer Res. 2008 68:9788-9798; Strumberg et al. Int J Clin Pharmacol Ther 2012 50:76-78; Santel et al., Gene Ther 2006 13:1222-1234; Santel et al., Gene Ther 2006 13:1360-1370; Gutbier et al., Pulm Pharmacol. Ther. 2010 23:334-344; Kaufmann et al. Microvasc Res 2010 80:286-293; Weide et al. J Immunother. 2009 32:498-507; Weide et al. J Immunother. 2008 31:180-188; Pascolo Expert Opin. Biol. Ther. 4:1285-1294; Fotin-Mleczek et al., 2011 J. Immunother. 34:1-15; Song et al., Nature Biotechnol. 2005, 23:709-717; Peer et al., Proc Natl Acad Sci U S A. 2007 6;104:4095-4100; deFougerolles Hum Gene Ther. 2008 19:125-132; all of which are incorporated herein by reference in its entirety).

In some embodiments, the polynucleotides described herein are Formulated as a solid lipid nanoparticle (SLN), which can be spherical with an average diameter between 10 to 1000 nm. SLN possess a solid lipid core matrix that can solubilize lipophilic molecules and

can be stabilized with surfactants and/or emulsifiers. Exemplary SLN can be those as described in Intl. Pub. No. WO2013105101, herein incorporated by reference in its entirety.

In some embodiments, the polynucleotides described herein can be Formulated for controlled release and/or targeted delivery. As used herein, "controlled release" refers to a pharmaceutical composition or compound release profile that conforms to a particular pattern of release to effect a therapeutic outcome. In some embodiments, the polynucleotides can be encapsulated into a delivery agent described herein and/or known in the art for controlled release and/or targeted delivery. As used herein, the term "encapsulate" means to enclose, surround or encase. As it relates to the formulation of the compounds of the present disclosure, encapsulation can be substantial, complete or partial. The term "substantially encapsulated" means that at least greater than 50, 60, 70, 80, 85, 90, 95, 96, 97, 98, 99, or greater than 99% of the pharmaceutical composition or compound of the present disclosure can be enclosed, surrounded or encased within the delivery agent. "Partial encapsulation" or "partially encapsulate" means that less than 10, 10, 20, 30, 40 50 or less of the pharmaceutical composition or compound of the present disclosure can be enclosed, surrounded or encased within the delivery agent.

Advantageously, encapsulation can be determined by measuring the escape or the activity of the pharmaceutical composition or compound of the present disclosure using fluorescence and/or electron micrograph. For example, at least 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 85, 90, 95, 96, 97, 98, 99, 99.9, or greater than 99% of the pharmaceutical composition or compound of the present disclosure are encapsulated in the delivery agent.

In some embodiments, the polynucleotides described herein can be encapsulated in a therapeutic nanoparticle, referred to herein as "therapeutic nanoparticle polynucleotides." Therapeutic nanoparticles can be Formulated by methods described in, *e.g.*, Intl. Pub. Nos. WO2010005740, WO2010030763, WO2010005721, WO2010005723, and WO2012054923; and U.S. Pub. Nos. US20110262491, US20100104645, US20100087337, US20100068285, US20110274759, US20100068286, US20120288541, US20120140790, US20130123351 and US20130230567; and U.S. Pat. Nos. 8,206,747, 8,293,276, 8,318,208 and 8,318,211, each of which is herein incorporated by reference in its entirety.

In some embodiments, the therapeutic nanoparticle polynucleotide can be Formulated for sustained release. As used herein, "sustained release" refers to a pharmaceutical

composition or compound that conforms to a release rate over a specific period of time. The period of time can include, but is not limited to, hours, days, weeks, months and years. As a non-limiting example, the sustained release nanoparticle of the polynucleotides described herein can be Formulated as disclosed in Intl. Pub. No. WO2010075072 and U.S. Pub. Nos. 5 US20100216804, US20110217377, US20120201859 and US20130150295, each of which is herein incorporated by reference in their entirety.

In some embodiments, the therapeutic nanoparticle polynucleotide can be Formulated to be target specific, such as those described in Intl. Pub. Nos. WO2008121949, WO2010005726, WO2010005725, WO2011084521 and WO2011084518; and U.S. Pub. 10 Nos. US20100069426, US20120004293 and US20100104655, each of which is herein incorporated by reference in its entirety.

The LNPs can be prepared using microfluidic mixers or micromixers. Exemplary microfluidic mixers can include, but are not limited to, a slit interdigital micromixer including, but not limited to those manufactured by Microinnova (Allerheiligen bei Wildon, 15 Austria) and/or a staggered herringbone micromixer (SHM) (see Zhigaltsevet al., "Bottom-up design and synthesis of limit size lipid nanoparticle systems with aqueous and triglyceride cores using millisecond microfluidic mixing," *Langmuir* 28:3633-40 (2012); Belliveau et al., "Microfluidic synthesis of highly potent limit-size lipid nanoparticles for *in vivo* delivery of siRNA," *Molecular Therapy-Nucleic Acids* 1:e37 (2012); Chen et al., "Rapid discovery of 20 potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation," *J. Am. Chem. Soc.* 134(16):6948-51 (2012); each of which is herein incorporated by reference in its entirety). Exemplary micromixers include Slit Interdigital Microstructured Mixer (SIMM-V2) or a Standard Slit Interdigital Micro Mixer (SSIMM) or Caterpillar (CPMM) or Impinging-jet (IJMM,) from the Institut für Mikrotechnik Mainz GmbH, Mainz Germany. In 25 some embodiments, methods of making LNP using SHM further comprise mixing at least two input streams wherein mixing occurs by microstructure-induced chaotic advection (MICA). According to this method, fluid streams flow through channels present in a herringbone pattern causing rotational flow and folding the fluids around each other. This method can also comprise a surface for fluid mixing wherein the surface changes orientations 30 during fluid cycling. Methods of generating LNPs using SHM include those disclosed in U.S. Pub. Nos. US20040262223 and US20120276209, each of which is incorporated herein by reference in their entirety.

In some embodiments, the polynucleotides described herein can be Formulated in lipid nanoparticles using microfluidic technology (see Whitesides, George M., "The Origins and the Future of Microfluidics," *Nature* 442: 368-373 (2006); and Abraham et al., "Chaotic Mixer for Microchannels," *Science* 295: 647-651 (2002); each of which is herein

5 incorporated by reference in its entirety). In some embodiments, the polynucleotides can be Formulated in lipid nanoparticles using a micromixer chip such as, but not limited to, those from Harvard Apparatus (Holliston, MA) or Dolomite Microfluidics (Royston, UK). A micromixer chip can be used for rapid mixing of two or more fluid streams with a split and recombine mechanism.

10 In some embodiments, the polynucleotides described herein can be Formulated in lipid nanoparticles having a diameter from about 1 nm to about 100 nm such as, but not limited to, about 1 nm to about 20 nm, from about 1 nm to about 30 nm, from about 1 nm to about 40 nm, from about 1 nm to about 50 nm, from about 1 nm to about 60 nm, from about 1 nm to about 70 nm, from about 1 nm to about 80 nm, from about 1 nm to about 90 nm, from
 15 about 5 nm to about 100 nm, from about 5 nm to about 10 nm, about 5 nm to about 20 nm, from about 5 nm to about 30 nm, from about 5 nm to about 40 nm, from about 5 nm to about 50 nm, from about 5 nm to about 60 nm, from about 5 nm to about 70 nm, from about 5 nm to about 80 nm, from about 5 nm to about 90 nm, about 10 to about 20 nm, about 10 to about 30 nm, about 10 to about 40 nm, about 10 to about 50 nm, about 10 to about 60 nm,
 20 about 10 to about 70 nm, about 10 to about 80 nm, about 10 to about 90 nm, about 20 to about 30 nm, about 20 to about 40 nm, about 20 to about 50 nm, about 20 to about 60 nm, about 20 to about 70 nm, about 20 to about 80 nm, about 20 to about 90 nm, about 20 to about 100 nm, about 30 to about 40 nm, about 30 to about 50 nm, about 30 to about 60 nm, about 30 to about 70 nm, about 30 to about 80 nm, about 30 to about 90 nm, about 30 to
 25 about 100 nm, about 40 to about 50 nm, about 40 to about 60 nm, about 40 to about 70 nm, about 40 to about 80 nm, about 40 to about 90 nm, about 40 to about 100 nm, about 50 to about 60 nm, about 50 to about 70 nm, about 50 to about 80 nm, about 50 to about 90 nm, about 50 to about 100 nm, about 60 to about 70 nm, about 60 to about 80 nm, about 60 to about 90 nm, about 60 to about 100 nm, about 70 to about 80 nm, about 70 to about 90 nm,
 30 about 70 to about 100 nm, about 80 to about 90 nm, about 80 to about 100 nm and/or about 90 to about 100 nm.

In some embodiments, the lipid nanoparticles can have a diameter from about 10 to 500 nm. In some embodiments, the lipid nanoparticle can have a diameter greater than 100 nm, greater than 150 nm, greater than 200 nm, greater than 250 nm, greater than 300 nm, greater than 350 nm, greater than 400 nm, greater than 450 nm, greater than 500 nm, greater than 550 nm, greater than 600 nm, greater than 650 nm, greater than 700 nm, greater than 750 nm, greater than 800 nm, greater than 850 nm, greater than 900 nm, greater than 950 nm or greater than 1000 nm.

In some embodiments, the polynucleotides can be delivered using smaller LNPs. Such particles can comprise a diameter from below 0.1 μm up to 100 nm such as, but not limited to, less than 0.1 μm , less than 1.0 μm , less than 5 μm , less than 10 μm , less than 15 μm , less than 20 μm , less than 25 μm , less than 30 μm , less than 35 μm , less than 40 μm , less than 50 μm , less than 55 μm , less than 60 μm , less than 65 μm , less than 70 μm , less than 75 μm , less than 80 μm , less than 85 μm , less than 90 μm , less than 95 μm , less than 100 μm , less than 125 μm , less than 150 μm , less than 175 μm , less than 200 μm , less than 225 μm , less than 250 μm , less than 275 μm , less than 300 μm , less than 325 μm , less than 350 μm , less than 375 μm , less than 400 μm , less than 425 μm , less than 450 μm , less than 475 μm , less than 500 μm , less than 525 μm , less than 550 μm , less than 575 μm , less than 600 μm , less than 625 μm , less than 650 μm , less than 675 μm , less than 700 μm , less than 725 μm , less than 750 μm , less than 775 μm , less than 800 μm , less than 825 μm , less than 850 μm , less than 875 μm , less than 900 μm , less than 925 μm , less than 950 μm , or less than 975 μm .

The nanoparticles and microparticles described herein can be geometrically engineered to modulate macrophage and/or the immune response. The geometrically engineered particles can have varied shapes, sizes and/or surface charges to incorporate the polynucleotides described herein for targeted delivery such as, but not limited to, pulmonary delivery (see, *e.g.*, Intl. Pub. No. WO2013082111, herein incorporated by reference in its entirety). Other physical features the geometrically engineering particles can include, but are not limited to, fenestrations, angled arms, asymmetry and surface roughness, charge that can alter the interactions with cells and tissues.

In some embodiment, the nanoparticles described herein are stealth nanoparticles or target-specific stealth nanoparticles such as, but not limited to, those described in U.S. Pub. No. US20130172406, herein incorporated by reference in its entirety. The stealth or target-specific stealth nanoparticles can comprise a polymeric matrix, which can comprise two or

more polymers such as, but not limited to, polyethylenes, polycarbonates, polyanhydrides, polyhydroxyacids, polypropylfumerates, polycaprolactones, polyamides, polyacetals, polyethers, polyesters, poly(orthoesters), polycyanoacrylates, polyvinyl alcohols, polyurethanes, polyphosphazenes, polyacrylates, polymethacrylates, polycyanoacrylates, polyureas, polystyrenes, polyamines, polyesters, polyanhydrides, polyethers, polyurethanes, polymethacrylates, polyacrylates, polycyanoacrylates, or combinations thereof.

g. mRNA-Lipid Adducts

It has been determined that certain ionizable lipids are susceptible to the formation of lipid-polynucleotide adducts. In particular, ionizable lipids that comprise a tertiary amine group may decompose into one or both of a secondary amine and a reactive aldehyde species capable of interacting with polynucleotides (such as mRNA) to form an ionizable lipid-polynucleotide adduct impurity that can be detected by reverse phase ion pair chromatography (RP-IP HPLC). For example, oxidation of the tertiary amine may lead to N-oxide formation that can undergo acid/base-catalyzed hydrolysis at the amine to generate aldehydes and secondary amines which may form adducts with mRNA. Thus, in some aspects, the ionizable lipid-polynucleotide adduct impurity is an aldehyde-mRNA adduct impurity.

It also has been determined that such adducts may disrupt mRNA translation and impact the activity of lipid nanoparticle (LNP) formulated mRNA products. Thus, it can be advantageous to prepare and use LNP compositions with a reduced content of ionizable lipid-polynucleotide adduct impurity, such as wherein less than about 20%, less than about 10%, less than about 5%, or less than about 1%, of the mRNA is in the form of ionizable lipid-polynucleotide adduct impurity, as may be measured by RP-IP HPLC. Thus, in accordance with some aspects, an LNP composition is provided wherein less than about 10%, less than about 5%, or less than about 1%, of the mRNA is in the form of ionizable lipid-polynucleotide adduct impurity, including less than 10%, less than 5%, or less than 1%, as may be measured by RP-IP HPLC.

In some aspects, an amount of lipid aldehydes in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects an amount of N-oxide compounds in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects an amount of transition metals, such as Fe, in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects an amount of alkyl halide compounds in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects

an amount of anhydride compounds in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects an amount of ketone compounds in the composition is less than about 50 ppm, including less than 50 ppm. Additionally or alternatively, in some aspects an amount of conjugated diene compounds in the composition is less than about 50 ppm, including less than 50 ppm.

In some aspects, the composition is stable against the formation of ionizable lipid-polynucleotide adduct impurity. In some aspects, an amount of ionizable lipid-polynucleotide adduct impurity in the composition increases at an average rate of less than about 2% per day when stored at a temperature of about 25 °C or below, including at an average rate of less than 2% per day. In some aspects, an amount of ionizable lipid-polynucleotide adduct impurity in the composition increases at an average rate of less than about 0.5% per day when stored at a temperature of about 5 °C or below, including at an average rate of less than 0.5% per day. In some aspects, an amount of ionizable lipid-polynucleotide adduct impurity in the composition increases at an average rate of less than about 0.5% per day when stored at a refrigerated temperature, optionally wherein the refrigerated temperature is about 5 °C.

Lipid vehicle (e.g., LNP) compositions with a reduced content of ionizable lipid-polynucleotide adduct impurity can be prepared by methods that inhibit formation of one or both of N-oxides and aldehydes. Such methods may comprise treating a composition comprising an ionizable lipid comprising a tertiary amine group to inhibit formation of one or both of N-oxides and aldehydes, such as by treating the composition with a reducing agent; treating the composition with a chelating agent; adjusting the pH of the composition; adjusting the temperature of the composition; and adjusting the buffer in the composition. Such methods may comprise, prior to combining the ionizable lipid with a polynucleotide, one or more of treating the ionizable lipid with a scavenging agent; treating the ionizable lipid with a reductive treatment agent; treating the ionizable lipid with a reducing agent; treating the ionizable lipid with a chelating agent; treating the polynucleotide with a reducing agent; and treating the polynucleotide with a chelating agent.

In accordance with any of the foregoing, the scavenging agent, reductive treatment agent, and/or reducing agent may be an agent that reacts with aldehyde, ketone, anhydride and/or diene compounds. A scavenging agent may comprise one or more selected from (O-(2,3,4,5,6-Pentafluorobenzyl)hydroxylamine hydrochloride) (PFBHA), methoxyamine (e.g., methoxyamine hydrochloride), benzyloxyamine (e.g., benzyloxyamine hydrochloride), ethoxyamine (e.g., ethoxyamine hydrochloride), 4-[2-(aminooxy)ethyl]morpholine dihydrochloride, butoxyamine (e.g., tert-butoxyamine hydrochloride), 4-

Dimethylaminopyridine (DMAP), 1,4-diazabicyclo[2.2.2]octane (DABCO), Triethylamine (TEA), Piperidine 4-carboxylate (BPPC), and combinations thereof. A reductive treatment agent may comprise a boron compound (*e.g.*, sodium borohydride and/or bis(pinacolato)diboron). A reductive treatment agent may comprise a boron compound, such as one or both of sodium borohydride and bis(pinacolato)diboron). A chelating agent may comprise immobilized iminodiacetic acid. A reducing agent may comprise an immobilized reducing agent, such as immobilized diphenylphosphine on silica (Si-DPP), immobilized thiol on agarose (Ag-Thiol), immobilized cysteine on silica (Si-Cysteine), immobilized thiol on silica (Si-Thiol), or a combination thereof. A reducing agent may comprise a free reducing agent, such as potassium metabisulfite, sodium thioglycolate, tris(2-carboxyethyl)phosphine (TCEP), sodium thiosulfate, N-acetyl cysteine, glutathione, dithiothreitol (DTT), cystamine, dithioerythritol (DTE), dichlorodiphenyltrichloroethane (DDT), homocysteine, lipoic acid, or a combination thereof.

In accordance with any of the foregoing, the pH may be, or adjusted to be, a pH of from about 7 to about 9.

In accordance with any of the foregoing, a buffer may be selected from sodium phosphate, sodium citrate, sodium succinate, histidine, histidine-HCl, sodium malate, sodium carbonate, and TRIS (tris(hydroxymethyl)aminomethane). In accordance with any of the foregoing, a buffer may be TRIS and may be, or adjusted to be, from about 20 mM to about 150 mM TRIS.

In accordance with any of the foregoing, the temperature of the composition may be, or adjusted to be, 25 °C or less.

The composition may also comprise a free reducing agent or antioxidant.

2. Argininosuccinate Lyase (ASL)

ASL converts argininosuccinic acid (ASA) into arginine and is integral to the liver-based urea cycle, which detoxifies neurotoxic ammonia. Argininosuccinic aciduria (argininosuccinate lyase deficiency) is a rare autosomal recessive condition caused by mutations in the gene encoding argininosuccinate lyase (ASL). Ammonia, which is formed when proteins are broken down in the digestive process, is toxic if blood ammonia levels become too high. Thus, ASL defects can cause accumulation of ammonia in the blood (hyperammonaemia). Argininosuccinic aciduria patients exhibit elevated levels of plasma

ammonia, elevated plasma argininosuccinic acid, elevated plasma citrulline, and elevated urinary orotic acid.

A variety of mutations can affect ASL function and activity in humans. Large deletions, frameshift, nonsense, and missense mutations can abolish ASL enzymatic activity or folding, causing severe neonatal onset disease in hemizygous individuals and argininosuccinic aciduria symptoms in heterozygous individuals. Missense mutations that retain ASL activity but destabilize the protein, reduce enzymatic activity, or decrease substrate affinity can lead to late onset disease in hemizygous individuals. Argininosuccinic aciduria may cause developmental delay and intellectual disability, chronic liver disease, neurodisability, high blood pressure (hypertension), skin lesions, and brittle hair. Milder forms of ASL deficiency may manifest in accumulation of plasma ammonia during periods of illness or stress, mild intellectual disability, or learning disabilities.

The wild type human ASL canonical mRNA sequence transcript variants are described at the NCBI Reference Sequence database (RefSeq) under accession numbers: NM_001024943.2 9 ("Homo sapiens argininosuccinate lyase (ASL), transcript variant 1 mRNA"), which is shown as SEQ ID NO:41; NM_000048.4 ("Homo sapiens argininosuccinate lyase (ASL), transcript variant 2, mRNA"), which is shown as SEQ ID NO:42; NM_001024944.2 ("Homo sapiens argininosuccinate lyase (ASL), transcript variant 3 mRNA"), which is shown as SEQ ID NO: 43; or NM_001024946.2 ("Homo sapiens argininosuccinate lyase (ASL), transcript variant 4 mRNA"), which is shown as SEQ ID NO:44.

In some embodiments, polynucleotides according to the present disclosure may encode the wild type human ASL polypeptide or variants thereof. In some embodiments, the ASL polypeptide (*e.g.*, ASL polypeptide) may be encoded by an mRNA sequence having an open reading frame (ORF) with at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, at least about 99.9%, or 100% identity to the nucleic acid sequence according to any one of SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44.

The wild type ASL canonical protein sequence is described at the RefSeq database under accession number NP_000039.2 (“Argininosuccinate lyase isoform 1 [Homo sapiens]”). The ASL protein is 464 amino acids long.

The amino acid sequence of human ASL is provided in SEQ ID NO:1:

```

5  MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
  ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
  MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
  LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
  CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRCA
10 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
  LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGH
  SV
  EQYGALGGTARSSVDWQIRQVRALLQAQQA.

```

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) K51R is provided in SEQ ID NO:2:

```

15 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
  ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
  MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
  LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
  CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRCA
20 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
  LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGH
  SV
  EQYGALGGTARSSVDWQIRQVRALLQAQQA.

```

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) K57R is provided in SEQ ID NO:3:

```

25 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTRAEMDQ
  ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
  MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
  LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
  CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRCA
30 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD

```

LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) K80R is provided in SEQ ID NO:4:

5 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFRNLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRACGL
10 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) K97R is provided in SEQ ID NO:5:

15 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLRELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRACGL
20 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) A104G is provided in SEQ ID NO:6:

25 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATGGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRACGL
30 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD

LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) C307L is provided in SEQ ID NO:7:

5 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRLAGL
10 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) G223D is provided in SEQ ID NO:8:

15 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFDAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRCAGL
20 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) T233V is provided in SEQ ID NO:9:

25 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSE
LLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDAVSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGRVFGRCAGL
30 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD

LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) K266P is provided in SEQ ID NO:10:

5 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSER
LLEVRKRINVLPGLSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTPEFSFVQLSDAYSTGSSLMPOKKNPDSLELIRSKAGRVFGRACGL
10 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYGALGGTARSSVDWQIRQVRALLQAQQA.

The amino acid sequence of a variant human ASL having the substitution (numbered according to SEQ ID NO:1) G438N is provided in SEQ ID NO:11:

15 MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGLEKAGLLTKAEMDQ
ILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELIGATAGKLHTGRSRNDQVVTDLRLW
MRQTCSTLSGLLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSER
LLEVRKRINVLPGLSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASL
CMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPOKKNPDSLELIRSKAGRVFGRACGL
20 LMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQALSPDMLATD
LAYYLVRKGMPFRQAHEASGKAVFMAETKGVALNQLSLQELQTISPLFSGDVICVWDYGHVS
EQYNALGGTARSSVDWQIRQVRALLQAQQA.

In certain aspects, the disclosure provides a polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) comprising a nucleotide sequence (*e.g.*, an open reading frame (ORF)) encoding an
25 ASL polypeptide. In some embodiments, the ASL polypeptide of the present disclosure is a wild type full length human ASL protein (SEQ ID NO:1). In some embodiments, the ASL polypeptide of the present disclosure is a variant, a peptide or a polypeptide containing a substitution (*e.g.*, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11), and
30 insertion and/or an addition, a deletion and/or a covalent modification with respect to a wild-type ASL sequence. In some embodiments, sequence tags or amino acids, can be added to the

sequences encoded by the polynucleotides of the present disclosure (*e.g.*, at the N-terminal or C-terminal ends), *e.g.*, for localization. In some embodiments, amino acid residues located at the carboxy, amino terminal, or internal regions of a polypeptide of the present disclosure can optionally be deleted providing for fragments.

5 In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a nucleotide sequence (*e.g.*, an ORF) of the present disclosure encodes a wild type full length human ASL protein (SEQ ID NO:1).

 In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a nucleotide sequence (*e.g.*, an ORF) of the present disclosure encodes a substitutional variant
10 of a human ASL sequence, which can comprise one, two, three, four, five, or more than five substitutions. In some embodiments, the substitutional variant can comprise one or more conservative amino acid substitutions. In some embodiments, the substitutional variant comprises the amino acid sequence of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ
15 ID NO:11. In some embodiments, the substitutional variant comprises the amino acid sequence of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5. In some embodiments, the variant is an insertional variant. In some embodiments, the variant is a deletional variant.

 ASL protein fragments, functional protein domains, variants, and homologous
20 proteins (orthologs) are also within the scope of the ASL polypeptides of the disclosure. A nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:2. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:3. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present
25 disclosure is shown in SEQ ID NO:4. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:5. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:6. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:7. Another
30 nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:8. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:9. Another

nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:10. Another nonlimiting example of a polypeptide encoded by the polynucleotides of the present disclosure is shown in SEQ ID NO:11.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a nucleotide sequence (*e.g.*, an ORF) of the present disclosure encodes a human ASL comprising (i) an amino acid other than lysine (K) at the position corresponding to position 51 of SEQ ID NO:1; (ii) an amino acid other than lysine (K) at the position corresponding to position 57 of SEQ ID NO:1; (iii) an amino acid other than lysine (K) at the position corresponding to position 80 of SEQ ID NO:1; (iv) an amino acid other than lysine (K) at the position corresponding to position 97 of SEQ ID NO:1; (v) an amino acid other than alanine (A) at the position corresponding to position 104 of SEQ ID NO:1; (vi) an amino acid other than cysteine (C) at the position corresponding to position 307 of SEQ ID NO:1; (vii) an amino acid other than glycine (G) at the position corresponding to position 223 of SEQ ID NO:1; (viii) an amino acid other than threonine (T) at the position corresponding to position 233 of SEQ ID NO:1; (ix) an amino acid other than lysine (K) at the position corresponding to position 266 of SEQ ID NO:1; and/or (x) an amino acid other than glycine (G) at the position corresponding to position 438 of SEQ ID NO:1.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a nucleotide sequence (*e.g.*, an ORF) of the present disclosure encodes a human ASL comprising one or more substitutions selected from the group consisting of K51R, K57R, K80R, K97R, A104G, C307L, G223D, T233V, K266R, and G438N (numbered according to SEQ ID NO:1).

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a nucleotide sequence (*e.g.*, an ORF) of the present disclosure encodes a human ASL comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11.

Certain compositions and methods presented in this disclosure refer to the protein or polynucleotide sequences of wild type human ASL (SEQ ID NO:1). Such disclosures are equally applicable to any other variants of ASL known in the art or described herein (*e.g.*, SEQ ID NO:2).

3. Polynucleotides and Open Reading Frames (ORFs)

The present disclosure features mRNAs for use in treating or preventing argininosuccinic aciduria. The mRNAs featured for use in the present disclosure are administered to subjects and encode human ASL protein or human ASL protein variants *in vivo*. Accordingly, the present disclosure relates to polynucleotides, *e.g.*, mRNA, comprising an open reading frame of linked nucleosides encoding human ASL (SEQ ID NO:1), isoforms thereof, variants thereof (*e.g.*, SEQ ID NO:2–11), functional fragments thereof, and fusion proteins comprising ASL. Specifically, the present disclosure provides sequence-optimized polynucleotides comprising nucleotides encoding the polypeptide sequence of human ASL (or a variant thereof, *e.g.*, SEQ ID NO:2), or sequences having high sequence identity with those sequence optimized polynucleotides.

In certain aspects, the invention provides polynucleotides (*e.g.*, a RNA such as an mRNA) that comprise a nucleotide sequence (*e.g.*, an ORF) encoding one or more ASL polypeptides. In some embodiments, the encoded ASL polypeptide of the present disclosure can be selected from:

- (i) a full length ASL polypeptide (*e.g.*, having the same or essentially the same length as wild-type ASL; *e.g.*, SEQ ID NO:1);
- (ii) a functional fragment of ASL described herein (*e.g.*, a truncated (*e.g.*, deletion of carboxy, amino terminal, or internal regions) sequence shorter than ASL; but still retaining ASL enzymatic activity); or
- (iii) a variant thereof (*e.g.*, full length or truncated ASL proteins in which one or more amino acids have been replaced, *e.g.*, variants that retain all or most of the ASL activity of the polypeptide with respect to a reference protein (*e.g.*, any natural or artificial variants known in the art or described herein (*e.g.*, SEQ ID NOs:2–11))); or

In certain embodiments, the encoded ASL polypeptide is a mammalian ASL polypeptide, such as a human ASL polypeptide, a functional fragment or a variant thereof.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure increases ASL protein expression levels and/or detectable ASL enzymatic activity levels in cells when introduced in those cells, *e.g.*, by at least 10%, at least 15%, at

least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 100%, compared to ASL protein expression levels and/or detectable ASL enzymatic activity levels in the cells prior to the administration of the polynucleotide of the present disclosure. ASL protein expression levels and/or ASL enzymatic activity can be measured according to methods known in the art. In some embodiments, the polynucleotide is introduced to the cells *in vitro*. In some embodiments, the polynucleotide is introduced to the cells *in vivo*.

In some embodiments, the polynucleotides (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprise a nucleotide sequence (*e.g.*, an ORF) that encodes a wild-type human ASL (*e.g.*, SEQ ID NO:1) or an isoform thereof.

In some embodiments, the polynucleotides (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprise a nucleotide sequence (*e.g.*, an ORF) that encodes a variant human ASL (*e.g.*, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11) or an isoform thereof.

The polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a codon optimized nucleic acid sequence, wherein the open reading frame (ORF) of the codon optimized nucleic acid sequence is derived from a wild-type ASL sequence (*e.g.*, wild-type human ASL). For example, for polynucleotides of invention comprising a sequence optimized ORF encoding ASL, the corresponding wild type sequence is the native human ASL. Similarly, for a sequence optimized mRNA encoding a functional fragment of human ASL, the corresponding wild type sequence is the corresponding fragment from human ASL.

In some embodiments, the polynucleotides (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprise a nucleotide sequence encoding ASL having the full-length sequence of human ASL (*i.e.*, including amino acids 1-464, *e.g.*, SEQ ID NO:1).

In some embodiments, the polynucleotides (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprise a nucleotide sequence (*e.g.*, an ORF) encoding a mutant ASL polypeptide. In some embodiments, the polynucleotides of the present disclosure comprise an ORF encoding an ASL polypeptide that comprises at least one point mutation in the ASL amino acid sequence and retains ASL enzymatic activity. In some embodiments, the mutant

ASL polypeptide has an ASL activity which is at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 100% of the ASL activity of the corresponding wild-type ASL (*e.g.*, SEQ ID NO:1). In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprising an ORF encoding a mutant ASL polypeptide is sequence optimized. In some embodiments, the mutant ASL polypeptide is SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) that encodes an ASL polypeptide with mutations that do not alter ASL enzymatic activity. Such mutant ASL polypeptides can be referred to as function-neutral. In some embodiments, the polynucleotide comprises an ORF that encodes a mutant ASL polypeptide comprising one or more function-neutral point mutations.

In some embodiments, the mutant ASL polypeptide has higher ASL enzymatic activity than the corresponding wild-type ASL. In some embodiments, the mutant ASL polypeptide has an ASL activity that is at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 100% higher than the activity of the corresponding wild-type ASL (*i.e.*, the same ASL protein but without the mutation(s)).

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof), wherein the nucleotide sequence has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, as shown in **Table 1** below.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof), wherein the nucleotide sequence has, 65% to 100%, 70% to 100%, 75% to 100%, 80% to 100%, 85% to 100%, 90% to 100%, 95% to 100%, 97% to 100%, 98% to 100%, 99% to 100%, 90% to 95%, 90% to 97%, 90% to 98%, 95% to 97%, 95% to 98%, or 95% to 99% sequence identity to the sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, as shown in **Table 1** below.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof), wherein the nucleotide sequence is at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, or 100% identical to the sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, as shown in **Table 1** below.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof), wherein the nucleotide sequence is 65% to 100%, 70% to 100%, 75% to 100%, 80% to 100%, 85% to 100%, 90% to 100%, 95% to 100%, 97% to 100%, 98% to 100%, 99% to 100%, 90% to 95%, 90% to 97%, 90% to 98%, 95% to 97%, 95% to 98%, or 95% to 99% identical to the sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, or SEQ ID NO:44, as shown in **Table 1**.

Table 1. Nucleotide Sequences (ORFs) encoding Human ASL or Variants Thereof

SEQ ID NO	Nucleic Acid Sequence
-----------	-----------------------

20 K51R	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AGAGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGGGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUCAACGUGCUGCCCCUGGGC AGCGGCGCCAUAAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUGGGCCAGGCCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
21 K57R	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGAGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGGGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUCAACGUGCUGCCCCUGGGC AGCGGCGCCAUAAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUGGGCCAGGCCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU

	UCCGGCAGGCCCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
22 K80R	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGA UCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAGACUGAACAGCA ACGACGAGGACA UCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGU UCCCCGGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGCCCUAGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUACAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAU UCCAUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGU UCCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCCUAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
23 K97R	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGA UCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACA UCCACACCGCCAACGAGAGAAGACUGAGAGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGU UCCCCGGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGCCCUAGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUACAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAU UCCAUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGU UCCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA

	GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCCUAGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
24 A104G	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGGCGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGCCGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGCCCUAGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUCAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUAUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCCUAGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
25 C307L	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGCCGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGCCCUAGACCAGAGUA GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUCAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG

	GGCCGAGCUGAACUUCGGUGCCAUACCCUGAAUUCCAUGGACGCCACUA GCGAGCGGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGCUGGGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUGGGCCAGGCCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
26 G223D	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGCCGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGAU GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUUAACGUGCUGCCCCUGGGC AGCGGCGCCAUAAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGACGCCAUACCCUGAAUUCCAUGGACGCCACUA GCGAGCGGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUGGGCCAGGCCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
27 T233V	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG

	GCCUUCUUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCCTCCGGCUACACCCACCUGCAGAGAGCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGAU GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUACAACGUGCUGCCCCUGGGC AGCGGCGCCAUAAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCCGUGA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
28 K266P	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCCTCCGGCUACACCCACCUGCAGAGAGCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGAU GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUACAACGUGCUGCCCCUGGGC AGCGGCGCCAUAAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCCUGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
29 G438N	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG

	AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGCCU GGACAAGGUGGCCGAGGAGUGGGCCCAGGGCACC UUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCG GCCUUCUUUGGGAAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUC CCGGCUACACCCACCUGCAGAGAGCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGCCUGACCAGAGAU GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUACAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAACCCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUCCAUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUUAUAGCACCCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUUGGGCCAGGCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCCAGUUGUCUCUCCAGGAACUGCAGACCAUCUC CCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACAACGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
30	AUGGCGAGCGAGAGCGGUAAAUUGUGGGGAGGACGGUUCGUGGGUGCCGU GGACCCUAUAAUGGAGAAAUUAACGCCAGUAUUGCAUACGAUCGUAUC UGUGGGAAGUGGACGUGCAAGGAAGCAAAGCUUACUCCAGAGGCCUCGAG AAAGCUGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCCUGCACGGACU GGACAAAGUGGCCGAGGAGUGGGCACAGGGAACUUAACUGAAUUA ACGACGAGGACAUACAUAACGCCAACGAGCGGAGACUGAAGGAACUGAUC GGCGCCACAGCUGGCAAGCUCCACACAGGACGAUCAAGGAACGACCAGGU GGUGACAGAUUCUGAGGCUGUGGAUGAGGCAGACCUGCUCGACCCUCUCAG GUCUGCUGUGGGAGUUAUCCGCACCAUGGUCGAUCGCGCCGAAGCCGAG CGCGACGUACUGUCCCAGGGUACACUCAUCUGCAGAGGGCUCAGCCUAU CCGCUGGUCGCACUGGAUCCUCAGUCACGCCGUGGCGCUGACGCGGACU CUGAGCGCCUGCUGGAAGUGCGUAAACGCAUCAACGUGCUGCCACUGGGC UCCGGCGCUAUUGCUGGCAAUCCACUAGGCGUGGACCGGGAGCUUCUGAG AGCCGAACUGAACUUCGGGGCGAUCACCUUAAAUAGCAUGGACGCCACUU CCGAACGCGAUUUCGUUGCAGAGUUUCUGUUUUGGGCAUCCUGUGCAUG ACUCACCUCAGUCGUAUGGCCGAAGACCUAUCCUCUAUUGCACAAGGA GUUUUCCUUUGUGCAGCUUUCUGACGCCUACAGCACAGGUUCCUCACUCA UGCCACAGAAGAAGAAUCCUGACAGCCUAGAGCUAAUCAGAUUAAGGCC GGCAGAGUCUUUGGCCGGUGUGCAGGGCUGCUGAUGACACUUAAGGACU GCCCUCUACUUAUAUAAGGACCUGCAAGAGGAUAAGGAGGCCGUGUUUG AGGUUAGCGACACCAUGUCAGCCGUCCUCCAGGUAGCAACGGGCGUGAUC AGCACUCUGCAGAUCCAUAAGAGAACAUGGGUCAGGCCUGUCACCAGA CAUGCUGGCCACAGAUUUGGCCUAUUUACCUGGUGCGCAAGGGGAUGCCCU UUCGGCAAGCCCACGAAGCAUCAGGCAAGGCUGUGUUUAUGGCAGAGACA AAGGGAGUGGCCCUCAACCAAUUGUCUCUGCAGGAAUUGCAAACCAUCUC UCCUCUGUUUCCGGGGACGUGAUCUGCGUGUGGGACUACGGGCAUUCUG

	UGGAACAGUACGGUGCGCUGGGAGGGACCGCUAGGUCCUCCGUCGACUGG CAGAUCCGGCAAGUGCGGGCCCUGCUGCAAGCGCAGCAGGCU
31	AUGGCAAGCGAGUCCGGGAAACUUUGGGGAGGUCGGUUUGUGGGCGCUGU GGACCCCAUUAUGGAGAAGUUCAACGCGUCCAUCGCAUACGACCGACACC UCUGGGAGGUGGACGUGCAGGGCUCUAAAAGCCUAUAGCAGGGGUCUGGAG AAGGCCGGGCUGCUCACUAAGGCUGAAAUGGACCAGAUACUGCACGGGCU GGAUAAGGUGGCUGAAGAGUGGGCCCAAGGCACGUUCAAGCUGAAUAGCA ACGACGAGGACAUACACACAGCCAACGAGAGAAGACUGAAGGAGUUAUU GGCGCGACAGCCGGGAAGUUGCACACUGGCAGGUCCCCGGAACGACCAAGU GGUAACCGACUUGCGGUUGUGGAUGCGUCAGACGUGUUCCACUCUGAGCG GCUUGCUGUGGGAGCUAAUCCGUACAAUGGUGGACAGGGCUGAAGCCGAG AGAGACGUCCUGUUCCCAGGAUACACCCACCUGCAGCGAGCACAGCCUAU ACGGUGGUCGCACUGGAUUUUGAGCCACGCGGUCGCGCUCACUAGAGAU GCGAGCGCCUUUUGGAAGUGCGCAAACGGAUCAACGUUCUGCCGCUCGGG UCCGGGGCAAUUGCCGGAAAUCCCCUGGGCGUAGACAGGGAGCUCCUAAG AGCAGAGUUGAAUUCGGUGCCAUAACCCUGAACAGCAUGGACGCCACA CGGAGCGGGAUUUCUGGGCAGAAUUUCUAUUCUGGGCAUCACUUUGUAUG ACUCAUCUGAGCCGUAUGGCGGAAGAUCUUAUAUUGUACUGUACCAAGGA GUUUUCUUUGUCCAGCUGUCAGACGCUUACUCCACAGGGUCAUCACUGA UGCCACAGAAGAAGAACCUGACAGCCUGGAAUUAUCCGCUCCAAGGCG GGCAGGGUAUUUGGGCGUUGUGCCGGCCUGCUGAUGACUCUGAAGGGUUU ACCUUCUACCUAUAACAAGGAUCUCCAGGAGGACAAGGAAGCUGUUUUCG AGGUGUCCGAUACCAUGUCCGCCGUGUUGCAGGUGGCAACAGGUGUUAU AGCACCCUCCAGAUUCACCAAGAGAACAUGGGCCAGGCACUGUCGCCCGA UAUGCUGGCUACUGAUCUGGCAUAUUAUCUCGUCCGGAAGGGCAUGCCAU UCCGGCAGGCGCACGAGGCCUCAGGGAAGGCUGUGUUAUGGCCGAAACC AAGGGCGUGGCACUCAACCAGUUGAGCUUGCAAGAGCUACAGACUAUCAG CCUCUCUUUUCUGGUGACGUGAUCUGUGUGUGGGAUUACGGCCACUCCG UGGAGCAGUACGGUGCCCUCGGUGGCACAGCCCGAAGCUCUGUCGAUUGG CAGAUUCGGCAGGUGCGGGCGCUACUGCAAGCCCAGCAGGCA
32	AUGGCGAGUGAAAGUGGGAAAUUGUGGGGUGGGCCGCUUCGUAGGGGCCGU GGACCCUAUCAUGGAGAAAUAACGCAAGCAUCGCCUACGAUAGGCAUC UUUGGGAGGUUGACGUUCAAGGUUCAAGGCGUAUUAAGAGGGCUAGAG AAGGCGGGACUGCUUACGAAAGCCGAGAUGGACCAGAUCCUACACGGGCU AGACAAGGUCGCAGAAGAGUGGGCUCAGGGGACUUUCAAACUCAACAGCA ACGACGAAGAUUAUACAUACUGCUAACGAGAGGAGACUGAAGGAACUGAUA GGGGCCACCGCUGGCAAGCUGCACACAGGGCGCAGCCGCAACGACCAGGU CGUGACUGACCUGAGACUUUGGAUGCGCCAGACGUGCAGUACACUGUCAG GCCUCCUUUGGGAGCUGAUCCGGACUAUGGUGGAUAGGGCCGAGGCUGAG CGCGACGUUUUGUUUCCAGGCUACACUCACCUGCAGCGCGCUCAACCUAU UCGUUGGUCACACUGGAUACUGAGUCACGCCGUUGCCCUACUAGAGACU CAGAACGGCUCCUGGAAGUACGGAAGCGUAUCAACGUACUACCCCUUGGA UCUGGAGCCAUCGCAGGGAACCCACUAGGCGUAGACAGAGAAUUGUUGCG AGCCGAACUGAAUUCGGCGCAAUAACACUGAAUUCUAUGGACGCCACCA GCGAGCGGGAUUUUGUCGAGAGUUUCUUUUGGGCAAGUUUGUGUAUG ACGCACCUUUCAGGAUGGCUGAAGACCUGAUCCUGUAUUGUACUAAGGA GUUUUCCUUCGUCCAACUUAGCGACGCAUACAGCACCGGCAGCUCUUUA UGCCCCAGAAGAAGAAUCCCGAUUCCCUUGGAGCUCAUCAGAAGCAAGGCU GGGCGGGUGUUUGGUAGGUGCGCCGGCCUCUUGAUGACACUGAAAGGCCU GCCUUCACCUACAUAUAAAGAUUUGCAGGAAGAUAAAGAGGCCGUCUUCG AGGUUUCUGACACAAUGUCUGCCGUGUUACAAGUGGCUACAGGUGUGAUU

	AGUACCUUGCAGAUCCACCAGGAGAACAUGGGUCAAGCCCUGUCGCCCGA UAUGCUCGCCACCGACUUGGCUUACUAUCUGGUGCGAAAGGGUAUGCCUU UCCGGCAGGCCCACGAGGCCAGUGGAAAGGCCGUGUUUAUGGCGGAAACU AAGGGAGUAGCUCUGAACCAACUAAGCCUUCAGGAGUUACAAACUAUCUC CCCGCUGUUCAGUGGUGACGUGAUCUGCGUGUGGGACUACGGCCAUUCUG UGGAACAGUACGGAGCGCUGGGUGGAACCGCCCGUUCUCAGUGGACUGG CAGAUUCGCCAGGUACGGGCCCUGCUGCAGGCGCAGCAGGCC
33	AUGGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCUGUGGGCGCCGU GGACCCCAUCAUGGAGAAGUUCAACGCCAGCAUCGCCUACGACAGACACC UGUGGGAGGUGGACGUGCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAG AAGGCCGGCCUGCUGACCAAGGCCGAGAUGGACCAGAUCUGCACGGCCU GGACAAGGUGGGCCGAGGAGUGGGCCCAGGGCACCUUCAAGCUGAACAGCA ACGACGAGGACAUCCACACCGCCAACGAGAGAAGACUGAAGGAGCUGAUC GGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAGGU GGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCUGAGCG GCCUUCUUUGGGAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCAGAG AGAGACGUGCUGUUCGCCGCUACACCCACCUGCAGAGAGCCCAGCCCAU CCGCUGGAGCCACUGGAUCCUGAGCCACGCCGUGGGCCUGACCAGAGAU GCGAGAGACUGCUGGAGGUGAGAAAGAGAAUCAAACGUGCUGCCCCUGGGC AGCGGCGCCAUAGCAGGUAAACCUUUAGGCGUUGACAGAGAGCUGCUGCG GGCCGAGCUGAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCCACUA GCGAGCGCGAUUUCGUGGCAGAGUUCUGUUCUGGGCCAGCCUGUGCAUG ACCCAUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGA GUUCAGCUUCGUGCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGA UGCCCCAGAAGAAGAACCCCGACAGCCUGGAACUUAUCAGAAGUAAGGCA GGCAGAGUGUUCGGCCGGUGCGCGGGUCUGCUAAUGACCCUGAAGGGCCU GCCCAGCACCUACAACAAGGACCUUCAGGAGGACAAGGAGGCCGUGUUCG AGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGCGUGAUU AGCACCCUUCAGAUCCACCAGGAGAACAUGGGCCAGGCCCUGAGCCCCGA CAUGUUAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAU UCCGGCAGGCCCACGAGGCCAGUGGCAAGGCCGUCUUAUGGCCGAAACG AAGGGAGUUGCGCUGAACCAAGUUGUCUUCUCCAGGAACUGCAGACCAUCUC CCCUCUGUUCAGCGGCGACGUGAUCUGCGUGUGGGACUACGGCCACAGCG UGGAGCAGUACGGCGCGCUCGGUGGCACCGCACGAUCCAGCGUAGACUGG CAGAUCCGCCAGGUGAGAGCCCUGCUGCAAGCCCAGCAGGCC
41	AGTGGCGGGCGCGACACTATCCGTGCGGCCAGGCGGAGGTGAGTGCGCGG CGGCCGGATGGGCGGGACGGGCGTGAGGACGCCGAGACCGTGGCGCGC GCTCACGTCCGCGTCCCCAAGGGCTGCGCTCCCTCAAGCGCAGTGCCCAG AACTCGGAGCCAGCCCGGGCCGGGGGACCCTGCTGGCCAAGGAGGTGCTC AGTCCGGTCTTGTCTTCCAGACCCGGAGGACCGAAGCTTCCGGACGACGA GGAACCGCCCAACATGGCCTCGGAGAGTGGGAAGCTTTGGGGTGGCCGGT TTGTGGGTGCAGTGGACCCCATCATGGAGAAGTTCAACGCGTCCATTGCC TACGACCGGCACCTTTGGGAGGTGGATGTTCAAGGCAGCAAAGCCTACAG CAGGGGCCTGGAGAAGGCAGGGCTCCTCACCAAGGCCGAGATGGACCAGA TACTCCATGGCCTAGACAAGGTGGCTGAGGAGTGGGCCAGGGCACCTTC AAACTGAACCTCAATGATGAGGACATCCACACAGCCAATGAGCGCCGCCT GAAGGAGCTCATTGGTGCAACGGCAGGGAAGCTGCACACGGGACGGAGCC GGAATGACCAGGTGGTCACAGACCTCAGGCTGTGGATGCGGCAGACCTGC TCCACGCTCTCGGGCCTCCTCTGGGAGCTCATTAGGACCATGGTGGATCG GGCAGAGGCGGAACGTGATGTTCTCTTCCCGGGGTACACCCATTTGCAGA GGGCCAGCCCATCCGCTGGAGCCACTGGATTCTGAGCCACGCCGTGGCA

	CTGACCCGAGACTCTGAGCGGCTGCTGGAGGTGCGGAAGCGGATCAATGT CCTGCCCCCTGGGGAGTGGGGCCATTGCAGGCAATCCCCTGGGTGTGGACC GAGAGCTGCTCCGAGCAGAACTCAACTTTGGGGCCATCACTCTCAACAGC ATGGATGCCACTAGTGAGCGGGACTTTGTGGCCGAGTTCCTGTTCTGGGC TTCGCTGTGCATGACCCATCTCAGCAGGATGGCCGAGGACCTCATCCTCT ACTGCACCAAGGAATTTCAGCTTCGTGCAGCTCTCAGATGCCTACAGCACG GGAAGCAGCCTGATGCCCCAGAAGAAAAACCCGACAGTTTGGAGCTGAT CCGGAGCAAGGCTGGGCGTGTGTTTGGGCGGTGTGCCGGGCTCCTGATGA CCCTCAAGGGACTTCCCAGCACCTACAACAAAGACTTACAGGAGGACAAG GAAGCTGTGTTTGAAGTGTGAGACACTATGAGTGCCGTGCTCCAGGTGGC CACTGGCGTCATCTCTACGCTGCAGATTCACCAAGAGAACATGGGACAGG CTCTCAGCCCCGACATGCTGGCCACTGACCTTGCCATTATACCTGGTCCGC AAAGGGATGCCATTCCGCCAGGCCCACGAGGCCTCCGGGAAAGCTGTGTT CATGGCCGAGACCAAGGGGGTTCGCCCTCAACCAGCTGTCACTGCAGGAGC TGCAGACCATCAGCCCCCTGTTCTCGGGCGACGTGATCTGCGTGTGGGAC TACGGGCACAGTGTGGAGCAGTATGGTGCCCTGGGCGGCACTGCGCGCTC CAGCGTCGACTGGCAGATCCGCCAGGTGCGGGCGCTACTGCAGGCACAGC AGGCCTAGGTCCCTCCACACCTGCCCCCTAATAAAGTGGGCGCGAGAGGA GGCTGCTGTGTGTTTCCCTGCCCCAGCCTGGCTCCCTCGTTGCTGGGCTTT CGGGGCTGGCCAGTGGGGACAGTCAGGGACTGGAGAGGCAGGGCAGGGTG GCCTGTAATCCCAGCACTTTGGAAGGGCAAGGTGCGAGGATGCTTGAGGC CAGGAGTTTGACACAGCCTGGGCAACACAGGGAGACCCCCATCTCTACTC AATAATAAAACAAATAGCCTGGCGTGGTGGCCCATGCATATAGTCCCAGC TACTTGTAAGGCTGAGGTGAGAGGACACTTGTGCCAGGAGTGGAGGCTG CAGTGAGCTATGATCACGCCACTGCATTCCAGCCTGGATAACAGAGTGAG AACCTATCTCTAAAAATAAATAAATAAACGAAAAATAAATGGAAGCGGAG AAAAC TGGGCAAGGGCAATGAGAGTCGTAGACGGGAAGGGAAGAGGGGGC TCCCATCATAGCCTCTGCTCTGTCCAGGCCCCATCCCCTTCAGACAGGGT ATCACAGTGACCTCCTAGGCCAGGAGCGATGGCTCACGCCTGTAATCCTA GCATTTTGGGAGGCCTGGGCAACAGGAAGACTCCAAC TACCATATTAAAA AAAACATGAATGAAAGCAAAAA
42	AGTGGCGGGCGCGACACTATCCGTGCGGCCAGGCGGAGACCCGGAGGACC GAAGCTTCCGGACGACGAGGAACCGCCCAACATGGCCTCGGAGAGTGGGA AGCTTTGGGGTGGCCGGTTTGTGGGTGCAGTGGACCCCATCATGGAGAAG TTCAACGCGTCCATTGCCTACGACCGGCACCTTTGGGAGGTGGATGTTCA AGGCAGCAAAGCCTACAGCAGGGGGCCTGGAGAAGGCAGGGCTCCTCACCA AGGCCGAGATGGACCAGATACTCCATGGCCTAGACAAGGTGGCTGAGGAG TGGGCCCAGGGCACCTTCAAAC TGAAC TCCAATGATGAGGACATCCACAC AGCCAATGAGCGCCGCTGAAGGAGCTCATTGGTGCAACGGCAGGGAAGC TGCACACGGGACGGAGCCGGAATGACCAGGTGGTACACAGACCTCAGGCTG TGGATGCGGCAGACCTGCTCCACGCTCTCGGGCCTCCTCTGGGAGCTCAT TAGGACCATGGTGGATCGGGCAGAGGCGGAACGTGATGTTCTCTTCCCGG GGTACACCCATTTGCAGAGGGCCCAGCCCATCCGCTGGAGCCACTGGATT CTGAGCCACGCCGTGGCACTGACCCGAGACTCTGAGCGGCTGCTGGAGGT GCGGAAGCGGATCAATGTCCTGCCCCCTGGGGAGTGGGGCCATTGCAGGCA ATCCCCTGGGTGTGGACCGAGAGCTGCTCCGAGCAGAACTCAACTTTGGG GCCATCACTCTCAACAGCATGGATGCCACTAGTGAGCGGGACTTTGTGGC CGAGTTCCTGTTCTGGGCTTCGCTGTGCATGACCCATCTCAGCAGGATGG CCGAGGACCTCATCCTCTACTGCACCAAGGAATTCAGCTTCGTGCAGCTC TCAGATGCC TACAGCACGGGAAGCAGCCTGATGCCCCAGAAGAAAAACCC CGACAGTTTGGAGCTGATCCGGAGCAAGGCTGGGCGTGTGTTTGGGCGGT

	GTGCCGGGCTCCTGATGACCCTCAAGGGACTTCCCAGCACCTACAACAAA GACTTACAGGAGGACAAGGAAGCTGTGTTTGAAGTGTGAGACACTATGAG TGCCGTGCTCCAGGTGGCCACTGGCGTCATCTCTACGCTGCAGATTACCC AAGAGAACATGGGACAGGCTCTCAGCCCCGACATGCTGGCCACTGACCTT GCCTATTACCTGGTCCGCAAAGGGATGCCATTCCGCCAGGCCACGAGGC CTCCGGGAAAGCTGTGTTTCATGGCCGAGACCAAGGGGGTTCGCCCTCAACC AGCTGTCACTGCAGGAGCTGCAGACCATCAGCCCCCTGTTCTCGGGCGAC GTGATCTGCGTGTGGGACTACGGGCACAGTGTGGAGCAGTATGGTGCCCT GGGCGGCACTGCGCGCTCCAGCGTCGACTGGCAGATCCGCCAGGTGCGGG CGCTACTGCAGGCACAGCAGGCCTAGGTCCTCCACACCTGCCCCCTAAT AAAGTGGGCGCGAGAGGAGGCTGCTGTGTGTTTCTGCCCCAGCCTGGCT CCCTCGTTGCTGGGCTTTCGGGGCTGGCCAGTGGGGACAGTCAGGGACTG GAGAGGCAGGGCAGGGTGGCCTGTAATCCCAGCACTTTGGAAGGGCAAGG TGCGAGGATGCTTGAGGCCAGGAGTTTGACACAGCCTGGGCAACACAGGG AGACCCCCATCTCTACTCAATAATAAAACAAATAGCCTGGCGTGGTGGCC CATGCATATAGTCCCAGCTACTTGTAAGGCTGAGGTGAGAGGACACTTGT GCCCAGGAGTGGAGGCTGCAGTGAGCTATGATCACGCCACTGCATTCCAG CCTGGATAACAGAGTGAGAACCCTATCTCTAAAAATAAATAAATAAACGAA AAATAAATGGAAGCGGAGAAAACCTGGGCAAGGGCAATGAGAGTCGTAGAC GGGAAGGGAAAGAGGGGGCTCCCATCATAGCCTCTGCTCTGTCCAGGCCCC ATCCCCTTCAGACAGGGTATCACAGTGACCTCCTAGGCCAGGAGCGATGG CTCACGCCTGTAATCCTAGCATTTTGGGAGGCCTGGGCAACAGGAAGACT CCAACCTACCATATTAAAAAAACATGAATGAAAGCAAAAA
43	AGTGGCGGGCGCGACACTATCCGTGCGGCCAGGCGGAGGTGAGTGCGCGG CGGCCGGATGGGCGGGACGGGCGTGGAGGACGCCGAGCACCGTGGCGCGC GCTCACGTCCGCGTCCCCAAGGGCTGCGCTCCCTCAAGCGCAGTGCCAG AACTCGGAGCCAGCCCCGGCCCCGGGGGACCCTGCTGGCCAAGGAGGTGCTC AGTCCGGTCTTGCTTCCAGACCCGGAGGACCGAAGCTTCCGGACGACGA GGAACCGCCCAACATGGCCTCGGAGAGTGGGAAGCTTTGGGGTGGCCGGT TTGTGGGTGCAGTGGACCCCATCATGGAGAAGTTCAACGCGTCCATTGCC TACGACCGGCACCTTTGGGAGGTGGATGTTCAAGGCAGCAAAGCCTACAG CAGGGGCCTGGAGAAGGCAGGGCTCCTCACCAAGGCCGAGATGGACCAGA TACTCCATGGCCTAGACAAGGTGGCTGAGGAGTGGGCCAGGGCACCTTC AACTGAACTCCAATGATGAGGACATCCACACAGCCAATGAGCGCCGCCT GAAGGAGCTCATTGGTGCAACGGCAGGGAAGCTGCACACGGGACGGAGCC GGAATGACCAGGTGGTTCACAGACCTCAGGCTGTGGATGCGGCAGACCTGC TCCACGCTCTCGGGCCTCCTCTGGGAGCTCATTAGGACCATGGTGGATCG GGCAGAGGCGGAACGTGATGTTCTCTTCCCGGGGTACACCCATTTGCAGA GGGCCAGCCCATCCGCTGGAGCCACTGGATTCTGAGCCACGCCGTGGCA CTGACCCGAGACTCTGAGCGGCTGCTGGAGGTGCGGAAGCGGATCAATGT CCTGCCCCCTGGGGAGTGGGGCCATTGCAGGCAATCCCCTGGGTGTGGACC GAGAGCTGCTCCGAGCAGAACTCAACTTTGGGGCCATCACTCTCAACAGC ATGGATGCCACTAGTGAGCGGGACTTTGTGGCCGAGTTCCTGTTCTGGGC TTCGCTGTGCATGACCCATCTCAGCAGGATGGCCGAGGACCTCATCCTCT ACTGCACCAAGGAATTGAGCTTCGTGCAGCTCTCAGATGCCTACAGCACG GGAAGCAGCCTGATGCCCCAGAAGAAAAACCCGACAGTTTGGAGCTGAT CCGGAGCAAGGCTGGGCGTGTGTTTGGGCGGGAGGACAAGGAAGCTGTGT TTGAAGTGTGAGACACTATGAGTGCCGTGCTCCAGGTGGCCACTGGCGTC ATCTCTACGCTGCAGATTACCAAGAGAACATGGGACAGGCTCTCAGCCC CGACATGCTGGCCACTGACCTTGCCATTACCTGGTCCGCAAAGGGATGC CATTCGCCAGGCCACGAGGCCTCCGGGAAAGCTGTGTTTCATGGCCGAG

	ACCAAGGGGGTTCGCCCTCAACCAGCTGTCACTGCAGGAGCTGCAGACCAT CAGCCCCCTGTTCTCGGGCGACGTGATCTGCGTGTGGGACTACGGGCACA GTGTGGAGCAGTATGGTGCCCTGGGCGGCACTGCGCGCTCCAGCGTCGAC TGGCAGATCCGCCAGGTGCGGGCGCTACTGCAGGCACAGCAGGCCTAGGT CCTCCCACACCTGCCCCCTAATAAAGTGGGCGCGAGAGGAGGCTGCTGTG TGTTTCCTGCCCCAGCCTGGCTCCCTCGTTGCTGGGCTTTTCGGGGCTGGC CAGTGGGGACAGTCAGGGACTGGAGAGGCAGGGCAGGGTGGCCTGTAATC CCAGCACTTTGGAAGGGCAAGGTGCGAGGATGCTTGAGGCCAGGAGTTTG ACACAGCCTGGGCAACACAGGGAGACCCCCATCTCTACTCAATAATAAAA CAAATAGCCTGGCGTGGTGGCCCATGCATATAGTCCCAGCTACTTGTAAG GCTGAGGTGAGAGGACACTTGTGCCAGGAGTGGAGGCTGCAGTGAGCTA TGATCACGCCACTGCATTCCAGCCTGGATAACAGAGTGAGAACCTATCTC TAAAAATAAATAAATAAACGAAAAATAAATGGAAGCGGAGAAAACCTGGGC AAGGGCAATGAGAGTCGTAGACGGGAAGGGAAGAGGGGGCTCCCATCATA GCCTCTGCTCTGTCCAGGCCCCATCCCTTCAGACAGGGTATCACAGTGA CCTCCTAGGCCAGGAGCGATGGCTCACGCCTGTAATCCTAGCATTTTGGG AGGCCTGGGCAACAGGAAGACTCCAACCTACCATATTAAAAAAAACATGAA TGAAAGCAAAAA
44	AGTGGCGGGCGCGACACTATCCGTGCGGCCAGGCGGAGGTGAGTGCGCGG CGGCCGATGGGCGGGACGGGCGTGAGGACGCCGAGACCGTGGCGCGC GCTCACGTCCGCGTCCCCAAGGGCTGCGCTCCCTCAAGCGCAGTGCCAG AACTCGGAGCCAGCCCGGCCCGGGGGACCCTGCTGGCCAAGGAGGTGCTC AGTCCGGTCTTGTCTTCCAGACCCGAGGACCGAAGCTTCCGGACGACGA GGAACCGCCCAACATGGCCTCGGAGAGTGGGAAGCTTTGGGGTGGCCGGT TTGTGGGTGCAGTGGACCCCATCATGGAGAAGTTCAACGCGTCCATTGCC TACGACCGGCACCTTTGGGAGGTGGATGTTCAAGGCAGCAAAGCCTACAG CAGGGGCCTGGAGAAGGCAGGGCTCCTCACCAAGGCCGAGATGGACCAGA TACTCCATGGCCTAGACAAGGTGGCTGAGGAGTGGGCCAGGGCACCTTC AACTGAACCTCAATGATGAGGACATCCACACAGCCAATGAGCGCCGCCT GAAGGAGCTCATTGGTGCAACGGCAGGGAAGCTGCACACGGGACGGAGCC GGAATGACCAGGTGGTCACAGACCTCAGGCTGTGGATGCGGCAGACCTGC TCCACGCTCTCGGGCCTCCTCTGGGAGCTCATTAGGACCATGGTGGATCG GGCAGAGGCGGAACGTGATGTTCTCTTCCCGGGGTACACCCATTTGCAGA GGGCCAGCCCATCCGCTGGAGCCACTGGATTCTGAGTGGGGCCATTGCA GGCAATCCCCTGGGTGTGGACCGAGAGCTGCTCCGAGCAGAACTCAACTT TGGGGCCATCACTCTCAACAGCATGGATGCCACTAGTGAGCGGGACTTTG TGGCCGAGTTCCTGTTCTGGGCTTCGCTGTGCATGACCCATCTCAGCAGG ATGGCCGAGGACCTCATCCTCTACTGCACCAAGGAATTCAGCTTCGTGCA GCTCTCAGATGCCTACAGCACGGGAAGCAGCCTGATGCCCCAGAAGAAAA ACCCCGACAGTTTGGAGCTGATCCGGAGCAAGGCTGGGCGTGTGTTTGGG CGGTGTGCCGGGCTCCTGATGACCCTCAAGGGACTTCCCAGCACCTACAA CAAAGACTTACAGGAGGACAAGGAAGCTGTGTTTGAAGTGTGAGACACTA TGAGTGCCGTGCTCCAGGTGGCCACTGGCGTCATCTCTACGCTGCAGATT CACCAAGAGAACATGGGACAGGCTCTCAGCCCCGACATGCTGGCCACTGA CCTTGCCATTATTACCTGGTCCGCAAAGGGATGCCATTCCGCCAGGCCCACG AGGCCTCCGGGAAAGCTGTGTTTATGGCCGAGACCAAGGGGGTCGCCCTC AACCAGCTGTCACTGCAGGAGCTGCAGACCATCAGCCCCCTGTTCTCGGG CGACGTGATCTGCGTGTGGGACTACGGGCACAGTGTGGAGCAGTATGGTG CCCTGGGCGGCACTGCGCGCTCCAGCGTCGACTGGCAGATCCGCCAGGTG CGGGCGCTACTGCAGGCACAGCAGGCCTAGGTCTCCACACCTGCCCCC TAATAAAGTGGGCGCGAGAGGAGGCTGCTGTGTGTTTCTGCCCCAGCCT

	GGCTCCCTCGTTGCTGGGCTTTCTGGGGCTGGCCAGTGGGGACAGTCAGGG ACTGGAGAGGCAGGGCAGGGTGGCCTGTAATCCCAGCACTTTGGAAGGGC AAGGTGCGAGGATGCTTGAGGCCAGGAGTTTGACACAGCCTGGGCAACAC AGGGAGACCCCCATCTCTACTCAATAATAAAACAAATAGCCTGGCGTGGT GGCCCATGCATATAGTCCCAGCTACTTGTAAGGCTGAGGTGAGAGGACAC TTGTGCCCAGGAGTGGAGGCTGCAGTGAGCTATGATCACGCCACTGCATT CCAGCCTGGATAACAGAGTGAGAACCTATCTCTAAAAATAAATAAATAAA CGAAAAATAAATGGAAGCGGAGAAAACCTGGGCAAGGGCAATGAGAGTCGT AGACGGGAAGGGAAGAGGGGGCTCCCATCATAGCCTCTGCTCTGTCCAGG CCCCATCCCCTTCAGACAGGGTATCACAGTGACCTCCTAGGCCAGGAGCG ATGGCTCACGCCTGTAATCCTAGCATTTTGGGAGGCCTGGGCAACAGGAA GACTCCAAC TACCATATTAAAAAAAACATGAATGAAAGCAAAAA
--	---

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises from about 1,000 to about 100,000 nucleotides (*e.g.*, from 1,000 to 2,500, from 1,000 to 2,600, from 1,000 to 2,700, from 1,000 to 2,800, from 1,000 to 2,900, from 1,000 to 3,000, from 1,000 to 5,000, from 1,000 to 10,000, from 1,000 to 25,000, from 1,000 to 50,000, from 1,000 to 70,000, from 1,392 to 100,000, from 1,392 to 2,700, from 1,392 to 2,800, from 1,392 to 2,900, from 1,392 to 5,000, from 1,392 to 7,000, from 1,392 to 10,000, from 1,392 to 25,000, from 1,392 to 50,000, from 1,392 to 70,000, or from 1,392 to 100,000).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof), wherein the length of the nucleotide sequence (*e.g.*, an ORF) is at least 500 nucleotides in length (*e.g.*, at least or greater than about 500, 600, 700, 80, 900, 1,000, 1,050, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, 3,600, 3,700, 3,800, 3,900, 4,000, 4,100, 4,200, 4,300, 4,400, 4,500, 4,600, 4,635, 4,700, 4,800, 4,900, 5,000, 5,100, 5,200, 5,300, 5,400, 5,500, 5,600, 5,700, 5,800, 5,900, 6,000, 7,000, 8,000, 9,000, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 or up to and including 100,000 nucleotides).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a nucleotide sequence (*e.g.*, an ORF, *e.g.*, any one of SEQ ID NOs:20–33 or 41–44) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof according to any one of SEQ ID NOs:1–11) and further

comprises a 5'-UTR (*e.g.*, any one of SEQ ID NOs:50–79) and a 3'-UTR (*e.g.*, any one of SEQ ID NOs:100–139).

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azidoguanosine, Cap2, Cap4, 5' methylG cap, or an analog thereof).

In some embodiments, the polynucleotide (*e.g.*, an mRNA) comprises a polyA tail (*e.g.*, about 100 nucleotides in length). In some instances, the poly A tail is 50-150 (SEQ ID NO:197), 75-150 (SEQ ID NO:198), 85-150 (SEQ ID NO:199), 90-120 (SEQ ID NO:193), 90-130 (SEQ ID NO:194), or 90-150 (SEQ ID NO:192) nucleotides in length. In some instances, the poly A tail is 100 nucleotides in length (SEQ ID NO:195). In some instances, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises A100-UCUAG-A20-inverted deoxy-thymidine. In some instances, the poly A tail is A100-UCUAG-A20-inverted deoxy-thymidine.

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, any one of SEQ ID NOs:20–33 or 41-44) encoding an ASL polypeptide (*e.g.*, the wild-type sequence (SEQ ID NO:1), functional fragment, or variant thereof according to any one of SEQ ID NOs:2–11); 5'-UTR (*e.g.*, any one of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78); a 3'-UTR (*e.g.*, any one of SEQ ID NO:108, SEQ ID NO:111, SEQ ID NO:128, SEQ ID NO:137, SEQ ID NO:138, or SEQ ID NO:139); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azidoguanosine, Cap2, Cap4, 5' methylG cap, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:50); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog

thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

5 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:50); a 3'-UTR (*e.g.*, SEQ ID NO:128); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
10 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding
15 an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:50); a 3'-UTR (*e.g.*, SEQ ID NO:138); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-
20 thymidine (SEQ ID NO:211).

 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:78); a 3'-UTR (*e.g.*, SEQ ID NO:137); a 5' terminal (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof);
25 and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
30 an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:78); a 3'-UTR

(*e.g.*, SEQ ID NO:139); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:20) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:2), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:21) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:3), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:22) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:4), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:23) encoding

an ASL polypeptide variant (*e.g.*, SEQ ID NO:5), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:24) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:6), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:25) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:7), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:26) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:8), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:27) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:9), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
5 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
10 an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:28) encoding an ASL polypeptide variant (*e.g.*, SEQ ID NO:10), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
15 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:29) encoding
an ASL polypeptide variant (*e.g.*, SEQ ID NO:11), a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-
20 UTR (*e.g.*, SEQ ID NO:108); a 5' terminal (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some
instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-
thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
25 an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:30) encoding a full-length wild-type human ASL (*e.g.*, SEQ ID NO:1); a 5'-UTR (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
30 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:31) encoding a full-length wild-type human ASL (*e.g.*, SEQ ID NO:1); a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:111); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
5 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
10 an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:32) encoding a full-length wild-type human ASL (*e.g.*, SEQ ID NO:1); a 5'-UTR (*e.g.*, SEQ ID NO:55); a 3'-UTR (*e.g.*, SEQ ID NO:111); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
15 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:33) encoding a full-length wild-type human ASL (*e.g.*, SEQ ID NO:1); a 5'-UTR (*e.g.*, SEQ ID NO:55); a 3'-
20 UTR (*e.g.*, SEQ ID NO:108); a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*,
25 an mRNA) comprises: a nucleotide sequence (*e.g.*, an ORF, *e.g.*, SEQ ID NO:33) encoding a full length wild-type human ASL (*e.g.*, SEQ ID NO:1); a 5'-UTR (*e.g.*, SEQ ID NO:56); a 3'-UTR (*e.g.*, SEQ ID NO:108); a 5' terminal (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); and a polyA tail (*e.g.*, about 100 nucleotides in length (SEQ ID NO:195). In some
30 embodiments, the poly A tail is protected (*e.g.*, with an inverted deoxy-thymidine). In some instances, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

In any of the embodiments disclosed above, the polynucleotides of the present disclosure may further comprise, between the ORF and the 3'UTR, one or more stop codons, if not already present at the 5' terminus of the 3'UTR.

5 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof) further comprises at least one nucleic acid sequence that is noncoding, *e.g.*, a microRNA binding site.

10 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide is single stranded or double stranded.

In some embodiments, the polynucleotide of the present disclosure comprising a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof) is DNA or RNA. In some embodiments, the polynucleotide of the present disclosure is RNA. In some embodiments, the polynucleotide of the present disclosure is, or functions as, a mRNA. In some embodiments, the mRNA comprises a nucleotide sequence (*e.g.*, an ORF) that encodes at least one ASL polypeptide, and is capable of being translated to produce the encoded ASL polypeptide *in vitro*, *in vivo*, *in situ* or *ex vivo*.

20 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a sequence-optimized nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof, see *e.g.*, any one of SEQ ID NO:20–33), wherein the polynucleotide comprises at least one chemically modified nucleobase, *e.g.*, N1-methylpseudouracil or 5-methoxyuracil. In certain embodiments, all uracils in the polynucleotide are N1-methylpseudouracils. In other 25 embodiments, all uracils in the polynucleotide are 5-methoxyuracils. In some embodiments, the polynucleotide further comprises a miRNA binding site, *e.g.*, a miRNA binding site that binds to miR-142 and/or a miRNA binding site that binds to miR-126.

30 In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) disclosed herein is Formulated with a delivery agent comprising, *e.g.*, a compound having the Formula (I), *e.g.*, Compound I-1, Compound I-2, or Compound I-3; a compound having the Formula (III), (IV), (V), or (VI), *e.g.*, Compound I-VI or Compound P-I, or any combination thereof.

In some embodiments, the delivery agent comprises an ionizable amino lipid (*e.g.*, Compound I-1, Compound I-2, or Compound I-3), a helper lipid (phospholipid) (*e.g.*, DSPC or DOPE), a sterol (*e.g.*, Cholesterol), and a PEG lipid (*e.g.*, Compound P-I or PEG-DMG), *e.g.*, with a mole ratio in the range of about (i) 40-50 mol.% ionizable amino lipid (*e.g.*, Compound I-1, Compound I-2, or Compound I-3), optionally 45-50 mol.% ionizable amino lipid, for example, 45-46 mol.%, 46-47 mol.%, 47-48 mol.%, 48-49 mol.%, or 49-50 mol.% for example about 45 mol.%, 45.5 mol.%, 46 mol.%, 46.5 mol.%, 47 mol.%, 47.5 mol.%, 48 mol.%, 48.5 mol.%, 49 mol.%, or 49.5 mol.%; (ii) 30-45 mol.% sterol (*e.g.*, cholesterol), optionally 35-42 mol.% sterol, for example, 30-31 mol.%, 31-32 mol.%, 32-33 mol.%, 33-34 mol.%, 34-35 mol.%, 35-36 mol.%, 36-37 mol.%, 37-38 mol.%, 38-39 mol.%, or 39-40 mol.%, 40-42 mol.%, or 35-45 mol.% sterol; (iii) 5-15 mol.% helper lipid (*e.g.*, DSPC or DOPE), optionally 10-15 mol.% helper lipid, for example, 5-6 mol.%, 6-7 mol.%, 7-8 mol.%, 8-9 mol.%, 9-10 mol.%, 10-11 mol.%, 11-12 mol.%, 12-13 mol.%, 13-14 mol.%, 14-15 mol.%, 8-15 mol.%, or 8-12 mol.% helper lipid (*e.g.*, phospholipid); and (iv) 1-5% PEG lipid (*e.g.*, Compound P-I or PEG-DMG), optionally 1-5 mol.% PEG lipid, for example 1.5 to 2.5 mol.%, 1.5-3.5 mol.%, 1-2 mol.%, 2-3 mol.%, 3-4 mol.%, or 4-5 mol.% PEG lipid.

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) disclosed herein is Formulated with a delivery agent comprising LNP-1A, LNP 1-B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-1A. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-1B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-2A. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-2B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-3A. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:50), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195). In some embodiments, the poly A tail comprises or consists of A100-UCUAG-A20-inverted deoxy-

thymidine (SEQ ID NO:211) wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

5 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:50), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some
10 embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:50), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID
15 NO:128), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that
20 comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:50), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID NO:138), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some
25 embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:78), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID
30 NO:137), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some

embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:78), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID NO:139), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:2, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:3, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:4, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:5, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all
5 uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a
10 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:6, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a
15 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:7, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some
20 embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a
25 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:8, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that
30 comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:9, a 3'UTR (*e.g.*, SEQ ID

NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

5 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:10, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some
10 embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:11, a 3'UTR (*e.g.*, SEQ ID
15 NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that
20 comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:55), an ORF sequence of SEQ ID NO:1, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some
25 embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

 In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:1, a 3'UTR (*e.g.*, SEQ ID
30 NO:111), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some

embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:55), an ORF sequence of SEQ ID NO:1, a 3'UTR (*e.g.*, SEQ ID NO:111), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:55), an ORF sequence of SEQ ID NO:1, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the polynucleotide of the disclosure is an mRNA that comprises, from 5' end to 3' end: a 5'-terminal cap (*e.g.*, Cap1, *e.g.*, m⁷Gp-ppGm-A), a 5'UTR (*e.g.*, SEQ ID NO:56), an ORF sequence of SEQ ID NO:1, a 3'UTR (*e.g.*, SEQ ID NO:108), and a poly A tail (*e.g.*, about 100 nt in length, *e.g.*, SEQ ID NO:195), wherein all uracils in the polynucleotide are N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In some embodiments, the uracil or thymine content of the ORF relative to the theoretical minimum uracil or thymine content of a nucleotide sequence encoding the ASL polypeptide (%U_{TM} or %T_{TM}), is between about 100% and about 150%.

In some embodiments, the polynucleotides, compositions or formulations disclosed herein are used to treat and/or prevent ASL-related diseases, disorders or conditions, *e.g.*, argininosuccinic aciduria.

4. Signal Sequences

The polynucleotides (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure can also comprise nucleotide sequences that encode additional features that facilitate trafficking of the encoded polypeptides to therapeutically relevant sites. One such feature that aids in protein trafficking is the signal sequence, or targeting sequence. The peptides encoded by these signal sequences are known by a variety of names, including targeting peptides, transit peptides, and signal peptides. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprises a nucleotide sequence (*e.g.*, an ORF) that encodes a signal peptide operably linked to a nucleotide sequence that encodes an ASL polypeptide described herein.

In some embodiments, the "signal sequence" or "signal peptide" is a polynucleotide or polypeptide, respectively, which is from about 30-210, *e.g.*, about 45-80 or 15-60 nucleotides (*e.g.*, about 20, 30, 40, 50, 60, or 70 amino acids) in length that, optionally, is incorporated at the 5' (or N-terminus) of the coding region or the polypeptide, respectively. Addition of these sequences results in trafficking the encoded polypeptide to a desired site, such as the endoplasmic reticulum or the mitochondria through one or more targeting pathways. Some signal peptides are cleaved from the protein, for example by a signal peptidase after the proteins are transported to the desired site.

In some embodiments, the polynucleotide of the present disclosure comprises a nucleotide sequence encoding an ASL polypeptide, wherein the nucleotide sequence further comprises a 5' nucleic acid sequence encoding a heterologous signal peptide.

5. Sequence Optimization of Nucleotide Sequence Encoding an ASL Polypeptide

The polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure is sequence optimized. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide, optionally, a nucleotide sequence (*e.g.*, an ORF) encoding another polypeptide of interest, a 5'-UTR, a 3'-UTR, the 5' UTR or 3' UTR optionally comprising at least one microRNA binding site, optionally a nucleotide sequence encoding a linker, a polyA tail, or any combination thereof), in which the ORF(s) are sequence optimized.

A sequence-optimized nucleotide sequence, *e.g.*, a codon-optimized mRNA sequence encoding an ASL polypeptide, is a sequence comprising at least one synonymous nucleobase

substitution with respect to a reference sequence (*e.g.*, a wild type nucleotide sequence encoding an ASL polypeptide).

A sequence-optimized nucleotide sequence can be partially or completely different in sequence from the reference sequence. For example, a reference sequence encoding polyserine uniformly encoded by UCU codons can be sequence-optimized by having 100% of its nucleobases substituted (for each codon, U in position 1 replaced by A, C in position 2 replaced by G, and U in position 3 replaced by C) to yield a sequence encoding polyserine which would be uniformly encoded by AGC codons. The percentage of sequence identity obtained from a global pairwise alignment between the reference polyserine nucleic acid sequence and the sequence-optimized polyserine nucleic acid sequence would be 0%. However, the protein products from both sequences would be 100% identical.

Some sequence optimization (also sometimes referred to codon optimization) methods are known in the art (and discussed in more detail below) and can be useful to achieve one or more desired results. These results can include, *e.g.*, matching codon frequencies in certain tissue targets and/or host organisms to ensure proper folding; biasing G/C content to increase mRNA stability or reduce secondary structures; minimizing tandem repeat codons or base runs that can impair gene construction or expression; customizing transcriptional and translational control regions; inserting or removing protein trafficking sequences; removing/adding post translation modification sites in an encoded protein (*e.g.*, glycosylation sites); adding, removing or shuffling protein domains; inserting or deleting restriction sites; modifying ribosome binding sites and mRNA degradation sites; adjusting translational rates to allow the various domains of the protein to fold properly; and/or reducing or eliminating problem secondary structures within the polynucleotide. Sequence optimization tools, algorithms and services are known in the art, non-limiting examples include services from GeneArt (Life Technologies), DNA2.0 (Menlo Park CA) and/or proprietary methods.

Codon options for each amino acid are given in **Table 2**.

TABLE 2. Codon Options

Amino Acid	Single Letter Code	Codon Options
Isoleucine	I	AUU, AUC, AUA
Leucine	L	CUU, CUC, CUA, CUG, UUA, UUG

Valine	V	GUU, GUC, GUA, GUG
Phenylalanine	F	UUU, UUC
Methionine	M	AUG
Cysteine	C	UGU, UGC
Alanine	A	GCU, GCC, GCA, GCG
Glycine	G	GGU, GGC, GGA, GGG
Proline	P	CCU, CCC, CCA, CCG
Threonine	T	ACU, ACC, ACA, ACG
Serine	S	UCU, UCC, UCA, UCG, AGU, AGC
Tyrosine	Y	UAU, UAC
Tryptophan	W	UGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAU, AAC
Histidine	H	CAU, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAU, GAC
Lysine	K	AAA, AAG
Arginine	R	CGU, CGC, CGA, CGG, AGA, AGG
Selenocysteine	Sec	UGA in mRNA in presence of Selenocysteine insertion element (SECIS)
Stop codons	Stop	UAA, UAG, UGA

In some embodiments, a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a sequence-optimized nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide, a functional fragment, or a variant thereof, wherein the ASL polypeptide, functional fragment, or a variant thereof encoded by the sequence-optimized nucleotide sequence has improved properties (*e.g.*, compared to an ASL polypeptide, functional fragment, or a variant thereof encoded by a reference nucleotide sequence that is not sequence optimized), *e.g.*, improved properties related to expression efficacy after administration *in vivo*. Such properties include, but are not limited to, improving nucleic acid stability (*e.g.*, mRNA stability), increasing translation efficacy in the target tissue, reducing the number of truncated proteins expressed, improving the folding or prevent misfolding of the expressed proteins, reducing toxicity of the expressed products, reducing cell death caused by the expressed products, increasing and/or decreasing protein aggregation.

In some embodiments, the sequence-optimized nucleotide sequence (*e.g.*, an ORF) is codon optimized for expression in human subjects, having structural and/or chemical features that avoid one or more of the problems in the art, for example, features which are useful for optimizing formulation and delivery of nucleic acid-based therapeutics while retaining structural and functional integrity; overcoming a threshold of expression; improving expression rates; half-life and/or protein concentrations; optimizing protein localization; and

avoiding deleterious bio-responses such as the immune response and/or degradation pathways.

In some embodiments, the polynucleotides of the present disclosure comprise a nucleotide sequence (*e.g.*, a nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide, a nucleotide sequence (*e.g.*, an ORF) encoding another polypeptide of interest, a 5'-UTR, a 3'-UTR, a microRNA binding site, a nucleic acid sequence encoding a linker, or any combination thereof) that is sequence-optimized according to a method comprising:

(i) substituting at least one codon in a reference nucleotide sequence (*e.g.*, an ORF encoding an ASL polypeptide) with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence;

(ii) substituting at least one codon in a reference nucleotide sequence (*e.g.*, an ORF encoding an ASL polypeptide) with an alternative codon having a higher codon frequency in the synonymous codon set;

(iii) substituting at least one codon in a reference nucleotide sequence (*e.g.*, an ORF encoding an ASL polypeptide) with an alternative codon to increase G/C content; or

(iv) a combination thereof.

In some embodiments, the sequence-optimized nucleotide sequence (*e.g.*, an ORF encoding an ASL polypeptide) has at least one improved property with respect to the reference nucleotide sequence.

In some embodiments, the sequence optimization method is multiparametric and comprises one, two, three, four, or more methods disclosed herein and/or other optimization methods known in the art.

Features, which can be considered beneficial in some embodiments of the present disclosure, can be encoded by or within regions of the polynucleotide and such regions can be upstream (5') to, downstream (3') to, or within the region that encodes the ASL polypeptide. These regions can be incorporated into the polynucleotide before and/or after sequence-optimization of the protein encoding region or open reading frame (ORF). Examples of such features include, but are not limited to, untranslated regions (UTRs), microRNA sequences,

Kozak sequences, oligo(dT) sequences, poly-A tail, and detectable tags and can include multiple cloning sites that can have XbaI recognition.

In some embodiments, the polynucleotide of the present disclosure comprises a 5' UTR, a 3' UTR and/or a microRNA binding site. In some embodiments, the polynucleotide comprises two or more 5' UTRs and/or 3' UTRs, which can be the same or different sequences. In some embodiments, the polynucleotide comprises two or more microRNA binding sites, which can be the same or different sequences. Any portion of the 5' UTR, 3' UTR, and/or microRNA binding site, including none, can be sequence-optimized and can independently contain one or more different structural or chemical modifications, before and/or after sequence optimization.

In some embodiments, after optimization, the polynucleotide is reconstituted and transformed into a vector such as, but not limited to, plasmids, viruses, cosmids, and artificial chromosomes. For example, the optimized polynucleotide can be reconstituted and transformed into chemically competent *E. coli*, yeast, neurospora, maize, drosophila, etc. where high copy plasmid-like or chromosome structures occur by methods described herein.

6. Sequence-Optimized Nucleotide Sequences Encoding ASL Polypeptides

In some embodiments, the polynucleotide of the present disclosure comprises a sequence-optimized nucleotide sequence encoding an ASL polypeptide disclosed herein. In some embodiments, the polynucleotide of the present disclosure comprises an open reading frame (ORF) encoding an ASL polypeptide, wherein the ORF has been sequence optimized.

An exemplary sequence-optimized nucleotide sequence encoding human full length ASL is set forth as SEQ ID NO:30. Another exemplary sequence-optimized nucleotide sequence encoding human full length ASL is set forth as SEQ ID NO:31. Another exemplary sequence-optimized nucleotide sequence encoding human full length ASL is set forth as SEQ ID NO:32. Another exemplary sequence-optimized nucleotide sequence encoding human full length ASL is set forth as SEQ ID NO:33. In some embodiments, the sequence optimized ASL sequence, fragment, and variant thereof are used to practice the methods disclosed herein.

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

- (i) a 5' cap such as provided herein, for example, m⁷Gp-ppGm-A or Cap1;
- 5 (ii) a 5' UTR, such as the sequences provided herein, for example, SEQ ID NO:58;
- (iii) an open reading frame encoding an ASL polypeptide, *e.g.*, a sequence optimized nucleic acid sequence encoding ASL set forth as SEQ ID NO:2;
- (iv) at least one stop codon (if not present at 5' terminus of 3'UTR);
- (v) a 3' UTR, such as the sequences provided herein, for example, SEQ ID NO:108;
- 10 and
- (vi) a poly-A tail provided above (*e.g.*, SEQ ID NO:195).

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

- 15 (i) a 5' cap such as provided herein, for example, m⁷Gp-ppGm-A or Cap 1;
- (ii) a 5' UTR, such as the sequences provided herein, for example, one of SEQ ID NOs:50-79;
- (iii) an open reading frame encoding an ASL polypeptide, *e.g.*, a sequence optimized nucleic acid sequence encoding ASL set forth as any one of SEQ ID NOs:2-11;
- 20 (iv) at least one stop codon (if not present at 5' terminus of 3'UTR);
- (v) a 3' UTR, such as the sequences provided herein, for example, one of SEQ ID NOs:100-136; and
- (vi) a poly-A tail provided above (*e.g.*, SEQ ID NO:195).

In certain embodiments, all uracils in the polynucleotide are N1-methylpseudouracil (G5). In
25 certain embodiments, all uracils in the polynucleotide are 5-methoxyuracil (G6).

The sequence-optimized nucleotide sequences disclosed herein are distinct from the corresponding wild type nucleotide acid sequences and from other known sequence-optimized nucleotide sequences, *e.g.*, these sequence-optimized nucleic acids have unique compositional characteristics.

30 In some embodiments, the percentage of uracil or thymine nucleobases in a sequence-optimized nucleotide sequence (*e.g.*, encoding an ASL polypeptide, a functional fragment, or

a variant thereof) is modified (*e.g.*, reduced) with respect to the percentage of uracil or thymine nucleobases in the reference wild-type nucleotide sequence. Such a sequence is referred to as a uracil-modified or thymine-modified sequence. The percentage of uracil or thymine content in a nucleotide sequence can be determined by dividing the number of uracils or thymines in a sequence by the total number of nucleotides and multiplying by 100. In some embodiments, the sequence-optimized nucleotide sequence has a lower uracil or thymine content than the uracil or thymine content in the reference wild-type sequence. In some embodiments, the uracil or thymine content in a sequence-optimized nucleotide sequence of the present disclosure is greater than the uracil or thymine content in the reference wild-type sequence and still maintain beneficial effects, *e.g.*, increased expression and/or reduced Toll-Like Receptor (TLR) response when compared to the reference wild-type sequence.

Methods for optimizing codon usage are known in the art. For example, an ORF of any one or more of the sequences provided herein may be codon optimized. Codon optimization, in some embodiments, may be used to match codon frequencies in target and host organisms to ensure proper folding; bias GC content to increase mRNA stability or reduce secondary structures; minimize tandem repeat codons or base runs that may impair gene construction or expression; customize transcriptional and translational control regions; insert or remove protein trafficking sequences; remove/add post translation modification sites in encoded protein (*e.g.*, glycosylation sites); add, remove or shuffle protein domains; insert or delete restriction sites; modify ribosome binding sites and mRNA degradation sites; adjust translational rates to allow the various domains of the protein to fold properly; or reduce or eliminate problem secondary structures within the polynucleotide. Codon optimization tools, algorithms and services are known in the art - non-limiting examples include services from GeneArt (Life Technologies), DNA2.0 (Menlo Park CA) and/or proprietary methods. In some embodiments, the open reading frame (ORF) sequence is optimized using optimization algorithms.

7. Characterization of Sequence Optimized Nucleic Acids

In some embodiments of the present disclosure, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a sequence optimized nucleic acid disclosed herein encoding an ASL polypeptide can be tested to determine whether at least one nucleic acid sequence property

(*e.g.*, stability when exposed to nucleases) or expression property has been improved with respect to the non-sequence optimized nucleic acid.

As used herein, "expression property" refers to a property of a nucleic acid sequence either *in vivo* (*e.g.*, translation efficacy of a synthetic mRNA after administration to a subject in need thereof) or *in vitro* (*e.g.*, translation efficacy of a synthetic mRNA tested in an *in vitro* model system). Expression properties include but are not limited to the amount of protein produced by an mRNA encoding an ASL polypeptide after administration, and the amount of soluble or otherwise functional protein produced. In some embodiments, sequence optimized nucleic acids disclosed herein can be evaluated according to the viability of the cells expressing a protein encoded by a sequence optimized nucleic acid sequence (*e.g.*, a RNA, *e.g.*, an mRNA) encoding an ASL polypeptide disclosed herein.

In a given embodiment, a plurality of sequence optimized nucleic acids disclosed herein (*e.g.*, a RNA, *e.g.*, an mRNA) containing codon substitutions with respect to the non-optimized reference nucleic acid sequence can be characterized functionally to measure a property of interest, for example an expression property in an *in vitro* model system, or *in vivo* in a target tissue or cell.

a. Optimization of Nucleic Acid Sequence Intrinsic Properties

In some embodiments of the present disclosure, the desired property of the polynucleotide is an intrinsic property of the nucleic acid sequence. For example, the nucleotide sequence (*e.g.*, a RNA, *e.g.*, an mRNA) can be sequence optimized for *in vivo* or *in vitro* stability. In some embodiments, the nucleotide sequence can be sequence optimized for expression in a given target tissue or cell. In some embodiments, the nucleic acid sequence is sequence optimized to increase its plasma half-life by preventing its degradation by endo and exonucleases.

In other embodiments, the nucleic acid sequence is sequence optimized to increase its resistance to hydrolysis in solution, for example, to lengthen the time that the sequence optimized nucleic acid or a pharmaceutical composition comprising the sequence optimized nucleic acid can be stored under aqueous conditions with minimal degradation.

In other embodiments, the sequence optimized nucleic acid can be optimized to increase its resistance to hydrolysis in dry storage conditions, for example, to lengthen the

time that the sequence optimized nucleic acid can be stored after lyophilization with minimal degradation.

b. Nucleic Acids Sequence Optimized for Protein Expression

In some embodiments of the present disclosure, the desired property of the polynucleotide is the level of expression of an ASL polypeptide encoded by a sequence optimized sequence disclosed herein. Protein expression levels can be measured using one or more expression systems. In some embodiments, expression can be measured in cell culture systems, *e.g.*, CHO cells or HEK293 cells. In some embodiments, expression can be measured using *in vitro* expression systems prepared from extracts of living cells, *e.g.*, rabbit reticulocyte lysates, or *in vitro* expression systems prepared by assembly of purified individual components. In other embodiments, the protein expression is measured in an *in vivo* system, *e.g.*, mouse, rabbit, monkey, etc.

In some embodiments, protein expression in solution form can be desirable. Accordingly, in some embodiments, a reference sequence can be sequence optimized to yield a sequence optimized nucleic acid sequence having optimized levels of expressed proteins in soluble form. Levels of protein expression and other properties such as solubility, levels of aggregation, and the presence of truncation products (*i.e.*, fragments due to proteolysis, hydrolysis, or defective translation) can be measured according to methods known in the art, for example, using electrophoresis (*e.g.*, native or SDS-PAGE) or chromatographic methods (*e.g.*, HPLC, size exclusion chromatography, etc.).

c. Optimization of Target Tissue or Target Cell Viability

In some embodiments, the expression of heterologous therapeutic proteins encoded by a nucleic acid sequence can have deleterious effects in the target tissue or cell, reducing protein yield, or reducing the quality of the expressed product (*e.g.*, due to the presence of protein fragments or precipitation of the expressed protein in inclusion bodies), or causing toxicity.

Accordingly, in some embodiments of the present disclosure, the sequence optimization of a nucleic acid sequence disclosed herein, *e.g.*, a nucleic acid sequence encoding an ASL polypeptide, can be used to increase the viability of target cells expressing the protein encoded by the sequence optimized nucleic acid.

Heterologous protein expression can also be deleterious to cells transfected with a nucleic acid sequence for autologous or heterologous transplantation. Accordingly, in some embodiments of the present disclosure the sequence optimization of a nucleic acid sequence disclosed herein can be used to increase the viability of target cells expressing the protein encoded by the sequence optimized nucleic acid sequence. Changes in cell or tissue viability, toxicity, and other physiological reaction can be measured according to methods known in the art.

d. Reduction of Immune and/or Inflammatory Response

In some cases, the administration of a sequence optimized nucleic acid encoding ASL polypeptide or a functional fragment thereof can trigger an immune response, which could be caused by (i) the therapeutic agent (*e.g.*, an mRNA encoding an ASL polypeptide), or (ii) the expression product of such therapeutic agent (*e.g.*, the ASL polypeptide encoded by the mRNA), or (iv) a combination thereof. Accordingly, in some embodiments of the present disclosure the sequence optimization of nucleic acid sequence (*e.g.*, an mRNA) disclosed herein can be used to decrease an immune or inflammatory response triggered by the administration of a nucleic acid encoding an ASL polypeptide or by the expression product of ASL encoded by such nucleic acid.

In some cases, an inflammatory response can be measured by detecting increased levels of one or more inflammatory cytokines using methods known in the art, *e.g.*, ELISA.

The term "inflammatory cytokine" refers to cytokines that are elevated in an inflammatory response. Examples of inflammatory cytokines include interleukin-6 (IL-6), CXCL1 (chemokine (C-X-C motif) ligand 1; also known as GRO α , interferon- γ (IFN γ), tumor necrosis factor α (TNF α), interferon γ -induced protein 10 (IP-10), or granulocyte-colony stimulating factor (G-CSF). The term inflammatory cytokines includes also other cytokines associated with inflammatory responses known in the art, *e.g.*, interleukin-1 (IL-1), interleukin-8 (IL-8), interleukin-12 (IL-12), interleukin-13 (IL-13), interferon α (IFN- α), etc.

8. Modified Nucleotide Sequences Encoding ASL Polypeptides

In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) of the present disclosure comprises a chemically modified nucleobase, for example, a chemically modified uracil, *e.g.*, pseudouracil, N1-methylpseudouracil, 5-methoxyuracil, or the like. In some embodiments, the mRNA is a uracil-modified sequence comprising an ORF encoding

an ASL polypeptide, wherein the mRNA comprises a chemically modified nucleobase, for example, a chemically modified uracil, *e.g.*, pseudouracil, N1-methylpseudouracil, or 5-methoxyuracil.

In certain aspects of the present disclosure, when the modified uracil base is
5 connected to a ribose sugar, as it is in polynucleotides, the resulting modified nucleoside or nucleotide is referred to as modified uridine. In some embodiments, uracil in the polynucleotide is at least about 25%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least 90%, at least 95%, at least 99%, or about 100% modified uracil. In some embodiments, uracil in the
10 polynucleotide is at least 95% modified uracil. In some embodiments, uracil in the polynucleotide is 100% modified uracil.

In embodiments where uracil in the polynucleotide is at least 95% modified uracil overall uracil content can be adjusted such that an mRNA provides suitable protein expression levels while inducing little to no immune response. In some embodiments, the
15 uracil content of the ORF is between about 100% and about 150%, between about 100% and about 110%, between about 105% and about 115%, between about 110% and about 120%, between about 115% and about 125%, between about 120% and about 130%, between about 125% and about 135%, between about 130% and about 140%, between about 135% and about 145%, between about 140% and about 150% of the theoretical minimum uracil content
20 in the corresponding wild-type ORF (%U_{TM}). In other embodiments, the uracil content of the ORF is between about 121% and about 136% or between 123% and 134% of the %U_{TM}. In some embodiments, the uracil content of the ORF encoding an ASL polypeptide is about 115%, about 120%, about 125%, about 130%, about 135%, about 140%, about 145%, or about 150% of the %U_{TM}. In this context, the term "uracil" can refer to modified uracil
25 and/or naturally occurring uracil.

In some embodiments, the uracil content in the ORF of the mRNA encoding an ASL polypeptide of the present disclosure is less than about 30%, about 25%, about 20%, about 15%, or about 10% of the total nucleobase content in the ORF. In some embodiments, the uracil content in the ORF is between about 10% and about 20% of the total nucleobase
30 content in the ORF. In other embodiments, the uracil content in the ORF is between about 10% and about 25% of the total nucleobase content in the ORF. In some embodiments, the uracil content in the ORF of the mRNA encoding an ASL polypeptide is less than about 20%

of the total nucleobase content in the open reading frame. In this context, the term "uracil" can refer to modified uracil and/or naturally occurring uracil.

In further embodiments, the ORF of the mRNA encoding an ASL polypeptide having modified uracil and adjusted uracil content has increased Cytosine (C), Guanine (G), or Guanine/Cytosine (G/C) content (absolute or relative). In some embodiments, the overall increase in C, G, or G/C content (absolute or relative) of the ORF is at least about 2%, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 10%, at least about 15%, at least about 20%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, or at least about 100% relative to the G/C content (absolute or relative) of the wild-type ORF. In some embodiments, the G, the C, or the G/C content in the ORF is less than about 100%, less than about 90%, less than about 85%, or less than about 80% of the theoretical maximum G, C, or G/C content of the corresponding wild type nucleotide sequence encoding the ASL polypeptide (%G_{TMX}, %C_{TMX}, or %G/C_{TMX}). In some embodiments, the increases in G and/or C content (absolute or relative) described herein can be conducted by replacing synonymous codons with low G, C, or G/C content with synonymous codons having higher G, C, or G/C content. In other embodiments, the increase in G and/or C content (absolute or relative) is conducted by replacing a codon ending with U with a synonymous codon ending with G or C.

In further embodiments, the ORF of the mRNA encoding an ASL polypeptide of the present disclosure comprises modified uracil and has an adjusted uracil content containing less uracil pairs (UU) and/or uracil triplets (UUU) and/or uracil quadruplets (UUUU) than the corresponding wild-type nucleotide sequence encoding the ASL polypeptide. In some embodiments, the ORF of the mRNA encoding an ASL polypeptide of the present disclosure contains no uracil pairs and/or uracil triplets and/or uracil quadruplets. In some embodiments, uracil pairs and/or uracil triplets and/or uracil quadruplets are reduced below a certain threshold, *e.g.*, no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 occurrences in the ORF of the mRNA encoding the ASL polypeptide. In a particular embodiment, the ORF of the mRNA encoding the ASL polypeptide of the present disclosure contains less than 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 non-phenylalanine uracil pairs and/or triplets. In some embodiments, the ORF of the

mRNA encoding the ASL polypeptide contains no non-phenylalanine uracil pairs and/or triplets.

In further embodiments, the ORF of the mRNA encoding an ASL polypeptide of the present disclosure comprises modified uracil and has an adjusted uracil content containing
5 less uracil-rich clusters than the corresponding wild-type nucleotide sequence encoding the ASL polypeptide. In some embodiments, the ORF of the mRNA encoding the ASL polypeptide of the present disclosure contains uracil-rich clusters that are shorter in length than corresponding uracil-rich clusters in the corresponding wild-type nucleotide sequence encoding the ASL polypeptide.

10 In further embodiments, alternative lower frequency codons are employed. At least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at
15 least about 99%, or 100% of the codons in the ASL polypeptide-encoding ORF of the modified uracil-comprising mRNA are substituted with alternative codons, each alternative codon having a codon frequency lower than the codon frequency of the substituted codon in the synonymous codon set. The ORF also has adjusted uracil content, as described above. In some embodiments, at least one codon in the ORF of the mRNA encoding the ASL
20 polypeptide is substituted with an alternative codon having a codon frequency lower than the codon frequency of the substituted codon in the synonymous codon set.

In some embodiments, the adjusted uracil content, ASL polypeptide-encoding ORF of the modified uracil-comprising mRNA exhibits expression levels of ASL when administered to a mammalian cell that are higher than expression levels of ASL from the corresponding
25 wild-type mRNA. In some embodiments, the mammalian cell is a mouse cell, a rat cell, or a rabbit cell. In other embodiments, the mammalian cell is a monkey cell or a human cell. In some embodiments, the human cell is a HeLa cell, a BJ fibroblast cell, or a peripheral blood mononuclear cell (PBMC). In some embodiments, ASL is expressed at a level higher than expression levels of ASL from the corresponding wild-type mRNA when the mRNA is
30 administered to a mammalian cell *in vivo*. In some embodiments, the mRNA is administered to mice, rabbits, rats, monkeys, or humans. In some embodiments, mice are null mice. In some embodiments, the mRNA is administered to mice in an amount of about 0.01 mg/kg,

about 0.05 mg/kg, about 0.1 mg/kg, or 0.2 mg/kg or about 0.5 mg/kg. In some embodiments, the mRNA is administered intravenously or intramuscularly. In other embodiments, the ASL polypeptide is expressed when the mRNA is administered to a mammalian cell *in vitro*. In some embodiments, the expression is increased by at least about 2-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 500-fold, at least about 1500-fold, or at least about 3000-fold. In other embodiments, the expression is increased by at least about 10%, about 20%, about 30%, about 40%, about 50%, 60%, about 70%, about 80%, about 90%, or about 100%.

In some embodiments, adjusted uracil content, ASL polypeptide-encoding ORF of the modified uracil-comprising mRNA exhibits increased stability. In some embodiments, the mRNA exhibits increased stability in a cell relative to the stability of a corresponding wild-type mRNA under the same conditions. In some embodiments, the mRNA exhibits increased stability including resistance to nucleases, thermal stability, and/or increased stabilization of secondary structure. In some embodiments, increased stability exhibited by the mRNA is measured by determining the half-life of the mRNA (*e.g.*, in a plasma, serum, cell, or tissue sample) and/or determining the area under the curve (AUC) of the protein expression by the mRNA over time (*e.g.*, *in vitro* or *in vivo*). An mRNA is identified as having increased stability if the half-life and/or the AUC is greater than the half-life and/or the AUC of a corresponding wild-type mRNA under the same conditions.

In some embodiments, the mRNA of the present disclosure induces a detectably lower immune response (*e.g.*, innate or acquired) relative to the immune response induced by a corresponding wild-type mRNA under the same conditions. In other embodiments, the mRNA of the present disclosure induces a detectably lower immune response (*e.g.*, innate or acquired) relative to the immune response induced by an mRNA that encodes for an ASL polypeptide but does not comprise modified uracil under the same conditions, or relative to the immune response induced by an mRNA that encodes for an ASL polypeptide and that comprises modified uracil but that does not have adjusted uracil content under the same conditions. The innate immune response can be manifested by increased expression of pro-inflammatory cytokines, activation of intracellular PRRs (RIG-I, MDA5, etc), cell death, and/or termination or reduction in protein translation. In some embodiments, a reduction in the innate immune response can be measured by expression or activity level of Type 1 interferons (*e.g.*, IFN- α , IFN- β , IFN- κ , IFN- δ , IFN- ϵ , IFN- τ , IFN- ω , and IFN- ζ) or the

expression of interferon-regulated genes such as the toll-like receptors (*e.g.*, TLR7 and TLR8), and/or by decreased cell death following one or more administrations of the mRNA of the present disclosure into a cell.

In some embodiments, the expression of Type-1 interferons by a mammalian cell in response to the mRNA of the present disclosure is reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or greater than 99.9% relative to a corresponding wild-type mRNA, to an mRNA that encodes an ASL polypeptide but does not comprise modified uracil, or to an mRNA that encodes an ASL polypeptide and that comprises modified uracil but that does not have adjusted uracil content. In some embodiments, the interferon is IFN- β . In some embodiments, cell death frequency caused by administration of mRNA of the present disclosure to a mammalian cell is 10%, 25%, 50%, 75%, 85%, 90%, 95%, or over 95% less than the cell death frequency observed with a corresponding wild-type mRNA, an mRNA that encodes for an ASL polypeptide but does not comprise modified uracil, or an mRNA that encodes for an ASL polypeptide and that comprises modified uracil but that does not have adjusted uracil content. In some embodiments, the mammalian cell is a BJ fibroblast cell. In other embodiments, the mammalian cell is a splenocyte. In some embodiments, the mammalian cell is that of a mouse or a rat. In other embodiments, the mammalian cell is that of a human. In some embodiments, the mRNA of the present disclosure does not substantially induce an innate immune response of a mammalian cell into which the mRNA is introduced.

9. Methods for Modifying Polynucleotides

The disclosure includes modified polynucleotides comprising a polynucleotide described herein (*e.g.*, a polynucleotide, *e.g.* mRNA, comprising a nucleotide sequence encoding an ASL polypeptide). The modified polynucleotides can be chemically modified and/or structurally modified. When the polynucleotides of the present disclosure are chemically and/or structurally modified the polynucleotides can be referred to as "modified polynucleotides."

The present disclosure provides for modified nucleosides and nucleotides of a polynucleotide (*e.g.*, RNA polynucleotides, such as mRNA polynucleotides) encoding an ASL polypeptide. A "nucleoside" refers to a compound containing a sugar molecule (*e.g.*, a pentose or ribose) or a derivative thereof in combination with an organic base (*e.g.*, a purine or pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). A

"nucleotide" refers to a nucleoside including a phosphate group. Modified nucleotides can be synthesized by any useful method, such as, for example, chemically, enzymatically, or recombinantly, to include one or more modified or non-natural nucleosides. Polynucleotides can comprise a region or regions of linked nucleosides. Such regions can have variable backbone linkages. The linkages can be standard phosphodiester linkages, in which case the polynucleotides would comprise regions of nucleotides.

The modified polynucleotides disclosed herein can comprise various distinct modifications. In some embodiments, the modified polynucleotides contain one, two, or more (optionally different) nucleoside or nucleotide modifications. In some embodiments, a modified polynucleotide, introduced to a cell can exhibit one or more desirable properties, *e.g.*, improved protein expression, reduced immunogenicity, or reduced degradation in the cell, as compared to an unmodified polynucleotide.

In some embodiments, a polynucleotide of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide) is structurally modified. As used herein, a "structural" modification is one in which two or more linked nucleosides are inserted, deleted, duplicated, inverted or randomized in a polynucleotide without significant chemical modification to the nucleotides themselves. Because chemical bonds will necessarily be broken and reformed to effect a structural modification, structural modifications are of a chemical nature and hence are chemical modifications. However, structural modifications will result in a different sequence of nucleotides. For example, the polynucleotide "ATCG" can be chemically modified to "AT-5meC-G". The same polynucleotide can be structurally modified from "ATCG" to "ATCCCG". Here, the dinucleotide "CC" has been inserted, resulting in a structural modification to the polynucleotide.

Therapeutic compositions of the present disclosure comprise, in some embodiments, at least one nucleic acid (*e.g.*, RNA) having an open reading frame encoding ASL, a fragment thereof, or a variant thereof (*e.g.*, SEQ ID NO:3), wherein the nucleic acid comprises nucleotides and/or nucleosides that can be standard (unmodified) or modified as is known in the art. In some embodiments, nucleotides and nucleosides of the present disclosure comprise modified nucleotides or nucleosides. Such modified nucleotides and nucleosides can be naturally-occurring modified nucleotides and nucleosides or non-naturally occurring modified nucleotides and nucleosides. Such modifications can include those at the sugar,

backbone, or nucleobase portion of the nucleotide and/or nucleoside as are recognized in the art.

In some embodiments, a naturally-occurring modified nucleotide or nucleotide of the disclosure is one as is generally known or recognized in the art. Non-limiting examples of such naturally occurring modified nucleotides and nucleotides can be found, inter alia, in the widely recognized MODOMICS database.

In some embodiments, a non-naturally occurring modified nucleotide or nucleoside of the disclosure is one as is generally known or recognized in the art. Non-limiting examples of such non-naturally occurring modified nucleotides and nucleosides can be found, inter alia, in published US Application Nos. PCT/US2012/058519; PCT/US2013/075177; PCT/US2014/058897; PCT/US2014/058891; PCT/US2014/070413; PCT/US2015/36773; PCT/US2015/36759; PCT/US2015/36771; or PCT/IB2017/051367 all of which are incorporated by reference herein.

In some embodiments, at least one RNA (*e.g.*, mRNA) of the present disclosure is not chemically modified and comprises the standard ribonucleotides consisting of adenosine, guanosine, cytosine and uridine. In some embodiments, nucleotides and nucleosides of the present disclosure comprise standard nucleoside residues such as those present in transcribed RNA (*e.g.* A, G, C, or U). In some embodiments, nucleotides and nucleosides of the present disclosure comprise standard deoxyribonucleosides such as those present in DNA (*e.g.* dA, dG, dC, or dT).

Hence, nucleic acids of the disclosure (*e.g.*, DNA nucleic acids and RNA nucleic acids, such as mRNA nucleic acids) can comprise standard nucleotides and nucleosides, naturally-occurring nucleotides and nucleosides, non-naturally-occurring nucleotides and nucleosides, or any combination thereof.

Nucleic acids of the disclosure (*e.g.*, DNA nucleic acids and RNA nucleic acids, such as mRNA nucleic acids), in some embodiments, comprise various (more than one) different types of standard and/or modified nucleotides and nucleosides. In some embodiments, a particular region of a nucleic acid contains one, two or more (optionally different) types of standard and/or modified nucleotides and nucleosides.

In some embodiments, a modified RNA nucleic acid (*e.g.*, a modified mRNA nucleic acid), introduced to a cell or organism, exhibits reduced degradation in the cell or organism, respectively, relative to an unmodified nucleic acid comprising standard nucleotides and nucleosides.

5 In some embodiments, a modified RNA nucleic acid (*e.g.*, a modified mRNA nucleic acid), introduced into a cell or organism, may exhibit reduced immunogenicity in the cell or organism, respectively (*e.g.*, a reduced innate response) relative to an unmodified nucleic acid comprising standard nucleotides and nucleosides.

10 Nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids), in some embodiments, comprise non-natural modified nucleotides that are introduced during synthesis or post-synthesis of the nucleic acids to achieve desired functions or properties. The modifications may be present on internucleotide linkages, purine or pyrimidine bases, or sugars. The modification may be introduced with chemical synthesis or with a polymerase enzyme at the terminal of a chain or anywhere else in the chain. Any of the regions of a
15 nucleic acid may be chemically modified.

 The present disclosure provides for modified nucleosides and nucleotides of a nucleic acid (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids). A “nucleoside” refers to a compound containing a sugar molecule (*e.g.*, a pentose or ribose) or a derivative thereof in combination with an organic base (*e.g.*, a purine or pyrimidine) or a derivative thereof (also
20 referred to herein as “nucleobase”). A “nucleotide” refers to a nucleoside, including a phosphate group. Modified nucleotides may be synthesized by any useful method, such as, for example, chemically, enzymatically, or recombinantly, to include one or more modified or non-natural nucleosides. Nucleic acids can comprise a region or regions of linked nucleosides. Such regions may have variable backbone linkages. The linkages can be
25 standard phosphodiester linkages, in which case the nucleic acids would comprise regions of nucleotides.

 Modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases,
30 wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two

complementary non-standard base structures, such as, for example, in those nucleic acids having at least one chemical modification. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil.

Any combination of base/sugar or linker may be incorporated into nucleic acids of the present disclosure.

In some embodiments, modified nucleobases in nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) comprise N1-methyl-pseudouridine (m1ψ), 1-ethyl-pseudouridine (e1ψ), 5-methoxy-uridine (mo5U), 5-methyl-cytidine (m5C), and/or pseudouridine (ψ). In some embodiments, modified nucleobases in nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) comprise 5-methoxymethyl uridine, 5-methylthio uridine, 1-methoxymethyl pseudouridine, 5-methyl cytidine, and/or 5-methoxy cytidine. In some embodiments, the polyribonucleotide includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of any of the aforementioned modified nucleobases, including but not limited to chemical modifications.

In some embodiments, a RNA nucleic acid of the disclosure comprises N1-methyl-pseudouridine (m1ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises N1-methyl-pseudouridine (m1ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises uridine at one or more or all uridine positions of the nucleic acid.

In some embodiments, nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) are uniformly modified (*e.g.*, fully modified, modified throughout the entire sequence)

for a particular modification. For example, a nucleic acid can be uniformly modified with N1-methyl-pseudouridine, meaning that all uridine residues in the mRNA sequence are replaced with N1-methyl-pseudouridine. Similarly, a nucleic acid can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as those set forth above.

The nucleic acids of the present disclosure may be partially or fully modified along the entire length of the molecule. For example, one or more or all or a given type of nucleotide (*e.g.*, purine or pyrimidine, or any one or more or all of A, G, U, C) may be uniformly modified in a nucleic acid of the disclosure, or in a predetermined sequence region thereof (*e.g.*, in the mRNA including or excluding the polyA tail). In some embodiments, all nucleotides X in a nucleic acid of the present disclosure (or in a sequence region thereof) are modified nucleotides, wherein X may be any one of nucleotides A, G, U, C, or any one of the combinations A+G, A+U, A+C, G+U, G+C, U+C, A+G+U, A+G+C, G+U+C or A+G+C.

The nucleic acid may contain from about 1% to about 100% modified nucleotides (either in relation to overall nucleotide content, or in relation to one or more types of nucleotide, *i.e.*, any one or more of A, G, U or C) or any intervening percentage (*e.g.*, from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100%). It will be understood that any remaining percentage is accounted for by the presence of unmodified A, G, U, or C.

The nucleic acids may contain at a minimum 1% and at maximum 100% modified nucleotides, or any intervening percentage, such as at least 5% modified nucleotides, at least 10% modified nucleotides, at least 25% modified nucleotides, at least 50% modified nucleotides, at least 80% modified nucleotides, or at least 90% modified nucleotides. For example, the nucleic acids may contain a modified pyrimidine such as a modified uracil or cytosine. In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least

80%, at least 90% or 100% of the uracil in the nucleic acid is replaced with a modified uracil (*e.g.*, a 5-substituted uracil). The modified uracil can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (*e.g.*, 2, 3, 4 or more unique structures). In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the cytosine in the nucleic acid is replaced with a modified cytosine (*e.g.*, a 5-substituted cytosine). The modified cytosine can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (*e.g.*, 2, 3, 4 or more unique structures).

10. Untranslated Regions (UTRs)

Untranslated regions (UTRs) are nucleic acid sections of a polynucleotide before a start codon (5' UTR) and after a stop codon (3' UTR) that are not translated. In some embodiments, a polynucleotide (*e.g.*, a ribonucleic acid (RNA), *e.g.*, a messenger RNA (mRNA)) of the present disclosure comprising an open reading frame (ORF) encoding an ASL polypeptide further comprises a UTR (*e.g.*, a 5' UTR or functional fragment thereof, a 3' UTR or functional fragment thereof, or a combination thereof).

A UTR (*e.g.*, 5' UTR or 3' UTR) can be homologous or heterologous to the coding region in a polynucleotide. In some embodiments, the UTR is homologous to the ORF encoding the ASL polypeptide. In some embodiments, the UTR is heterologous to the ORF encoding the ASL polypeptide.

In some embodiments, the polynucleotide comprises two or more 5' UTRs or functional fragments thereof, each of which has the same or different nucleotide sequences. In some embodiments, the polynucleotide comprises two or more 3' UTRs or functional fragments thereof, each of which has the same or different nucleotide sequences.

In some embodiments, the 5' UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof is sequence optimized.

In some embodiments, the 5'UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof comprises at least one chemically modified nucleobase, *e.g.*, N1-methylpseudouracil or 5-methoxyuracil.

UTRs can have features that provide a regulatory role, *e.g.*, increased or decreased stability, localization and/or translation efficiency. A polynucleotide comprising a UTR can be administered to a cell, tissue, or organism, and one or more regulatory features can be measured using routine methods. In some embodiments, a functional fragment of a 5' UTR or 3' UTR comprises one or more regulatory features of a full length 5' or 3' UTR, respectively.

Natural 5'UTRs bear features that play roles in translation initiation. They harbor signatures like Kozak sequences that are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus CCR(A/G)CCAUGG (SEQ ID NO:214), where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5' UTRs also have been known to form secondary structures that are involved in elongation factor binding.

By engineering the features typically found in abundantly expressed genes of specific target organs, one can enhance the stability and protein production of a polynucleotide. For example, introduction of 5' UTR of liver-expressed mRNA, such as albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, or Factor VIII, can enhance expression of polynucleotides in hepatic cell lines or liver. Likewise, use of 5'UTR from other tissue-specific mRNA to improve expression in that tissue is possible for muscle (*e.g.*, MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (*e.g.*, Tie-1, CD36), for myeloid cells (*e.g.*, C/EBP, AML1, G-CSF, GM-CSF, CD11b, MSR, Fr-1, i-NOS), for leukocytes (*e.g.*, CD45, CD18), for adipose tissue (*e.g.*, CD36, GLUT4, ACRP30, adiponectin) and for lung epithelial cells (*e.g.*, SP-A/B/C/D).

In some embodiments, UTRs are selected from a family of transcripts whose proteins share a common function, structure, feature or property. For example, an encoded polypeptide can belong to a family of proteins (*i.e.*, that share at least one function, structure, feature, localization, origin, or expression pattern), which are expressed in a particular cell, tissue or at some time during development. The UTRs from any of the genes or mRNA can be swapped for any other UTR of the same or different family of proteins to create a new polynucleotide.

In some embodiments, the 5' UTR and the 3' UTR can be heterologous. In some embodiments, the 5' UTR can be derived from a different species than the 3' UTR. In some embodiments, the 3' UTR can be derived from a different species than the 5' UTR.

Co-owned International Patent Application No. PCT/US2014/021522 (Publ. No. WO/2014/164253, incorporated herein by reference in its entirety) provides a listing of exemplary UTRs that can be utilized in the polynucleotide of the present disclosure as flanking regions to an ORF.

Additional exemplary UTRs of the application include, but are not limited to, one or more 5'UTR and/or 3'UTR derived from the nucleic acid sequence of: a globin, such as an α - or β -globin (*e.g.*, a *Xenopus*, mouse, rabbit, or human globin); a strong Kozak translational initiation signal; a CYBA (*e.g.*, human cytochrome b-245 α polypeptide); an albumin (*e.g.*, human albumin7); a HSD17B4 (hydroxysteroid (17- β) dehydrogenase); a virus (*e.g.*, a tobacco etch virus (TEV), a Venezuelan equine encephalitis virus (VEEV), a Dengue virus, a cytomegalovirus (CMV) (*e.g.*, CMV immediate early 1 (IE1)), a hepatitis virus (*e.g.*, hepatitis B virus), a sindbis virus, or a PAV barley yellow dwarf virus); a heat shock protein (*e.g.*, hsp70); a translation initiation factor (*e.g.*, eIF4G); a glucose transporter (*e.g.*, hGLUT1 (human glucose transporter 1)); an actin (*e.g.*, human α or β actin); a GAPDH; a tubulin; a histone; a citric acid cycle enzyme; a topoisomerase (*e.g.*, a 5'UTR of a TOP gene lacking the 5' TOP motif (the oligopyrimidine tract)); a ribosomal protein Large 32 (L32); a ribosomal protein (*e.g.*, human or mouse ribosomal protein, such as, for example, rps9); an ATP synthase (*e.g.*, ATP5A1 or the β subunit of mitochondrial H^+ -ATP synthase); a growth hormone (*e.g.*, bovine (bGH) or human (hGH)); an elongation factor (*e.g.*, elongation factor 1 α 1 (EEF1A1)); a manganese superoxide dismutase (MnSOD); a myocyte enhancer factor 2A (MEF2A); a β -F1-ATPase, a creatine kinase, a myoglobin, a granulocyte-colony stimulating factor (G-CSF); a collagen (*e.g.*, collagen type I, alpha 2 (Col1A2), collagen type I, alpha 1 (Col1A1), collagen type VI, alpha 2 (Col6A2), collagen type VI, alpha 1 (Col6A1)); a ribophorin (*e.g.*, ribophorin I (RPNI)); a low density lipoprotein receptor-related protein (*e.g.*, LRP1); a cardiotrophin-like cytokine factor (*e.g.*, Nnt1); calreticulin (Calr); a procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 (Plod1); and a nucleobindin (*e.g.*, Nucb1).

In some embodiments, the 5' UTR is selected from the group consisting of a β -globin 5' UTR; a 5'UTR containing a strong Kozak translational initiation signal; a cytochrome b-245 α polypeptide (CYBA) 5' UTR; a hydroxysteroid (17- β) dehydrogenase (HSD17B4) 5'

UTR; a Tobacco etch virus (TEV) 5' UTR; a Venezuelan equine encephalitis virus (VEEV) 5' UTR; a 5' proximal open reading frame of rubella virus (RV) RNA encoding nonstructural proteins; a Dengue virus (DEN) 5' UTR; a heat shock protein 70 (Hsp70) 5' UTR; a eIF4G 5' UTR; a GLUT1 5' UTR; functional fragments thereof and any combination thereof.

5 In some embodiments, the 3' UTR is selected from the group consisting of a β -globin 3' UTR; a CYBA 3' UTR; an albumin 3' UTR; a growth hormone (GH) 3' UTR; a VEEV 3' UTR; a hepatitis B virus (HBV) 3' UTR; α -globin 3'UTR; a DEN 3' UTR; a PAV barley yellow dwarf virus (BYDV-PAV) 3' UTR; an elongation factor 1 α 1 (EEF1A1) 3' UTR; a manganese superoxide dismutase (MnSOD) 3' UTR; a β subunit of mitochondrial H(+)-ATP synthase (β -mRNA) 3' UTR; a GLUT1 3' UTR; a MEF2A 3' UTR; a β -F1-ATPase 3' UTR; 10 functional fragments thereof, and combinations thereof.

Wild-type UTRs derived from any gene or mRNA can be incorporated into the polynucleotides of the present disclosure. In some embodiments, a UTR can be altered relative to a wild type or native UTR to produce a variant UTR, *e.g.*, by changing the 15 orientation or location of the UTR relative to the ORF; or by inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides. In some embodiments, variants of 5' or 3' UTRs can be utilized, for example, mutants of wild type UTRs, or variants wherein one or more nucleotides are added to or removed from a terminus of the UTR.

20 Additionally, one or more synthetic UTRs can be used in combination with one or more non-synthetic UTRs. *See, e.g.*, Mandal and Rossi, *Nat. Protoc.* 2013 8(3):568-82, the contents of which are incorporated herein by reference in their entirety.

UTRs or portions thereof can be placed in the same orientation as in the transcript from which they were selected or can be altered in orientation or location. Hence, a 5' and/or 25 3' UTR can be inverted, shortened, lengthened, or combined with one or more other 5' UTRs or 3' UTRs.

In some embodiments, the polynucleotide comprises multiple UTRs, *e.g.*, a double, a triple or a quadruple 5' UTR or 3' UTR. For example, a double UTR comprises two copies of the same UTR either in series or substantially in series. For example, a double beta-globin 30 3'UTR can be used (see US2010/0129877, the contents of which are incorporated herein by reference in its entirety).

The polynucleotides of the present disclosure can comprise combinations of features. For example, the ORF can be flanked by a 5'UTR that comprises a strong Kozak translational initiation signal and/or a 3'UTR comprising an oligo(dT) sequence for templated addition of a poly-A tail. A 5'UTR can comprise a first polynucleotide fragment and a second
5 polynucleotide fragment from the same and/or different UTRs (see, *e.g.*, US2010/0293625, herein incorporated by reference in its entirety).

Other non-UTR sequences can be used as regions or subregions within the polynucleotides of the present disclosure. For example, introns or portions of intron sequences can be incorporated into the polynucleotides of the present disclosure.

10 Incorporation of intronic sequences can increase protein production as well as polynucleotide expression levels. In some embodiments, the polynucleotide of the present disclosure comprises an internal ribosome entry site (IRES) instead of or in addition to a UTR (see, *e.g.*, Yakubov et al., *Biochem. Biophys. Res. Commun.* 2010 394(1):189-193, the contents of which are incorporated herein by reference in their entirety). In some embodiments, the
15 polynucleotide comprises an IRES instead of a 5' UTR sequence. In some embodiments, the polynucleotide comprises an ORF and a viral capsid sequence. In some embodiments, the polynucleotide comprises a synthetic 5' UTR in combination with a non-synthetic 3' UTR.

In some embodiments, the UTR can also include at least one translation enhancer polynucleotide, translation enhancer element, or translational enhancer elements
20 (collectively, "TEE," which refers to nucleic acid sequences that increase the amount of polypeptide or protein produced from a polynucleotide. As a non-limiting example, the TEE can be located between the transcription promoter and the start codon. In some embodiments, the 5' UTR comprises a TEE.

In one aspect, a TEE is a conserved element in a UTR that can promote translational
25 activity of a nucleic acid such as, but not limited to, cap-dependent or cap-independent translation.

a. 5' UTR sequences

5' UTR sequences are important for ribosome recruitment to the mRNA and have been reported to play a role in translation (Hinnebusch A, et al., (2016) *Science*, 352:6292:
30 1413-6).

Disclosed herein, *inter alia*, is a polynucleotide, *e.g.*, mRNA, comprising an open reading frame (*e.g.*, any of of SEQ ID NO:20–33) encoding an ASL polypeptide (*e.g.*, any of of SEQ ID NO:1–11), which polynucleotide has a 5' UTR that confers an increased half-life, increased expression and/or increased activity of the polypeptide encoded by said

5 polynucleotide, or of the polynucleotide itself. In some embodiments, a polynucleotide disclosed herein comprises: (a) a 5'-UTR (*e.g.*, as provided in **Table 3** or a variant or fragment thereof); (b) a coding region comprising a stop element (*e.g.*, as described herein); and (c) a 3'-UTR (*e.g.*, as described herein), and LNP compositions comprising the same. In some embodiments, the polynucleotide comprises a 5'-UTR comprising a sequence provided
10 in **Table 3** or a variant or fragment thereof (*e.g.*, a functional variant or fragment thereof).

In some embodiments, the polynucleotide having a 5' UTR sequence provided in **Table 3** or a variant or fragment thereof, has an increase in the half-life of the polynucleotide, *e.g.*, about 1.5-20-fold increase in half-life of the polynucleotide. In some embodiments, the increase in half-life is about 1.5-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-,
15 17-, 18-, 19- or 20-fold, or more. In some embodiments, the increase in half life is about 1.5-fold or more. In some embodiments, the increase in half life is about 2-fold or more. In some embodiments, the increase in half life is about 3-fold or more. In some embodiments, the increase in half life is about 4-fold or more. In some embodiments, the increase in half life is about 5-fold or more.

20 In some embodiments, the polynucleotide having a 5' UTR sequence provided in **Table 3** or a variant or fragment thereof, results in an increased level and/or activity, *e.g.*, output, of the polypeptide encoded by the polynucleotide. In some embodiments, the 5'UTR results in about 1.5-20-fold increase in level and/or activity, *e.g.*, output, of the polypeptide encoded by the polynucleotide. In some embodiments, the increase in level and/or activity is
25 about 1.5-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-, 17-, 18-, 19- or 20-fold, or more. In some embodiments, the increase in level and/or activity is about 1.5-fold or more. In some embodiments, the increase in level and/or activity is about 2-fold or more. In some embodiments, the increase in level and/or activity is about 3-fold or more. In some
30 embodiments, the increase in level and/or activity is about 4-fold or more. In some embodiments, the increase in level and/or activity is about 5-fold or more.

In some embodiments, the increase is compared to an otherwise similar polynucleotide which does not have a 5' UTR, has a different 5' UTR, or does not have a 5' UTR described in **Table 3** or a variant or fragment thereof.

5 In some embodiments, the increase in half-life of the polynucleotide is measured according to an assay that measures the half-life of a polynucleotide.

In some embodiments, the increase in level and/or activity, *e.g.*, output, of the polypeptide encoded by the polynucleotide is measured according to an assay that measures the level and/or activity of a polypeptide.

10 In some embodiments, the 5' UTR comprises a sequence provided in **Table 3** or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to a 5' UTR sequence provided in **Table 3**, or a variant or a fragment thereof. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, or SEQ ID
15 NO:78.

In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:50. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:51. In some embodiments, the 5' UTR comprises a sequence
20 with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:52. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:53. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:54. In some embodiments, the 5' UTR comprises a sequence
25 with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:55. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:56. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:57. In some embodiments, the 5' UTR comprises a sequence
30 with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID

NO:58. In some embodiments, the 5' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:78.

In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO:50. In some embodiments, the 5' UTR consists of the sequence of SEQ ID NO:50. In some
 5 embodiments, the 5' UTR comprises the sequence of SEQ ID NO:55. In some embodiments, the 5' UTR consists of the sequence of SEQ ID NO:55. In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO:56. In some embodiments, the 5' UTR consists of the sequence of SEQ ID NO:56. In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO:78. In some embodiments, the 5' UTR consists of the sequence of SEQ ID
 10 NO:78.

In some embodiments, a 5' UTR sequence provided in **Table 3** has a first nucleotide (not shown) which is an A. In some embodiments, a 5' UTR sequence provided in **Table 3** has a first nucleotide (not shown) which is a G.

Table 3: 5' UTR sequences

SEQ ID NO:	Sequence name	Sequence
50	A1	GGAAAUCGCAAAAUUUGCUCUUCGCGUUAGAUUUCUUUUAG UUUUCUCGCAACUAGCAAGCUUUUUGUUCUCGCC
51	A5	GGAAAUCCCCACAACCGCCUCAUAUCCAGGCUCAAGAAUAG AGCUCAGUGUUUUGUUGUUUAAUCAUCCGACGUGUUUUGC GAUAUUCGCGCAAAGCAGCCAGUCGCGCGCUUGC UUUUAAG UAGAGUUGUUUUUCCACCCGUUUGCCAGGCAUCUUUAAUUU AACAUAUUUUUAUUUUUCAGGCUAACCUACGCCGCCACC
52	A6	GGAAUAAGAGAGAAAAGAAGAGUAAGAAGAAUAUAAGAU CUCCCUGAGCUUCAGGGAGCCCCGGCGCCGCCACC
53	A7	GGAAACCCCCACCCCCGUAAGAGAGAAAAGAAGAGUAAGA AGAAUAUAAGAUCUCCCUGAGCUUCAGGGAGCCCCGGCGC CGCCACC
54	A8	GGAGAACUCCGCUUCCGUUGGCGCAAGCGCUUUCAUUUUU UCUGCUACCGUGACUAAG

55	A9	GGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAG CCACC
56	A11 (Reference)	GGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAC CCCGGCGCCGCCACC
57	A2	GGAAAUUCGCAAAAUUUGCUCUUCGCGUUAGAUUUCUUUUAG UUUUCUCGCAACUAGCAAGCUUUUUGUUCUCGCCGCCGCC
58	A3	GGAAAUUCGCAAAAUUUCUUUUCGCGUUAGAUUUCUUUUAG UUUUCUUUCAACUAGCAAGCUUUUUGUUCUCGCCGCCGCC
59	A4	G G A A A U C G C A A A A (N ₂) _x (N ₃) _x C U (N ₄) _x (N ₅) _x C G C G U U A G A U U U C U U U U A G U U U U C U N ₆ N ₇ C A A C U A G C A A G C U U U U U G U U C U C G C C (N ₈ C C) _x (N ₂) _x is a uracil and x is an integer from 0 to 5, <i>e.g.</i> , wherein x =3 or 4; (N ₃) _x is a guanine and x is an integer from 0 to 1; (N ₄) _x is a cytosine and x is an integer from 0 to 1; (N ₅) _x is a uracil and x is an integer from 0 to 5, <i>e.g.</i> , wherein x =2 or 3; N ₆ is a uracil or cytosine; N ₇ is a uracil or guanine; N ₈ is adenine or guanine and x is an integer from 0 to 1.
60	A27	GGAAAAUUUUAGCCUGGAACGUUAGAUAAACUGUCCUGUUGU CUUUAAUUAACUUGGUCGCCCAAGUAGUUUGUCUCCAAA
61	A12	GGAAACUUUAUUUAGUGUUACUUUAUUUUCUGUUUAUUUGU GUUUCUUCAGUGGGUUUGUUCUAAUUUCCUUGGCCGCC
62	A13	GGAAAAUCUGUAUUAGGUUGGCGUGUUCUUUGGUCGGUUGU UAGUAUUGUUGUUGAUUCGUUUGUGGUCGGUUGGCCGCC
63	A14	GGAAAAUUAUUAACAUCUUGGUUAUUCUGAUAACCAUUCGU UGGAUUUUUAUUGUAUUCGUAGUUUGGGUUCUGGCCGCC
64	A15	GGAAAUUAUUUAUUUUCUAGCUACAAUUUAUCAUUGUAUU AUUUUAGCUAUUCAUCAUUAUUUACUUGGUGAUCAACA
65	A16	GGAAAUAGGUUGUUAACCAAGUUCAAGCCUAAUAAGCUUGG AUUCUGGUGACUUGCUUCACCGUUGGCGGGCACCGAUC

66	A17	GGAAAUCGUAGAGAGUCGUACUUAGUACAUUUCGACUAUCG GUGGACACCAUCAAGAUUAUAAACCAGGCCAGA
67	A18	GGAAACCCGCCCAAGCGACCCCAACAUUUCAGCAGUUGCCC AAUCCCAACUCCCAACACAAUCCCAAGCAACGCCGCC
68	A19	GGAAAGCGAUUGAAGGCGUCUUUUAACUACUCGAUUAAGG UUGGGUAUCGUCGUGGGACUUGGAAAUUUGUUGUUUCC
69	A20	GGAAACUAAUCGAAAUAAAAGAGCCCCGUACUCUUUUAUUU CUAUUAGGUUAGGAGCCUUAGCAUUUGUAUCUUAGGUA
70	A21	GGAAAUGUGAUUUUCCAGCAACUUCUUUUGAAUUAUUAUUGAAU UCCUAAUUCAAAGCGAACAAAUUCUACAAGCCAUUAUACC
71	A22	GGAAAUCGUAGAGAGUCGUACUUACGUGGUCGCCAUUGCAU AGCGCGCGAAAGCAACAGGAACAAGAACGCGCC
72	A23	GGAAAUCGUAGAGAGUCGUACUUAGAAUAAACAGAGUCGGG UCGACUUGUCUCUGAUACUACGACGUCACAAUC
73	A24	GGAAAUUUUGCCUUCGGAGUUGCGUAUCCUGAACUGCCCAG CCUCCUGAUUAUACAACUGUUCGCGCUUAUUCGGGCCGCC
74	A25	GGAAAUCUGAGCAGGAAUCCUUUGUGCAUUGAAGACUUUAG AUUCCUCUCUGCGGUAGACGUGCACUUUAUAAGUAUUUG
75	A26	GGAAAGCGAUUGAAGGCGUCUUUUAACUACUCGAUUAAGG UUGGGUAUCGUCGUGGGACUUGGAAAUUUGUUGCCACC
76	A28	GGAAAUUUUUUUUGAUUAUUAAGAGUUUUUUUUUGAUUAU UAAGAAAAUUUUUUUUGAUUAUUAAGAGUAAGAAGAAAU AUAAGACCCCGGCGCCGCCACC
77	A29	GGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAG CCAAAAAATAAAACC
78	A30	GGAAAUCUCCUGAGCUUCAGGGAGUAAGAGAGAAAAGAAG AGUAAGAAGAAAUUAAGACCCCGGCGCCGCCACC
79	A31	GCCRCC, wherein R= A or G

In some embodiments, the 5' UTR comprises a variant of SEQ ID NO:50. In some embodiments, the variant of SEQ ID NO:50 comprises a nucleic acid sequence of Formula A:

5 G G A A A U C G C A A A A (N₂)_x (N₃)_x C U (N₄)_x (N₅)_x C G C G U U A G A U U U C U
U U U A G U U U U C U N₆ N₇ C A A C U A G C A A G C U U U U U G U U C U C G C
C (N₈ C C)_x (SEQ ID NO:59),

wherein:

$(N_2)_x$ is a uracil and x is an integer from 0 to 5, *e.g.*, wherein $x = 3$ or 4;

$(N_3)_x$ is a guanine and x is an integer from 0 to 1;

$(N_4)_x$ is a cytosine and x is an integer from 0 to 1;

5 $(N_5)_x$ is a uracil and x is an integer from 0 to 5, *e.g.*, wherein $x = 2$ or 3;

N_6 is a uracil or cytosine;

N_7 is a uracil or guanine;

N_8 is adenine or guanine and x is an integer from 0 to 1.

10 In some embodiments $(N_2)_x$ is a uracil and x is 0. In some embodiments $(N_2)_x$ is a uracil and x is 1. In some embodiments $(N_2)_x$ is a uracil and x is 2. In some embodiments $(N_2)_x$ is a uracil and x is 3. In some embodiments, $(N_2)_x$ is a uracil and x is 4. In some embodiments $(N_2)_x$ is a uracil and x is 5.

In some embodiments, $(N_3)_x$ is a guanine and x is 0. In some embodiments, $(N_3)_x$ is a
15 guanine and x is 1.

In some embodiments, $(N_4)_x$ is a cytosine and x is 0. In some embodiments, $(N_4)_x$ is a cytosine and x is 1.

In some embodiments, $(N_5)_x$ is a uracil and x is 0. In some embodiments, $(N_5)_x$ is a uracil and x is 1. In some embodiments, $(N_5)_x$ is a uracil and x is 2. In some embodiments,
20 $(N_5)_x$ is a uracil and x is 3. In some embodiments, $(N_5)_x$ is a uracil and x is 4. In some embodiments $(N_5)_x$ is a uracil and x is 5.

In some embodiments, N_6 is a uracil. In some embodiments, N_6 is a cytosine.

In some embodiments, N_7 is a uracil. In some embodiments, N_7 is a guanine.

In some embodiments, N_8 is an adenine and x is 0. In some embodiments, N_8 is an
25 adenine and x is 1.

In some embodiments, N_8 is a guanine and x is 0. In some embodiments, N_8 is a guanine and x is 1.

In some embodiments, the 5' UTR comprises a variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID

NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 50% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 60% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 70% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 80% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 90% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 95% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 96% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 97% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 98% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a sequence with at least 99% identity to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78.

In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 5%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ

ID NO:78 comprises a uridine content of at least 10%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 20%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 30%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 40%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 50%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 60%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 70%. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises a uridine content of at least 80%.

In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises at least 2, 3, 4, 5, 6 or 7 consecutive uridines (*e.g.*, a polyuridine tract). In some embodiments, the polyuridine tract in the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises at least 1-7, 2-7, 3-7, 4-7, 5-7, 6-7, 1-6, 1-5, 1-4, 1-3, 1-2, 2-6, or 3-5 consecutive uridines. In some embodiments, the polyuridine tract in the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 4 consecutive uridines. In some embodiments, the polyuridine tract in the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 5 consecutive uridines.

In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 polyuridine tracts. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 3 polyuridine tracts. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 4 polyuridine tracts. In some embodiments, the variant of SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78 comprises 5 polyuridine tracts.

In some embodiments, one or more of the polyuridine tracts are adjacent to a different polyuridine tract. In some embodiments, each of, *e.g.*, all, the polyuridine tracts are adjacent to each other, *e.g.*, all of the polyuridine tracts are contiguous.

In some embodiments, one or more of the polyuridine tracts are separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 or 60 nucleotides. In some embodiments, each of, *e.g.*, all of, the polyuridine tracts are separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 or 60 nucleotides.

5 In some embodiments, a first polyuridine tract and a second polyuridine tract are adjacent to each other.

In some embodiments, a subsequent, *e.g.*, third, fourth, fifth, sixth or seventh, eighth, ninth, or tenth, polyuridine tract is separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 or 60 nucleotides from the first polyuridine tract, the second
10 polyuridine tract, or any one of the subsequent polyuridine tracts.

In some embodiments, a first polyuridine tract is separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 or 60 nucleotides from a subsequent polyuridine tract, *e.g.*, a second, third, fourth, fifth, sixth or seventh, eighth, ninth, or tenth polyuridine tract. In some embodiments, one or more of the subsequent polyuridine tracts are
15 adjacent to a different polyuridine tract.

In some embodiments, the 5' UTR comprises a Kozak sequence, *e.g.*, a GCCRCC nucleotide sequence (SEQ ID NO:79) wherein R is an adenine or guanine. In some embodiments, the Kozak sequence is disposed at the 3' end of the 5'UTR sequence.

In an aspect, the polynucleotide (*e.g.*, mRNA) comprising an open reading frame
20 (*e.g.*, SEQ ID NO:20) encoding an ASL polypeptide (*e.g.*, SEQ ID NO:2) and comprising a 5' UTR sequence disclosed herein is formulated as an LNP. In some embodiments, the LNP composition comprises: (i) an ionizable lipid, *e.g.*, an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.

In another aspect, the LNP compositions of the disclosure are used in a method of
25 treating argininosuccinic aciduria in a subject.

In another aspect, an LNP composition comprising a polynucleotide disclosed herein encoding an ASL polypeptide, *e.g.*, as described herein, can be administered with an additional agent, *e.g.*, as described herein.

b. 3' UTR sequences

3'UTR sequences have been shown to influence translation, half-life, and subcellular localization of mRNAs (Mayr C., *Cold Spring Harb. Persp. Biol.* 2019 Oct 1;11(10):a034728).

Disclosed herein, *inter alia*, is a polynucleotide, *e.g.*, mRNA, comprising an open
5 reading frame (*e.g.*, any one of SEQ ID NO:20–33) encoding an ASL polypeptide (*e.g.*, any one of SEQ ID NO:1–11), which polynucleotide has a 3' UTR that confers an increased half-life, increased expression and/or increased activity of the polypeptide encoded by said polynucleotide, or of the polynucleotide itself., a polynucleotide disclosed herein comprises:
(a) a 5'-UTR (*e.g.*, as described herein); (b) a coding region comprising a stop element (*e.g.*,
10 as described herein); and (c) a 3'-UTR (*e.g.*, as provided in **Table 4** or a variant or fragment thereof), and LNP compositions comprising the same. In some embodiments, the polynucleotide comprises a 3'-UTR comprising a sequence provided in **Table 4** or a variant or fragment thereof.

In some embodiments, the polynucleotide having a 3' UTR sequence provided in
15 **Table 4** or a variant or fragment thereof, results in an increased half-life of the polynucleotide, *e.g.*, about 1.5-10-fold increase in half-life of the polynucleotide. In some embodiments, the increase in half-life is about 1.5-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, or 10-fold, or more. In some embodiments, the increase in half-life is about 1.5-fold or more. In some
20 embodiments, the increase in half-life is about 2-fold or more. In some embodiments, the increase in half-life is about 3-fold or more. In some embodiments, the increase in half-life is about 4-fold or more. In some embodiments, the increase in half-life is about 5-fold or more. In some
25 embodiments, the increase in half-life is about 6-fold or more. In some embodiments, the increase in half-life is about 7-fold or more. In some embodiments, the increase in half-life is about 8-fold. In some embodiments, the increase in half-life is about 9-fold or more. In some embodiments, the increase in half-life is about 10-fold or more.

In some embodiments, the polynucleotide having a 3' UTR sequence provided in **Table 4** or a variant or fragment thereof, results in a polynucleotide with a mean half-life score of greater than 10.

In some embodiments, the polynucleotide having a 3' UTR sequence provided in
30 **Table 4** or a variant or fragment thereof, results in an increased level and/or activity, *e.g.*, output, of the polypeptide encoded by the polynucleotide.

In some embodiments, the increase is compared to an otherwise similar polynucleotide which does not have a 3' UTR, has a different 3' UTR, or does not have a 3' UTR of **Table 4** or a variant or fragment thereof.

In some embodiments, the polynucleotide comprises a 3' UTR sequence provided in **Table 4** or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to a 3' UTR sequence provided in **Table 4**, or a fragment thereof. In some embodiments, the 3' UTR comprises a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:128, SEQ ID NO:137, SEQ ID NO:138, and SEQ ID NO:139.

In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:100, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:100. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:101, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:101. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:102, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:102. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:103, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:103. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:104, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:104. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:105, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:105. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:106, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:106. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:107, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:107. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:108, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:108. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:109, or a

sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:109. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:110, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:110. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:111, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:111. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:112, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:112. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:113, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:113. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:114, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:114. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:115, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:115. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:128, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:128. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:137, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:137. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:138, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:138. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:139, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:139. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:140, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:140. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:141, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:141. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:142, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:142. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:143, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:143. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:144, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:144. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:145, or a sequence with at least 80%, 85%,

90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:145. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:146, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:146. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:147, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:147. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO:148, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO:148.

Table 4: 3' UTR sequences

SEQ ID NO	Sequence information	Sequence
100	B1	UAAUAGUAAGCUGGAGCCUCCUGAGAGACCUGUGUGAACUAUU GAGAAGAUCGGAACAGCUCCUACUCUGAGGAAGUUGGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
101	B3	UAAUAGUAAGCUGGAGCCUCACUCUCCUCUCCAUCCCGUAUCC AGGCUGUGAAUUUUUCAAGGAAUAUAAAGAUCCGGAUGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
102	B4	UGAUAGUAAGCUGGAGCCUCUAGUGACGGCAACAGGGCUUGGU UUUUCUUGUUGUGAAAUCGACAUCUCUGAAGACAGGGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
103	B5	UGAUAGUAAGCUGGAGCCUCCUCCAUCUAGUCACAAAGACUC CUUCGUCCCCAGUUGCCGUCUAGGAUUGGGCCUCCAGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
104	B6	UGAUAGUAAGCUGGAGCCUCCCAUAACAUGACAUAUCUGGAUU UUGUGCUUAGAACCUUAAAUUGGAAGCAUUCUAAUUGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
105	B7	UAAUAGUAAGCUGGAGCCUCCGAAAACUAAAUAAGAGUAUU UCAAGAUUUUAUAAUUUUCAAAGACCUUUGAAAUUUGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
106	B8	UAAUAGUAAGCUGGAGCCUCUACACAUUGCUCUAGUUGGCAG AAAUAAUUGAUUAAAAGACCAGAAACUGUGAUAAACUGGUACCC CCGUGGUCUUUAAUAAAGUCUAAGUGGGCGGC

107	B9	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
108	B10	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
109	B11	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGC
110	B12	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGC
111	B13	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCUCCAUAAAGUAGGAAACACUACAGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC
112	B14	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUU GGGCCUCCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUACCC CCCGCAUUAUUACUCACGGUACGAGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGC
113	B15	UGAUAAUAGUCCAUAAAGUAGGAAACACUACAGCUGGAGCCUC GGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUC CUCCCCUUCUGCACCCGUACCCCCCGCAUUAUUACUCACGGU ACGAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
114	B16	UGAUAAUAGUCCAUAAAGUAGGAAACACUACAGCUGGAGCCUC GGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAAGUAGGAA ACACUACAUCCCCCCAGCCCCUCCUCCCCUUCUGCACCCGUA CCCCCUCCAUAAAGUAGGAAACACUACAGUGGUCUUUGAAUAA AGUCUGAGUGGGCGGC
115	B2	CUGAGAGACCUGUGUGAACUAUUGAGAAGAUCGGAACAGCUCC UUACUCUGAGGAAGUUG

116	B17	UGAUAAUAGGCUGGAGCCUCACACACCUCUGCCCCUUGGGC CUCCCACUCCCAUGGCUCUGGGCGGUCCAGAAGGAGCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
117	B18	UGAUAAUAGGCUGGAGCCUCCACCGCGUUAUCCGUUCCUCGUA GGCUGGUCUUGGGGAACGGGUCGGCGGGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC
118	B19	UGAUAAUAGGCUGGAGCCUCUGCCCCGGCAACGGCCAGGUCUGU GCCAAGUGUUUGCUGACGCAACCCCCACUGGCUGGGGCUUGGU CAUGGGGCCAUCAGCGCGUGCGUGGAACCUUUUCGGCUCUCUG CCGAUCCAUAUCUGCGGAACUCCUAGCCGCUUGUUUUGCUCGCA GCAGGUCUGGAGCAAACAUAUCGGGACUGAUAACUCUGUUGU CCUGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
119	B20	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCU GGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCC UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU CGGC
120	B21	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCU GGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCC UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU CGGC
133	B22	UAAGUCUCCAUAAGUAGGAAACACUACAGCUGGAGCCUCGGU GGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAGUAGGAAACA CUACAUCCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCC CCUCCAUAAGUAGGAAACACUACAGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC
134	B23	UAAAGCGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGG CCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCU CCAUAAGUAGGAAACACUACAGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGC

135	B24	UAAAGCUCCCCGGGGUCCAUAAGUAGGAAACACUACAGCUGG AGCCUCCUGAGAGACCUGUGUGAACUAUUGAGAAGAUCGGAAC AGCUCCUACUCUGAGGAAGUUGUCCAUAAGUAGGAAACACU ACAGUACCCCUCCAUAAGUAGGAAACACUACAGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGC
136	B25	UAAUAGUAAACCUCACUCACGGCCACAUUGAGUGCCAGGCUCC GGGCUGGUUUUAUAGUAGUGUAGAGCAUUGCAGCACUUAGACUG GGGUGCUGUAGUCUUUAUUGUAGUCUUUCCACAUACCUGAUAA UUCUUAGAUAAUUCUUAUUUUAAUUCCAUAAAGUAGGAAACA CUACAUAUAUCUCCAUAAGUAGGAAACACUACAUAUUCUCC AUAAAGUAGGAAACACUACAUAGGCU
137	B26	UAAUAGUAAACCUCACUCACGGCCACAUUGAGUGCCAGGCUCC GGGCUGGUUUUAUAGUAGUGUAGAGCAUUGCAGCACUUAGACUG GGGUGCUGUAGUCUUUAUUGUAGUCUUUCCACAUACCUGAUAA UUCUUAGAUAAUUCUUAUUUUAAUUAUAUCUAUUCUUAGGCU
138	B27	UAAAGCUAAGCUGGAGCCUCCUGAGAGACCUGUGUGAACUAUU GAGAAGAUCGGAACAGCUCCUACUCUGAGGAAGUUGGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
139	B28	UAAAGCUAAACCUCACUCACGGCCACAUUGAGUGCCAGGCUCC GGGCUGGUUUUAUAGUAGUGUAGAGCAUUGCAGCACUUAGACUG GGGUGCUGUAGUCUUUAUUGUAGUCUUUCCACAUACCUGAUAA UUCUUAGAUAAUUCUUAUUUUAAUUAUAUCUAUUCUUAGGCU
140	D1	UAAAGCUCCCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCCCUU GGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
141	D2	UAAGUCUAAGCUGGAGCCUCCUGAGAGACCUGUGUGAACUAUU GAGAAGAUCGGAACAGCUCCUACUCUGAGGAAGUUGGUACCC CCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
142	D3	UAAAGCUCCCCGGGGCAAACACCAUUGUCACACUCCAGCCUCG GUGGCCUAGCUUCUUGCCCCUUGGGCCCAAACACCAUUGUCAC ACUCCAUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACC CCCCAAACACCAUUGUCACACUCCAGUGGUCUUUGAAUAAAGU CUGAGUGGGCGGC

		(miR122 binding sites boldened)
143	D4	UAAAGCUC~~CCCCGGGG~~UCCAUA~~AAGUAGGAAACACUACAGCUGG~~ AGCCUCGGUGGCCUAGCUUCUUG~~CCCCUUGGGCCCAAACACCA~~ UUGUCACACUCCA~~UCCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCAC CCGUAC~~CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC~~ (miR-142-3p and miR122 binding sites boldened)
144	D5	UAAAGCUC~~CCCCGGGG~~GCCUCGGUGGCCUAGCUUCUUG~~CCCCU~~ GGGCCU~~CCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCACCCGUACCC CCCAAACACCAUUGUCACACUCCA~~GUGGUCUUUGAAUAAAGUC~~ UGAGUGGGCGGC (miR122 binding site boldened)
145	D6	UAAG~~CCCCUCCGGGG~~CAAACACCAUUGUCACACUCCA~~GCCUCG~~ GUGGCCUAGCUUCUUG~~CCCCUUGGGCCCAAACACCAUUGUCAC~~ ACUCCA~~UCCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCACCCGUACC CCCCAAACACCAUUGUCACACUCCA~~GUGGUCUUUGAAUAAAGU~~ CUGAGUGGGCGGC (miR122 binding sites boldened)
146	D7	UAAG~~CCCCUCCGGGG~~UCCAUA~~AAGUAGGAAACACUACAGCCUC~~ GGUGGCCUAGCUUCUUG~~CCCCUUGGGCCUCCAUA~~AAGUAGGAA ACACUACA~~UCCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCACCCGUA CCCCCGCAUUAUUAUCACGGUACGAGUGGUCUUUGAAUAA GUCUGAGUGGGCGGC (miR-142-3p and miR-126-3p binding sites boldened)
147	D8	UAAG~~CCCCUCCGGGG~~UCCAUA~~AAGUAGGAAACACUACAGCUGG~~ AGCCUCGGUGGCCUAGCUUCUUG~~CCCCUUGGGCCCAAACACCA~~ UUGUCACACUCCA~~UCCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCAC CCGUAC~~CCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC~~ (miR-142-3p and miR122 binding sites boldened)
148	D9	UAAG~~CCCCUCCGGGG~~GCCUCGGUGGCCUAGCUUCUUG~~CCCCU~~ GGGCCU~~CCCCC~~CAG~~CCCCUCCUCCCCU~~UCCUGCACCCGUACCC CCCAAACACCAUUGUCACACUCCA~~GUGGUCUUUGAAUAAAGUC~~ UGAGUGGGCGGC (miR122 binding site boldened)

In some embodiments, the 3' UTR comprises a micro RNA (miRNA) binding site, *e.g.*, as described herein, which binds to a miR present in a human cell. In some embodiments, the 3' UTR comprises a miRNA binding site of SEQ ID NO: 212, SEQ ID NO:174, SEQ ID NO:152 or a combination thereof. In some embodiments, the 3' UTR comprises a plurality of miRNA binding sites (*e.g.*, 2, 3, 4, 5, 6, 7 or 8 miRNA binding sites). In some embodiments, the plurality of miRNA binding sites comprises the same or different miRNA binding sites.

miR122 bs = CAAACACCAUUGUCACACUCCA (SEQ ID NO:212)

miR-142-3p bs = UCCAUAAGUAGGAAACACUACA (SEQ ID NO:174)

miR-126 bs = CGCAUUAUACUCACGGUACGA (SEQ ID NO:152)

In an aspect, disclosed herein is a polynucleotide encoding a polypeptide, wherein the polynucleotide comprises: (a) a 5'-UTR, *e.g.*, as described herein; (b) a coding region comprising a stop element (*e.g.*, as described herein); and (c) a 3'-UTR (*e.g.*, as described herein).

In an aspect, an LNP composition comprising a polynucleotide comprising an open reading frame (*e.g.*, any one of SEQ ID NO:20–33 or 41–44) encoding an ASL polypeptide (*e.g.*, any one of SEQ ID NO:1–11) and comprising a 3' UTR disclosed herein comprises: (i) an ionizable lipid, *e.g.*, an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.

In some embodiments, the present disclosure relates to an polynucleotide sequence (*e.g.*, an mRNA construct) comprising an ORF encoding a polypeptide, wherein the polynucleotide sequence further comprises a UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the nucleotide sequence according to SEQ ID NO:137, SEQ ID NO:138, or SEQ ID NO:139. In some embodiments, the present disclosure relates to an polynucleotide sequence (*e.g.*, an mRNA construct) comprising an ORF encoding a polypeptide, wherein the polynucleotide sequence further comprises a UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the nucleotide sequence according to SEQ ID NO:137. In some embodiments, the present disclosure relates to an polynucleotide sequence (*e.g.*, an mRNA construct) comprising an ORF encoding a polypeptide, wherein the polynucleotide sequence further comprises a UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the

nucleotide sequence according to SEQ ID NO:138. In some embodiments, the present disclosure relates to an polynucleotide sequence (*e.g.*, an mRNA construct) comprising an ORF encoding a polypeptide, wherein the polynucleotide sequence further comprises a UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the

5 nucleotide sequence according to SEQ ID NO:139.

In some embodiments, the present disclosure relates to a polynucleotide sequence comprising, from 5' end to 3' end: (i) a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); (ii) a 5'UTR; (iii) an ORF; (iv) a stop codon (if not present in the ORF); (v) a 3'UTR having a nucleotide sequence consisting of, consisting essentially of, or

10 comprising the nucleic acid sequence according to SEQ ID NO:137, SEQ ID NO:138, or SEQ ID NO:139; and (vi) a polyA tail (*e.g.*, SEQ ID NO:195 or SEQ ID NO:211). In some embodiments, the present disclosure relates to a polynucleotide sequence comprising, from 5' end to 3' end: (i) a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); (ii) a 5'UTR; (iii) an ORF; (iv) a stop codon (if not present in the ORF); (v) a 3'UTR having

15 a nucleotide sequence consisting of, consisting essentially of, or comprising the nucleic acid sequence according to SEQ ID NO:137; and (vi) a polyA tail (*e.g.*, SEQ ID NO:195 or SEQ ID NO:211). In some embodiments, the present disclosure relates to a polynucleotide sequence comprising, from 5' end to 3' end: (i) a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); (ii) a 5'UTR; (iii) an ORF; (iv) a stop codon (if not present in the

20 ORF); (v) a 3'UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the nucleic acid sequence according to SEQ ID NO:138; and (vi) a polyA tail (*e.g.*, SEQ ID NO:195 or SEQ ID NO:211). In some embodiments, the present disclosure relates to a polynucleotide sequence comprising, from 5' end to 3' end: (i) a 5' terminal cap (*e.g.*, m⁷Gp-ppGm-A, *e.g.*, Cap1, or an analog thereof); (ii) a 5'UTR; (iii) an ORF; (iv) a stop

25 codon (if not present in the ORF); (v) a 3'UTR having a nucleotide sequence consisting of, consisting essentially of, or comprising the nucleic acid sequence according to SEQ ID NO:139; and (vi) a polyA tail (*e.g.*, SEQ ID NO:195 or SEQ ID NO:211).

In some embodiments, the 5'UTR has a nucleic acid sequence consisting of, consisting essentially of, or comprising the nucleic acid sequence according to any one of

30 SEQ ID NOs:50–79. In some embodiments, the ORF encodes a polypeptide having an amino acid sequence according to any one of SEQ ID NOs:1–11. In some embodiments, the ORF has an amino acid sequence having at least about 60%, at least about 65%, at least 70%, at

least 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 99%, or 100% identity to any one of SEQ ID NOs:20–33 or 41–44. In some embodiments, the polyA tail is a 100-nucleotide polyA tail (*e.g.*, SEQ ID NO: 195 or SEQ ID NO:211). In some embodiments, all uracils in the polynucleotide are

5 N1-methylpseudouracils or 5-methoxyuracil. In some embodiments, the polynucleotide is included in a delivery agent comprising LNP-1A, LNP-1B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B.

In another aspect, the LNP compositions of the disclosure are used in a method of treating argininosuccinic aciduria in a subject.

10 In an aspect, an LNP composition comprising a polynucleotide disclosed herein encoding an ASL polypeptide, *e.g.*, as described herein, can be administered with an additional agent, *e.g.*, as described herein.

11. MicroRNA (miRNA) Binding Sites

Polynucleotides of the present disclosure can include regulatory elements, for
15 example, microRNA (miRNA) binding sites, transcription factor binding sites, structured mRNA sequences and/or motifs, artificial binding sites engineered to act as pseudo-receptors for endogenous nucleic acid binding molecules, and combinations thereof. In some embodiments, polynucleotides including such regulatory elements are referred to as including “sensor sequences”.

20 In some embodiments, a polynucleotide (*e.g.*, a ribonucleic acid (RNA), *e.g.*, a messenger RNA (mRNA)) of the present disclosure comprises an open reading frame (ORF) encoding a polypeptide of interest and further comprises one or more miRNA binding site(s). Inclusion or incorporation of miRNA binding site(s) provides for regulation of polynucleotides of the present disclosure, and in turn, of the polypeptides encoded therefrom,
25 based on tissue-specific and/or cell-type specific expression of naturally-occurring miRNAs.

The present disclosure also provides pharmaceutical compositions and formulations that comprise any of the polynucleotides described above. In some embodiments, the composition or formulation further comprises a delivery agent.

In some embodiments, the composition or formulation can contain a polynucleotide
30 comprising a sequence optimized nucleic acid sequence disclosed herein which encodes a

polypeptide. In some embodiments, the composition or formulation can contain a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a polynucleotide (*e.g.*, an ORF) having significant sequence identity to a sequence optimized nucleic acid sequence disclosed herein which encodes a polypeptide. In some embodiments, the polynucleotide further
 5 comprises a miRNA binding site, *e.g.*, a miRNA binding site that binds to miR-126, miR-142, miR-144, miR-146, miR-150, miR-155, miR-16, miR-21, miR-223, miR-24, miR-27 and miR-26a.

A miRNA, *e.g.*, a natural-occurring miRNA, is a 19-25 nucleotide long noncoding RNA that binds to a polynucleotide and down-regulates gene expression either by reducing
 10 stability or by inhibiting translation of the polynucleotide. A miRNA sequence comprises a “seed” region, *i.e.*, a sequence in the region of positions 2-8 of the mature miRNA. A miRNA seed can comprise positions 2-8 or 2-7 of the mature miRNA.

MicroRNAs derive enzymatically from regions of RNA transcripts that fold back on themselves to form short hairpin structures often termed a pre-miRNA (precursor-miRNA).
 15 A pre-miRNA typically has a two-nucleotide overhang at its 3' end, and has 3' hydroxyl and 5' phosphate groups. This precursor-mRNA is processed in the nucleus and subsequently transported to the cytoplasm where it is further processed by DICER (a RNase III enzyme), to form a mature microRNA of approximately 22 nucleotides. The mature microRNA is then incorporated into a ribonuclear particle to form the RNA-induced silencing complex, RISC,
 20 which mediates gene silencing. Art-recognized nomenclature for mature miRNAs typically designates the arm of the pre-miRNA from which the mature miRNA derives; “5p” means the microRNA is from the 5 prime arm of the pre-miRNA hairpin and “3p” means the microRNA is from the 3 prime end of the pre-miRNA hairpin. A miR referred to by number herein can refer to either of the two mature microRNAs originating from opposite arms of the same pre-
 25 miRNA (*e.g.*, either the 3p or 5p microRNA). All miRs referred to herein are intended to include both the 3p and 5p arms/sequences, unless particularly specified by the 3p or 5p designation.

As used herein, the term “microRNA (miRNA or miR) binding site” refers to a sequence within a polynucleotide, *e.g.*, within a DNA or within an RNA transcript, including
 30 in the 5'UTR and/or 3'UTR, that has sufficient complementarity to all or a region of a miRNA to interact with, associate with or bind to the miRNA. In some embodiments, a polynucleotide of the present disclosure comprising an ORF encoding a polypeptide of

interest and further comprises one or more miRNA binding site(s). In exemplary embodiments, a 5' UTR and/or 3' UTR of the polynucleotide (*e.g.*, a ribonucleic acid (RNA), *e.g.*, a messenger RNA (mRNA)) comprises the one or more miRNA binding site(s).

5 A miRNA binding site having sufficient complementarity to a miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated regulation of a polynucleotide, *e.g.*, miRNA-mediated translational repression or degradation of the polynucleotide. In exemplary aspects of the present disclosure, a miRNA binding site having sufficient complementarity to the miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated degradation of the polynucleotide, *e.g.*, miRNA-guided RNA-
10 induced silencing complex (RISC)-mediated cleavage of mRNA. The miRNA binding site can have complementarity to, for example, a 19-25 nucleotide long miRNA sequence, to a 19-23 nucleotide long miRNA sequence, or to a 22 nucleotide long miRNA sequence. A miRNA binding site can be complementary to only a portion of a miRNA, *e.g.*, to a portion less than 1, 2, 3, or 4 nucleotides of the full length of a naturally-occurring miRNA sequence,
15 or to a portion less than 1, 2, 3, or 4 nucleotides shorter than a naturally-occurring miRNA sequence. Full or complete complementarity (*e.g.*, full complementarity or complete complementarity over all or a significant portion of the length of a naturally-occurring miRNA) is preferred when the desired regulation is mRNA degradation.

In some embodiments, a miRNA binding site includes a sequence that has
20 complementarity (*e.g.*, partial or complete complementarity) with an miRNA seed sequence. In some embodiments, the miRNA binding site includes a sequence that has complete complementarity with a miRNA seed sequence. In some embodiments, a miRNA binding site includes a sequence that has complementarity (*e.g.*, partial or complete complementarity) with an miRNA sequence. In some embodiments, the miRNA binding site includes a
25 sequence that has complete complementarity with a miRNA sequence. In other embodiments, the sequence is not completely complementary. In some embodiments, a miRNA binding site has complete complementarity with a miRNA sequence but for 1, 2, or 3 nucleotide substitutions, terminal additions, and/or truncations.

In some embodiments, the miRNA binding site is the same length as the
30 corresponding miRNA. In other embodiments, the miRNA binding site is one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve nucleotide(s) shorter than the corresponding miRNA at the 5' terminus, the 3' terminus, or both. In still other embodiments,

the microRNA binding site is two nucleotides shorter than the corresponding microRNA at the 5' terminus, the 3' terminus, or both. The miRNA binding sites that are shorter than the corresponding miRNAs are still capable of degrading the mRNA incorporating one or more of the miRNA binding sites or preventing the mRNA from translation.

5 In some embodiments, the miRNA binding site binds the corresponding mature miRNA that is part of an active RISC containing Dicer. In some embodiments, binding of the miRNA binding site to the corresponding miRNA in RISC degrades the mRNA containing the miRNA binding site or prevents the mRNA from being translated. In some embodiments, the miRNA binding site has sufficient complementarity to miRNA so that a RISC complex
10 comprising the miRNA cleaves the polynucleotide comprising the miRNA binding site. In other embodiments, the miRNA binding site has imperfect complementarity so that a RISC complex comprising the miRNA induces instability in the polynucleotide comprising the miRNA binding site. In some embodiments, the miRNA binding site has imperfect complementarity so that a RISC complex comprising the miRNA represses transcription of
15 the polynucleotide comprising the miRNA binding site.

In some embodiments, the miRNA binding site has one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve mismatch(es) from the corresponding miRNA.

In some embodiments, the miRNA binding site has at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about
20 fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least about twenty-one contiguous nucleotides complementary to at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least
25 about twenty-one, respectively, contiguous nucleotides of the corresponding miRNA.

By engineering one or more miRNA binding sites into a polynucleotide of the present disclosure, the polynucleotide can be targeted for degradation or reduced translation, provided the miRNA in question is available. This can reduce off-target effects upon delivery of the polynucleotide. For example, if a polynucleotide of the present disclosure is
30 not intended to be delivered to a tissue or cell but ends up in said tissue or cell, then a miRNA abundant in the tissue or cell can inhibit the expression of the gene of interest if one or

multiple binding sites of the miRNA are engineered into the 5' UTR and/or 3' UTR of the polynucleotide. Thus, in some embodiments, incorporation of one or more miRNA binding sites into an mRNA of the disclosure may reduce the hazard of off-target effects upon nucleic acid molecule delivery and/or enable tissue-specific regulation of expression of a polypeptide encoded by the mRNA. In yet other embodiments, incorporation of one or more miRNA binding sites into an mRNA of the disclosure can modulate immune responses upon nucleic acid delivery *in vivo*. In further embodiments, incorporation of one or more miRNA binding sites into an mRNA of the disclosure can modulate accelerated blood clearance (ABC) of lipid-comprising compounds and compositions described herein.

Conversely, miRNA binding sites can be removed from polynucleotide sequences in which they naturally occur to increase protein expression in specific tissues. For example, a binding site for a specific miRNA can be removed from a polynucleotide to improve protein expression in tissues or cells containing the miRNA.

Regulation of expression in multiple tissues can be accomplished through introduction or removal of one or more miRNA binding sites, *e.g.*, one or more distinct miRNA binding sites. The decision whether to remove or insert a miRNA binding site can be made based on miRNA expression patterns and/or their profilings in tissues and/or cells in development and/or disease. Identification of miRNAs, miRNA binding sites, and their expression patterns and role in biology have been reported (*e.g.*, Bonauer et al., Curr. Drug Targets 2010 11:943-949; Anand and Cheresch Curr. Op. Hematol. 2011 18:171-176; Contreras and Rao Leukemia 2012 26:404-413 (2011 Dec 20. doi: 10.1038/leu.2011.356); Bartel, Cell, 2009 136:215-233; Landgraf et al, Cell, 2007 129:1401-1414; Gentner and Naldini, Tissue Antigens, 2012 80:393-403 and all references therein; each of which is incorporated herein by reference in its entirety).

Examples of tissues where miRNA are known to regulate mRNA, and thereby protein expression, include, but are not limited to, liver (miR-122), muscle (miR-133, miR-206, miR-208), endothelial cells (miR-17-92, miR-126), myeloid cells (miR-142-3p, miR-142-5p, miR-16, miR-21, miR-223, miR-24, miR-27), adipose tissue (let-7, miR-30c), heart (miR-1d, miR-149), kidney (miR-192, miR-194, miR-204), and lung epithelial cells (let-7, miR-133, miR-126).

Specifically, miRNAs are known to be differentially expressed in immune cells (also called hematopoietic cells), such as antigen presenting cells (APCs) (*e.g.*, dendritic cells and macrophages), macrophages, monocytes, B lymphocytes, T lymphocytes, granulocytes, natural killer cells, etc. Immune cell specific miRNAs are involved in immunogenicity, autoimmunity, the immune-response to infection, inflammation, as well as unwanted immune response after gene therapy and tissue/organ transplantation. Immune cells specific miRNAs also regulate many aspects of development, proliferation, differentiation and apoptosis of hematopoietic cells (immune cells). For example, miR-142 and miR-146 are exclusively expressed in immune cells, particularly abundant in myeloid dendritic cells. It has been demonstrated that the immune response to a polynucleotide can be shut-off by adding miR-142 binding sites to the 3'-UTR of the polynucleotide, enabling more stable gene transfer in tissues and cells. miR-142 efficiently degrades exogenous polynucleotides in antigen presenting cells and suppresses cytotoxic elimination of transduced cells (*e.g.*, Annoni A et al., *Blood*, 2009, 114, 5152-5161; Brown BD, et al., *Nat. Med.* 2006, 12(5), 585-591; Brown BD, et al., *Blood*, 2007, 110(13): 4144-4152, each of which is incorporated herein by reference in its entirety).

An antigen-mediated immune response can refer to an immune response triggered by foreign antigens, which, when entering an organism, are processed by the antigen presenting cells and displayed on the surface of the antigen presenting cells. T cells can recognize the presented antigen and induce a cytotoxic elimination of cells that express the antigen.

Introducing one or more (*e.g.*, one, two, or three) miR-142 binding sites into the 5' UTR and/or 3'UTR of a polynucleotide of the present disclosure can selectively repress gene expression in antigen presenting cells through miR-142 mediated degradation, limiting antigen presentation in antigen presenting cells (*e.g.*, dendritic cells) and thereby preventing antigen-mediated immune response after the delivery of the polynucleotide. The polynucleotide is then stably expressed in target tissues or cells without triggering cytotoxic elimination.

In some embodiments, it may be beneficial to target the same cell type with multiple miRs and to incorporate binding sites to each of the 3p and 5p arm if both are abundant (*e.g.*, both miR-142-3p and miR142-5p are abundant in hematopoietic stem cells). Thus, in certain embodiments, polynucleotides of the present disclosure contain two or more (*e.g.*, two, three, four or more) miR bindings sites from: (i) the group consisting of miR-142, miR-144, miR-

150, miR-155 and miR-223 (which are expressed in many hematopoietic cells); or (ii) the group consisting of miR-142, miR150, miR-16 and miR-223 (which are expressed in B cells); or the group consisting of miR-223, miR-451, miR-26a, miR-16 (which are expressed in progenitor hematopoietic cells).

5 In some embodiments, it may also be beneficial to combine various miRs such that multiple cell types of interest are targeted at the same time (*e.g.*, miR-142 and miR-126 to target many cells of the hematopoietic lineage and endothelial cells). Thus, for example, in certain embodiments, polynucleotides of the present disclosure comprise two or more (*e.g.*, two, three, four or more) miRNA bindings sites, wherein: (i) at least one of the miRs targets
 10 cells of the hematopoietic lineage (*e.g.*, miR-142, miR-144, miR-150, miR-155 or miR-223) and at least one of the miRs targets plasmacytoid dendritic cells, platelets or endothelial cells (*e.g.*, miR-126); or (ii) at least one of the miRs targets B cells (*e.g.*, miR-142, miR150, miR-16 or miR-223) and at least one of the miRs targets plasmacytoid dendritic cells, platelets or endothelial cells (*e.g.*, miR-126); or (iii) at least one of the miRs targets progenitor
 15 hematopoietic cells (*e.g.*, miR-223, miR-451, miR-26a or miR-16) and at least one of the miRs targets plasmacytoid dendritic cells, platelets or endothelial cells (*e.g.*, miR-126); or (iv) at least one of the miRs targets cells of the hematopoietic lineage (*e.g.*, miR-142, miR-144, miR-150, miR-155 or miR-223), at least one of the miRs targets B cells (*e.g.*, miR-142, miR150, miR-16 or miR-223) and at least one of the miRs targets plasmacytoid dendritic
 20 cells, platelets or endothelial cells (*e.g.*, miR-126); or any other possible combination of the foregoing four classes of miR binding sites (*i.e.*, those targeting the hematopoietic lineage, those targeting B cells, those targeting progenitor hematopoietic cells and/or those targeting plasmacytoid dendritic cells/platelets/endothelial cells).

In some embodiments, to modulate immune responses, polynucleotides of the present
 25 disclosure can comprise one or more miRNA binding sequences that bind to one or more miRs that are expressed in conventional immune cells or any cell that expresses TLR7 and/or TLR8 and secrete pro-inflammatory cytokines and/or chemokines (*e.g.*, in immune cells of peripheral lymphoid organs and/or splenocytes and/or endothelial cells). It has now been
 30 discovered that incorporation into an mRNA of one or more miRs that are expressed in conventional immune cells or any cell that expresses TLR7 and/or TLR8 and secrete pro-inflammatory cytokines and/or chemokines (*e.g.*, in immune cells of peripheral lymphoid organs and/or splenocytes and/or endothelial cells) reduces or inhibits immune cell activation

(*e.g.*, B cell activation, as measured by frequency of activated B cells) and/or cytokine production (*e.g.*, production of IL-6, IFN- γ and/or TNF α). Furthermore, it has now been discovered that incorporation into an mRNA of one or more miRs that are expressed in conventional immune cells or any cell that expresses TLR7 and/or TLR8 and secrete pro-inflammatory cytokines and/or chemokines (*e.g.*, in immune cells of peripheral lymphoid organs and/or splenocytes and/or endothelial cells) can reduce or inhibit an anti-drug antibody (ADA) response against a protein of interest encoded by the mRNA.

In some embodiments, to modulate accelerated blood clearance of a polynucleotide delivered in a lipid-comprising compound or composition, polynucleotides of the present disclosure can comprise one or more miR binding sequences that bind to one or more miRNAs expressed in conventional immune cells or any cell that expresses TLR7 and/or TLR8 and secrete pro-inflammatory cytokines and/or chemokines (*e.g.*, in immune cells of peripheral lymphoid organs and/or splenocytes and/or endothelial cells). It has now been discovered that incorporation into an mRNA of one or more miR binding sites reduces or inhibits accelerated blood clearance (ABC) of the lipid-comprising compound or composition for use in delivering the mRNA. Furthermore, it has now been discovered that incorporation of one or more miR binding sites into an mRNA reduces serum levels of anti-PEG anti-IgM (*e.g.*, reduces or inhibits the acute production of IgMs that recognize polyethylene glycol (PEG) by B cells) and/or reduces or inhibits proliferation and/or activation of plasmacytoid dendritic cells following administration of a lipid-comprising compound or composition comprising the mRNA.

In some embodiments, miR sequences may correspond to any known microRNA expressed in immune cells, including but not limited to those taught in US Publication US2005/0261218 and US Publication US2005/0059005, the contents of which are incorporated herein by reference in their entirety. Non-limiting examples of miRs expressed in immune cells include those expressed in spleen cells, myeloid cells, dendritic cells, plasmacytoid dendritic cells, B cells, T cells and/or macrophages. For example, miR-142-3p, miR-142-5p, miR-16, miR-21, miR-223, miR-24 and miR-27 are expressed in myeloid cells, miR-155 is expressed in dendritic cells, B cells and T cells, miR-146 is upregulated in macrophages upon TLR stimulation and miR-126 is expressed in plasmacytoid dendritic cells. In certain embodiments, the miR(s) is expressed abundantly or preferentially in immune cells. For example, miR-142 (miR-142-3p and/or miR-142-5p), miR-126 (miR-126-

3p and/or miR-126-5p), miR-146 (miR-146-3p and/or miR-146-5p) and miR-155 (miR-155-3p and/or miR155-5p) are expressed abundantly in immune cells. These microRNA sequences are known in the art and, thus, one of ordinary skill in the art can readily design binding sequences or target sequences to which these microRNAs will bind based upon

5 Watson-Crick complementarity.

In some embodiments, the polynucleotide of the present disclosure comprises three copies of the same miRNA binding site. In certain embodiments, use of three copies of the same miR binding site can exhibit beneficial properties as compared to use of a single miRNA binding site.

10 In some embodiments, the polynucleotide of the present disclosure comprises two or more (*e.g.*, two, three, four) copies of at least two different miR binding sites expressed in immune cells.

In some embodiments, the polynucleotide of the present disclosure comprises at least two miR binding sites for microRNAs expressed in immune cells, wherein one of the miR
15 binding sites is for miR-142-3p. In various embodiments, the polynucleotide of the present disclosure comprises binding sites for miR-142-3p and miR-155 (miR-155-3p or miR-155-5p), miR-142-3p and miR-146 (miR-146-3 or miR-146-5p), or miR-142-3p and miR-126 (miR-126-3p or miR-126-5p).

In some embodiments, the polynucleotide of the present disclosure comprises at least
20 two miR binding sites for microRNAs expressed in immune cells, wherein one of the miR binding sites is for miR-126-3p. In various embodiments, the polynucleotide of the present disclosure comprises binding sites for miR-126-3p and miR-155 (miR-155-3p or miR-155-5p), miR-126-3p and miR-146 (miR-146-3p or miR-146-5p), or miR-126-3p and miR-142 (miR-142-3p or miR-142-5p).

25 In some embodiments, the polynucleotide of the present disclosure comprises at least two miR binding sites for microRNAs expressed in immune cells, wherein one of the miR binding sites is for miR-142-5p. In various embodiments, the polynucleotide of the present disclosure comprises binding sites for miR-142-5p and miR-155 (miR-155-3p or miR-155-5p), miR-142-5p and miR-146 (miR-146-3 or miR-146-5p), or miR-142-5p and miR-126
30 (miR-126-3p or miR-126-5p).

In some embodiments, the polynucleotide of the present disclosure comprises at least two miR binding sites for microRNAs expressed in immune cells, wherein one of the miR binding sites is for miR-155-5p. In various embodiments, the polynucleotide of the present disclosure comprises binding sites for miR-155-5p and miR-142 (miR-142-3p or miR-142-5p), miR-155-5p and miR-146 (miR-146-3 or miR-146-5p), or miR-155-5p and miR-126 (miR-126-3p or miR-126-5p).

In some embodiments, a polynucleotide of the present disclosure comprises a miRNA binding site, wherein the miRNA binding site comprises one or more nucleotide sequences selected from Table 5, including one or more copies of any one or more of the miRNA binding site sequences. In some embodiments, a polynucleotide of the present disclosure further comprises at least one, two, three, four, five, six, seven, eight, nine, ten, or more of the same or different miRNA binding sites selected from Table 5, including any combination thereof.

In some embodiments, the miRNA binding site binds to miR-142 or is complementary to miR-142. In some embodiments, the miR-142 comprises SEQ ID NO:172. In some embodiments, the miRNA binding site binds to miR-142-3p or miR-142-5p. In some embodiments, the miR-142-3p binding site comprises SEQ ID NO:174. In some embodiments, the miR-142-5p binding site comprises SEQ ID NO:210. In some embodiments, the miRNA binding site comprises a nucleotide sequence at least 80%, at least 85%, at least 90%, at least 95%, or 100% identical to SEQ ID NO:174 or SEQ ID NO:210.

In some embodiments, the miRNA binding site binds to miR-126 or is complementary to miR-126. In some embodiments, the miR-126 comprises SEQ ID NO:150. In some embodiments, the miRNA binding site binds to miR-126-3p or miR-126-5p. In some embodiments, the miR-126-3p binding site comprises SEQ ID NO:152. In some embodiments, the miR-126-5p binding site comprises SEQ ID NO:154. In some embodiments, the miRNA binding site comprises a nucleotide sequence at least 80%, at least 85%, at least 90%, at least 95%, or 100% identical to SEQ ID NO:152 or SEQ ID NO:154.

In some embodiments, the 3' UTR comprises two miRNA binding sites, wherein a first miRNA binding site binds to miR-142 and a second miRNA binding site binds to miR-126.

TABLE 5. miR-142, miR-126, and miR-142 and miR-126 binding sites

SEQ ID NO.	Description	Sequence
172	miR-142	GACAGUGCAGUCACCCAUAAAGUAGAAAGCACUACUAA CAGCACUGGAGGGUGUAGUGUUUCCUACUUUAUGGAUG AGUGUACUGUG
173	miR-142-3p	UGUAGUGUUUCCUACUUUAUGGA
174	miR-142-3p binding site	UCCAUAAGUAGGAAACACUACA
175	miR-142-5p	CAUAAAGUAGAAAGCACUACU
210	miR-142-5p binding site	AGUAGUGCUUUCUACUUUAUG
150	miR-126	CGCUGGCGACGGGACAUUAUACUUUUGGUACGCGCUG UGACACUUCAAACUCGUACCGUGAGUAAUAAUGCGCCG UCCACGGCA
151	miR-126-3p	UCGUACCGUGAGUAAUAAUGCG
152	miR-126-3p binding site	CGCAUUAUACUCACGGUACGA
153	miR-126-5p	CAUUAUACUUUUGGUACGCG
154	miR-126-5p binding site	CGCGUACCAAAAGUAAUAAUG

In some embodiments, a miRNA binding site is inserted in the polynucleotide of the present disclosure in any position of the polynucleotide (*e.g.*, the 5' UTR and/or 3' UTR). In some embodiments, the 5' UTR comprises a miRNA binding site. In some embodiments, the 3' UTR comprises a miRNA binding site. In some embodiments, the 5' UTR and the 3' UTR comprise a miRNA binding site. The insertion site in the polynucleotide can be anywhere in the polynucleotide as long as the insertion of the miRNA binding site in the polynucleotide does not interfere with the translation of a functional polypeptide in the absence of the corresponding miRNA; and in the presence of the miRNA, the insertion of the miRNA binding site in the polynucleotide and the binding of the miRNA binding site to the corresponding miRNA are capable of degrading the polynucleotide or preventing the translation of the polynucleotide.

In some embodiments, a miRNA binding site is inserted in at least about 30 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the present disclosure comprising the ORF. In some embodiments, a miRNA binding site is inserted in at least about 10 nucleotides, at least about 15 nucleotides, at least about 20 nucleotides, at least about 25 nucleotides, at least about 30 nucleotides, at least about 35 nucleotides, at least about 40 nucleotides, at least about 45 nucleotides, at least about 50 nucleotides, at least about 55 nucleotides, at least about 60 nucleotides, at least about 65 nucleotides, at least about 70 nucleotides, at least about 75 nucleotides, at least about 80 nucleotides, at least about 85 nucleotides, at least about 90 nucleotides, at least about 95 nucleotides, or at least about 100 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the

present disclosure. In some embodiments, a miRNA binding site is inserted in about 10 nucleotides to about 100 nucleotides, about 20 nucleotides to about 90 nucleotides, about 30 nucleotides to about 80 nucleotides, about 40 nucleotides to about 70 nucleotides, about 50 nucleotides to about 60 nucleotides, about 45 nucleotides to about 65 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the present disclosure.

In some embodiments, a miRNA binding site is inserted within the 3' UTR immediately following the stop codon of the coding region within the polynucleotide of the present disclosure, *e.g.*, mRNA. In some embodiments, if there are multiple copies of a stop codon in the construct, a miRNA binding site is inserted immediately following the final stop codon. In some embodiments, a miRNA binding site is inserted further downstream of the stop codon, in which case there are 3' UTR bases between the stop codon and the miRNA binding site(s).

In some embodiments, one or more miRNA binding sites can be positioned within the 5' UTR at one or more possible insertion sites.

In some embodiments, a codon optimized open reading frame encoding a polypeptide of interest comprises a stop codon and the at least one microRNA binding site is located within the 3' UTR 1-100 nucleotides after the stop codon. In some embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a stop codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 3' UTR 30-50 nucleotides after the stop codon. In some embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a stop codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 3' UTR at least 50 nucleotides after the stop codon. In other embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a stop codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 3' UTR immediately after the stop codon, or within the 3' UTR 15-20 nucleotides after the stop codon or within the 3' UTR 70-80 nucleotides after the stop codon. In other embodiments, the 3' UTR comprises more than one miRNA binding site (*e.g.*, 2-4 miRNA binding sites), wherein there can be a spacer region (*e.g.*, of 10-100, 20-70 or 30-50 nucleotides in length) between each miRNA binding site. In some embodiments, the 3' UTR comprises a spacer region between the end of the miRNA binding site(s) and the poly A tail nucleotides. For example, a spacer region of 10-100, 20-70 or 30-50 nucleotides in length

can be situated between the end of the miRNA binding site(s) and the beginning of the poly A tail.

In some embodiments, a codon optimized open reading frame encoding a polypeptide of interest comprises a start codon and the at least one microRNA binding site is located within the 5' UTR 1-100 nucleotides before (upstream of) the start codon. In some embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a start codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 5' UTR 10-50 nucleotides before (upstream of) the start codon. In some embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a start codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 5' UTR at least 25 nucleotides before (upstream of) the start codon. In other embodiments, the codon optimized open reading frame encoding the polypeptide of interest comprises a start codon and the at least one microRNA binding site for a miR expressed in immune cells is located within the 5' UTR immediately before the start codon, or within the 5' UTR 15-20 nucleotides before the start codon or within the 5' UTR 70-80 nucleotides before the start codon. In other embodiments, the 5' UTR comprises more than one miRNA binding site (*e.g.*, 2-4 miRNA binding sites), wherein there can be a spacer region (*e.g.*, of 10-100, 20-70 or 30-50 nucleotides in length) between each miRNA binding site.

In some embodiments, the 3' UTR comprises more than one stop codon, wherein at least one miRNA binding site is positioned downstream of the stop codons. For example, a 3' UTR can comprise 1, 2 or 3 stop codons. Non-limiting examples of triple stop codons that can be used include: UGAUAAUAG (SEQ ID NO:182), UGAUAGUAA (SEQ ID NO:183), UAAUGAUAG (SEQ ID NO:184), UGAUAAUAA (SEQ ID NO:185), UGAUAGUAG (SEQ ID NO:186), UAAUGAUGA (SEQ ID NO:187), UAAUAGUAG (SEQ ID NO:188), UGAUGAUGA (SEQ ID NO:179), UAAUAAUAA (SEQ ID NO:180), and UAGUAGUAG (SEQ ID NO:181). Within a 3' UTR, for example, 1, 2, 3 or 4 miRNA binding sites, *e.g.*, miR-142-3p binding sites, can be positioned immediately adjacent to the stop codon(s) or at any number of nucleotides downstream of the final stop codon. When the 3' UTR comprises multiple miRNA binding sites, these binding sites can be positioned directly next to each other in the construct (*i.e.*, one after the other) or, alternatively, spacer nucleotides can be positioned between each binding site.

In some embodiments, the 3' UTR comprises three stop codons with a single miR-142-3p binding site located downstream of the 3rd stop codon.

In some embodiments, the polynucleotide of the present disclosure comprises a 5' UTR, a codon optimized open reading frame encoding a polypeptide of interest, a 3' UTR comprising the at least one miRNA binding site for a miR expressed in immune cells, and a 3' tailing region of linked nucleosides. In various embodiments, the 3' UTR comprises 1-4, at least two, one, two, three or four miRNA binding sites for miRs expressed in immune cells, preferably abundantly or preferentially expressed in immune cells.

In some embodiments, the at least one miRNA expressed in immune cells is a miR-142-3p microRNA binding site. In some embodiments, the miR-142-3p microRNA binding site comprises the sequence shown in SEQ ID NO:174.

In some embodiments, the at least one miRNA expressed in immune cells is a miR-126 microRNA binding site. In some embodiments, the miR-126 binding site is a miR-126-3p binding site. In some embodiments, the miR-126-3p microRNA binding site comprises the sequence shown in SEQ ID NO:152.

Non-limiting exemplary sequences for miRs to which a microRNA binding site(s) of the disclosure can bind include the following: miR-142-3p (SEQ ID NO:173), miR-142-5p (SEQ ID NO:175), miR-146-3p (SEQ ID NO:155), miR-146-5p (SEQ ID NO:156), miR-155-3p (SEQ ID NO:157), miR-155-5p (SEQ ID NO:158), miR-126-3p (SEQ ID NO:151), miR-126-5p (SEQ ID NO:153), miR-16-3p (SEQ ID NO:159), miR-16-5p (SEQ ID NO:160), miR-21-3p (SEQ ID NO:161), miR-21-5p (SEQ ID NO:162), miR-223-3p (SEQ ID NO:163), miR-223-5p (SEQ ID NO:164), miR-24-3p (SEQ ID NO:165), miR-24-5p (SEQ ID NO:166), miR-27-3p (SEQ ID NO:167) and miR-27-5p (SEQ ID NO:168). Other suitable miR sequences expressed in immune cells (*e.g.*, abundantly or preferentially expressed in immune cells) are known and available in the art, for example at the University of Manchester's microRNA database, miRBase. Sites that bind any of the aforementioned miRs can be designed based on Watson-Crick complementarity to the miR, typically 100% complementarity to the miR, and inserted into an mRNA construct of the disclosure as described herein.

In some embodiments, a polynucleotide of the present disclosure (*e.g.*, and mRNA, *e.g.*, the 3' UTR thereof) can comprise at least one miRNA binding site to thereby reduce or

inhibit accelerated blood clearance, for example by reducing or inhibiting production of IgMs, *e.g.*, against PEG, by B cells and/or reducing or inhibiting proliferation and/or activation of pDCs, and can comprise at least one miRNA binding site for modulating tissue expression of an encoded protein of interest.

5 miRNA gene regulation can be influenced by the sequence surrounding the miRNA such as, but not limited to, the species of the surrounding sequence, the type of sequence (*e.g.*, heterologous, homologous, exogenous, endogenous, or artificial), regulatory elements in the surrounding sequence and/or structural elements in the surrounding sequence. The miRNA can be influenced by the 5'UTR and/or 3'UTR. As a non-limiting example, a non-
10 human 3'UTR can increase the regulatory effect of the miRNA sequence on the expression of a polypeptide of interest compared to a human 3' UTR of the same sequence type.

In some embodiments, other regulatory elements and/or structural elements of the 5' UTR can influence miRNA mediated gene regulation. One example of a regulatory element and/or structural element is a structured IRES (Internal Ribosome Entry Site) in the 5' UTR,
15 which is necessary for the binding of translational elongation factors to initiate protein translation. EIF4A2 binding to this secondarily structured element in the 5'-UTR is necessary for miRNA mediated gene expression (Meijer HA et al., *Science*, 2013, 340, 82-85, herein incorporated by reference in its entirety). The polynucleotides of the present disclosure can further include this structured 5' UTR in order to enhance microRNA mediated gene
20 regulation.

At least one miRNA binding site can be engineered into the 3' UTR of a polynucleotide of the present disclosure. In this context, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or more miRNA binding sites can be engineered into a 3' UTR of a polynucleotide of the present
25 disclosure. For example, 1 to 10, 1 to 9, 1 to 8, 1 to 7, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 2, or 1 miRNA binding sites can be engineered into the 3' UTR of a polynucleotide of the present disclosure. In some embodiments, miRNA binding sites incorporated into a polynucleotide of the present disclosure can be the same or can be different miRNA sites. A combination of different miRNA binding sites incorporated into a polynucleotide of the present disclosure
30 can include combinations in which more than one copy of any of the different miRNA sites are incorporated. In some embodiments, miRNA binding sites incorporated into a polynucleotide of the present disclosure can target the same or different tissues in the body.

As a non-limiting example, through the introduction of tissue-, cell-type-, or disease-specific miRNA binding sites in the 3'-UTR of a polynucleotide of the present disclosure, the degree of expression in specific cell types (*e.g.*, myeloid cells, endothelial cells, etc.) can be reduced.

5 In some embodiments, a miRNA binding site can be engineered near the 5' terminus of the 3' UTR, about halfway between the 5' terminus and 3' terminus of the 3'UTR and/or near the 3' terminus of the 3' UTR in a polynucleotide of the present disclosure. As a non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR and about halfway between the 5' terminus and 3' terminus of the 3'UTR. As another non-limiting example, a miRNA binding site can be engineered near the 3' terminus of the 3'UTR
10 and about halfway between the 5' terminus and 3' terminus of the 3' UTR. As yet another non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3' UTR and near the 3' terminus of the 3' UTR.

In some embodiments, a 3'UTR can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 miRNA binding sites. The miRNA binding sites can be complementary to a miRNA, miRNA seed
15 sequence, and/or miRNA sequences flanking the seed sequence.

In some embodiments, the expression of a polynucleotide of the present disclosure can be controlled by incorporating at least one sensor sequence in the polynucleotide and Formulating the polynucleotide for administration. As a non-limiting example, a polynucleotide of the present disclosure can be targeted to a tissue or cell by incorporating a
20 miRNA binding site and Formulating the polynucleotide in a lipid nanoparticle comprising an ionizable amino lipid, including any of the lipids described herein.

A polynucleotide of the present disclosure can be engineered for more targeted expression in specific tissues, cell types, or biological conditions based on the expression patterns of miRNAs in the different tissues, cell types, or biological conditions. Through
25 introduction of tissue-specific miRNA binding sites, a polynucleotide of the present disclosure can be designed for optimal protein expression in a tissue or cell, or in the context of a biological condition.

In some embodiments, a polynucleotide of the present disclosure can be designed to incorporate miRNA binding sites that either have 100% identity to known miRNA seed
30 sequences or have less than 100% identity to miRNA seed sequences. In some embodiments, a polynucleotide of the present disclosure can be designed to incorporate miRNA binding

sites that have at least: 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to known miRNA seed sequences. The miRNA seed sequence can be partially mutated to decrease miRNA binding affinity and as such result in reduced downmodulation of the polynucleotide. In essence, the degree of match or mis-match between the miRNA binding site and the miRNA seed can act as a rheostat to more finely tune the ability of the miRNA to modulate protein expression. In addition, mutation in the non-seed region of a miRNA binding site can also impact the ability of a miRNA to modulate protein expression.

In some embodiments, a miRNA sequence can be incorporated into the loop of a stem loop.

In some embodiments, a miRNA seed sequence can be incorporated in the loop of a stem loop and a miRNA binding site can be incorporated into the 5' or 3' stem of the stem loop.

In some embodiments, the miRNA sequence in the 5' UTR can be used to stabilize a polynucleotide of the present disclosure described herein.

In some embodiments, a miRNA sequence in the 5' UTR of a polynucleotide of the present disclosure can be used to decrease the accessibility of the site of translation initiation such as, but not limited to a start codon. See, *e.g.*, Matsuda et al., PLoS One. 2010 11(5):e15057; incorporated herein by reference in its entirety, which used antisense locked nucleic acid (LNA) oligonucleotides and exon-junction complexes (EJC) around a start codon (-4 to +37 where the A of the AUG codons is +1) in order to decrease the accessibility to the first start codon (AUG). Matsuda showed that altering the sequence around the start codon with an LNA or EJC affected the efficiency, length and structural stability of a polynucleotide. A polynucleotide of the present disclosure can comprise a miRNA sequence, instead of the LNA or EJC sequence described by Matsuda et al, near the site of translation initiation in order to decrease the accessibility to the site of translation initiation. The site of translation initiation can be prior to, after or within the miRNA sequence. As a non-limiting example, the site of translation initiation can be located within a miRNA sequence such as a seed sequence or binding site.

In some embodiments, a polynucleotide of the present disclosure can include at least one miRNA in order to dampen the antigen presentation by antigen presenting cells. The miRNA can be the complete miRNA sequence, the miRNA seed sequence, the miRNA

sequence without the seed, or a combination thereof. As a non-limiting example, a miRNA incorporated into a polynucleotide of the present disclosure can be specific to the hematopoietic system. As another non-limiting example, a miRNA incorporated into a polynucleotide of the present disclosure to dampen antigen presentation is miR-142-3p.

5 In some embodiments, a polynucleotide of the present disclosure can include at least one miRNA in order to dampen expression of the encoded polypeptide in a tissue or cell of interest. As a non-limiting example a polynucleotide of the present disclosure can include at least one miR-142-3p binding site, miR-142-3p seed sequence, miR-142-3p binding site without the seed, miR-142-5p binding site, miR-142-5p seed sequence, miR-142-5p binding
10 site without the seed, miR-146 binding site, miR-146 seed sequence and/or miR-146 binding site without the seed sequence.

 In some embodiments, a polynucleotide of the present disclosure can comprise at least one miRNA binding site in the 3'UTR in order to selectively degrade mRNA therapeutics in the immune cells to subdue unwanted immunogenic reactions caused by therapeutic delivery.
15 As a non-limiting example, the miRNA binding site can make a polynucleotide of the present disclosure more unstable in antigen presenting cells. Non-limiting examples of these miRNAs include miR-142-5p, miR-142-3p, miR-146a-5p, and miR-146-3p.

 In some embodiments, a polynucleotide of the present disclosure comprises at least one miRNA sequence in a region of the polynucleotide that can interact with a RNA binding
20 protein.

 In some embodiments, the polynucleotide of the present disclosure (*e.g.*, a RNA, *e.g.*, an mRNA) comprising (i) a sequence-optimized nucleotide sequence (*e.g.*, an ORF) encoding an ASL polypeptide (*e.g.*, the wild-type sequence, functional fragment, or variant thereof) and (ii) a miRNA binding site (*e.g.*, a miRNA binding site that binds to miR-142) and/or a
25 miRNA binding site that binds to miR-126.

12. Regions having a 5' Cap

 The disclosure also includes a polynucleotide that comprises both a 5' cap and a polynucleotide of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide to be expressed).

The 5' cap structure of a natural mRNA is involved in nuclear export, increasing mRNA stability and binds the mRNA Cap Binding Protein (CBP), which is responsible for mRNA stability in the cell and translation competency through the association of CBP with poly(A) binding protein to form the mature cyclic mRNA species. The cap further assists the removal of 5' proximal introns during mRNA splicing.

Endogenous mRNA molecules can be 5'-end capped generating a 5'-ppp-5'-triphosphate linkage between a terminal guanosine cap residue and the 5'-terminal transcribed sense nucleotide of the mRNA molecule. This 5'-guanylate cap can then be methylated to generate an N7-methyl-guanylate residue. The ribose sugars of the terminal and/or antiterminal transcribed nucleotides of the 5' end of the mRNA can optionally also be 2'-O-methylated. 5'-decapping through hydrolysis and cleavage of the guanylate cap structure can target a nucleic acid molecule, such as an mRNA molecule, for degradation.

In some embodiments, the polynucleotides of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide) incorporate a cap moiety. In any of the embodiments disclosed herein, a 5' terminal cap may terminate at the 3' end with an A or G, even if not shown in the disclosure below.

In some embodiments, polynucleotides of the present disclosure comprise a non-hydrolyzable cap structure preventing decapping and thus increasing mRNA half-life. Because cap structure hydrolysis requires cleavage of 5'-ppp-5' phosphodiester linkages, modified nucleotides can be used during the capping reaction. For example, a Vaccinia Capping Enzyme from New England Biolabs (Ipswich, MA) can be used with α -thio-guanosine nucleotides according to the manufacturer's instructions to create a phosphorothioate linkage in the 5'-ppp-5' cap. Additional modified guanosine nucleotides can be used such as α -methyl-phosphonate and seleno-phosphate nucleotides.

Additional modifications include, but are not limited to, 2'-O-methylation of the ribose sugars of 5'-terminal and/or 5'-antiterminal nucleotides of the polynucleotide (as mentioned above) on the 2'-hydroxyl group of the sugar ring. Multiple distinct 5'-cap structures can be used to generate the 5'-cap of a nucleic acid molecule, such as a polynucleotide that functions as an mRNA molecule. Cap analogs, which herein are also referred to as synthetic cap analogs, chemical caps, chemical cap analogs, or structural or functional cap analogs, differ from natural (*i.e.*, endogenous, wild-type or physiological) 5'-

caps in their chemical structure, while retaining cap function. Cap analogs can be chemically (*i.e.*, non-enzymatically) or enzymatically synthesized and/or linked to the polynucleotides of the present disclosure.

For example, the Anti-Reverse Cap Analog (ARCA) cap contains two guanines linked by a 5'-5'-triphosphate group, wherein one guanine contains an N7 methyl group as well as a 3'-O-methyl group (*i.e.*, N7,3'-O-dimethyl-guanosine-5'-triphosphate-5'-guanosine (m⁷G-3'mppp-G; which can equivalently be designated 3' O-Me-m⁷G(5')ppp(5')G). The 3'-O atom of the other, unmodified, guanine becomes linked to the 5'-terminal nucleotide of the capped polynucleotide. The N7- and 3'-O-methylated guanine provides the terminal moiety of the capped polynucleotide.

Another exemplary cap is mCAP, which is similar to ARCA but has a 2'-O-methyl group on guanosine (*i.e.*, N7,2'-O-dimethyl-guanosine-5'-triphosphate-5'-guanosine, m⁷Gm-ppp-G).

Another exemplary cap is m⁷G-ppp-Gm-A (*i.e.*, N7,guanosine-5'-triphosphate-2'-O-dimethyl-guanosine-adenosine).

In some embodiments, the cap is a dinucleotide cap analog. As a non-limiting example, the dinucleotide cap analog can be modified at different phosphate positions with a boranophosphate group or a phosphoroselenoate group such as the dinucleotide cap analogs described in U.S. Patent No. US 8,519,110, the contents of which are herein incorporated by reference in its entirety.

In some embodiments, the cap is a cap analog is a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog known in the art and/or described herein. Non-limiting examples of a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog include a N7-(4-chlorophenoxyethyl)-G(5')ppp(5')G and a N7-(4-chlorophenoxyethyl)-m^{3'-O}G(5')ppp(5')G cap analog (*See, e.g.*, the various cap analogs and the methods of synthesizing cap analogs described in Kore et al. *Bioorganic & Medicinal Chemistry* 2013 21:4570-4574; the contents of which are herein incorporated by reference in its entirety). In some embodiments, a cap analog of the present disclosure is a 4-chloro/bromophenoxyethyl analog.

Polynucleotides of the present disclosure can also be capped post-manufacture (whether IVT or chemical synthesis), using enzymes, in order to generate more authentic 5'-cap structures. As used herein, the phrase "more authentic" refers to a feature that closely mirrors or mimics, either structurally or functionally, an endogenous or wild type feature.

5 That is, a "more authentic" feature is better representative of an endogenous, wild-type, natural or physiological cellular function and/or structure as compared to synthetic features or analogs, etc., of the prior art, or which outperforms the corresponding endogenous, wild-type, natural or physiological feature in one or more respects. Non-limiting examples of more authentic 5'cap structures of the present disclosure are those that, among other things, have
10 enhanced binding of cap binding proteins, increased half-life, reduced susceptibility to 5' endonucleases and/or reduced 5'decapping, as compared to synthetic 5'cap structures known in the art (or to a wild-type, natural or physiological 5'cap structure). For example, recombinant Vaccinia Virus Capping Enzyme and recombinant 2'-O-methyltransferase enzyme can create a canonical 5'-5'-triphosphate linkage between the 5'-terminal nucleotide
15 of a polynucleotide and a guanine cap nucleotide wherein the cap guanine contains an N7 methylation and the 5'-terminal nucleotide of the mRNA contains a 2'-O-methyl. Such a structure is termed the Cap1 structure. This cap results in a higher translational-competency and cellular stability and a reduced activation of cellular pro-inflammatory cytokines, as compared, *e.g.*, to other 5'cap analog structures known in the art. Cap structures include, but
20 are not limited to, 7mG(5')ppp(5')N1pN2p (cap 0), 7mG(5')ppp(5')N1mpNp (cap 1), and 7mG(5')-ppp(5')N1mpN2mp (cap 2).

As a non-limiting example, capping chimeric polynucleotides post-manufacture can be more efficient as nearly 100% of the chimeric polynucleotides can be capped. This is in contrast to ~80% when a cap analog is linked to a chimeric polynucleotide in the course of an
25 *in vitro* transcription reaction.

According to the present disclosure, 5' terminal caps can include endogenous caps or cap analogs. According to the present disclosure, a 5' terminal cap can comprise a guanine analog. Useful guanine analogs include, but are not limited to, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-
30 guanosine, and 2-azido-guanosine.

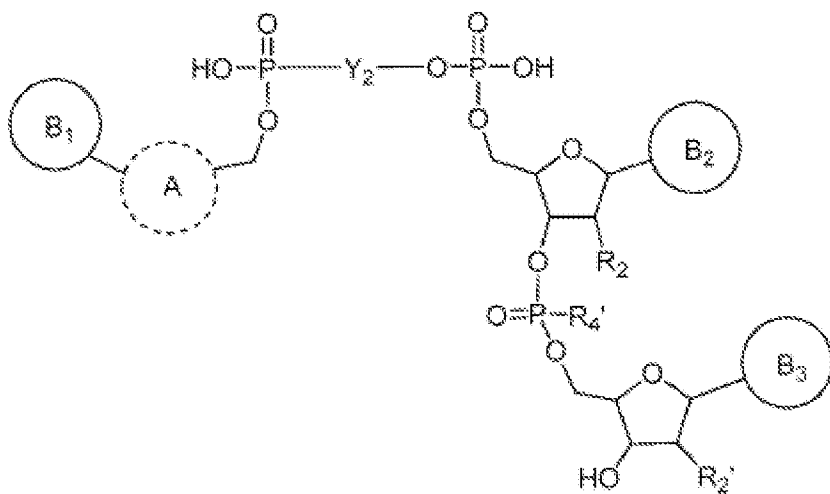
Also provided herein are exemplary caps including those that can be used in co-transcriptional capping methods for ribonucleic acid (RNA) synthesis, using RNA

polymerase, *e.g.*, wild type RNA polymerase or variants thereof, *e.g.*, such as those variants described herein. In some embodiments, caps can be added when RNA is produced in a “one-pot” reaction, without the need for a separate capping reaction. Thus, the methods, in some embodiments, comprise reacting a polynucleotide template with an RNA polymerase variant, nucleoside triphosphates, and a cap analog under *in vitro* transcription reaction conditions to produce RNA transcript.

As used here the term “cap” includes the inverted G nucleotide and can comprise one or more additional nucleotides 3' of the inverted G nucleotide, *e.g.*, 1, 2, 3, or more nucleotides 3' of the inverted G nucleotide and 5' to the 5' UTR, *e.g.*, a 5' UTR described herein.

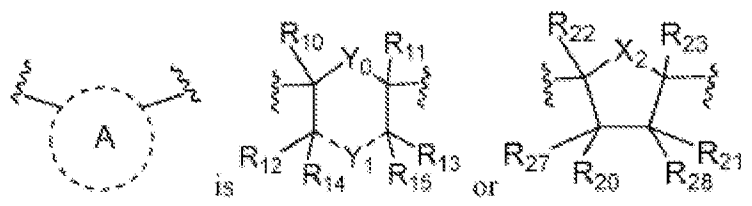
Exemplary caps comprise a sequence of GG, GA, or GGA, wherein the underlined, italicized G is an inverted G nucleotide followed by a 5'-5'-triphosphate group.

In some embodiments, a cap comprises a compound of Formula (C-I)



(C-I), or a

stereoisomer, tautomer or salt thereof, wherein



ring B₁ is a modified or unmodified Guanine;

ring B₂ and ring B₃ each independently is a nucleobase or a modified nucleobase;

X₂ is O, S(O)_p, NR₂₄ or CR₂₅R₂₆ in which p is 0, 1, or 2;

Y₀ is O or CR₆R₇;

Y₁ is O, S(O)_n, CR₆R₇, or NR₈, in which n is 0, 1, or 2;

5 each --- is a single bond or absent, wherein when each --- is a single bond, Y_i is O, S(O)_n, CR₆R₇, or NR₈; and when each --- is absent, Y₁ is void;

Y₂ is (OP(O)R₄)_m in which m is 0, 1, or 2, or -O-(CR₄₀R₄₁)_u-Q₀-(CR₄₂R₄₃)_v-, in which Q₀ is a bond, O, S(O)_r, NR₄₄, or CR₄₅R₄₆, r is 0, 1, or 2, and each of u and v independently is 1, 2, 3 or 4;

10 each R₂ and R₂' independently is halo, LNA, or OR₃;

each R₃ independently is H, C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl and R₃, when being C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl, is optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy that is optionally substituted with one or more OH or OC(O)-C₁-C₆ alkyl;

15 each R₄ and R₄' independently is H, halo, C₁-C₆ alkyl, OH, SH, SeH, or BH₃⁻;

each of R₆, R₇, and R₈, independently, is -Q₁-T₁, in which Q₁ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₁ is H, halo, OH, COOH, cyano, or R_{s1}, in which R_{s1} is C₁-C₃ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₁-C₆ alkoxy, C(O)O-C₁-C₆ alkyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, NR₃₁R₃₂, (NR₃₁R₃₂R₃₃)⁺, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R_{s1} is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, NR₃₁R₃₂, (NR₃₁R₃₂R₃₃)⁺, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl;

25 each of R₁₀, R₁₁, R₁₂, R₁₃, R₁₄, and R₁₅, independently, is -Q₂-T₂, in which Q₂ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₂ is H, halo, OH, NH₂, cyano, NO₂, N₃, R_{s2}, or OR_{s2}, in which R_{s2} is C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, NHC(O)-C₁-C₆ alkyl,

NR₃₁R₃₂, (NR₃₁R₃₂R₃₃)⁺, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R₃₂ is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, NR₃₁R₃₂, (NR₃₁R₃₂R₃₃)⁺, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl; or alternatively R₁₂ together with R₁₄ is oxo, or R₁₃ together with R₁₅ is oxo,

each of R₂₀, R₂₁, R₂₂, and R₂₃ independently is -Q₃-T₃, in which Q₃ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₃ is H, halo, OH, NH₂, cyano, NO₂, N₃, R_{S3}, or OR_{S3}, in which R_{S3} is C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, NHC(O)-C₁-C₆ alkyl, mono-C₁-C₆ alkylamino, di-C₁-C₆ alkylamino, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R_{S3} is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, amino, mono-C₁-C₆ alkylamino, di-C₁-C₆ alkylamino, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl;

each of R₂₄, R₂₅, and R₂₆ independently is H or C₁-C₆ alkyl;

each of R₂₇ and R₂₈ independently is H or OR₂₉; or R₂₇ and R₂₈ together form O-R₃₀-O; each R₂₉ independently is H, C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl and R₂₉, when being C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl, is optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy that is optionally substituted with one or more OH or OC(O)-C₁-C₆ alkyl;

R₃₀ is C₁-C₆ alkylene optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy;

each of R₃₁, R₃₂, and R₃₃, independently is H, C₁-C₆ alkyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl;

each of R₄₀, R₄₁, R₄₂, and R₄₃ independently is H, halo, OH, cyano, N₃, OP(O)R₄₇R₄₈, or C₁-C₆ alkyl optionally substituted with one or more OP(O)R₄₇R₄₈, or one R₄₁ and one R₄₃, together with the carbon atoms to which they are attached and Q₀, form C₄-C₁₀ cycloalkyl, 4- to 14-membered heterocycloalkyl, C₆-C₁₀ aryl, or 5- to 14-membered heteroaryl, and each of the cycloalkyl, heterocycloalkyl, phenyl, or 5- to 6-membered heteroaryl is optionally

substituted with one or more of OH, halo, cyano, N₃, oxo, OP(O)R₄₇R₄₈, C₁-C₆ alkyl, C₁-C₆ haloalkyl, COOH, C(O)O-C₁-C₆ alkyl, C₁-C₆ alkoxy, C₁-C₆ haloalkoxy, amino, mono-C₁-C₆ alkylamino, and di-C₁-C₆ alkylamino;

R₄₄ is H, C₁-C₆ alkyl, or an amine protecting group;

5 each of R₄₅ and R₄₆ independently is H, OP(O)R₄₇R₄₈, or C₁-C₆ alkyl optionally substituted with one or more OP(O)R₄₇R₄₈, and

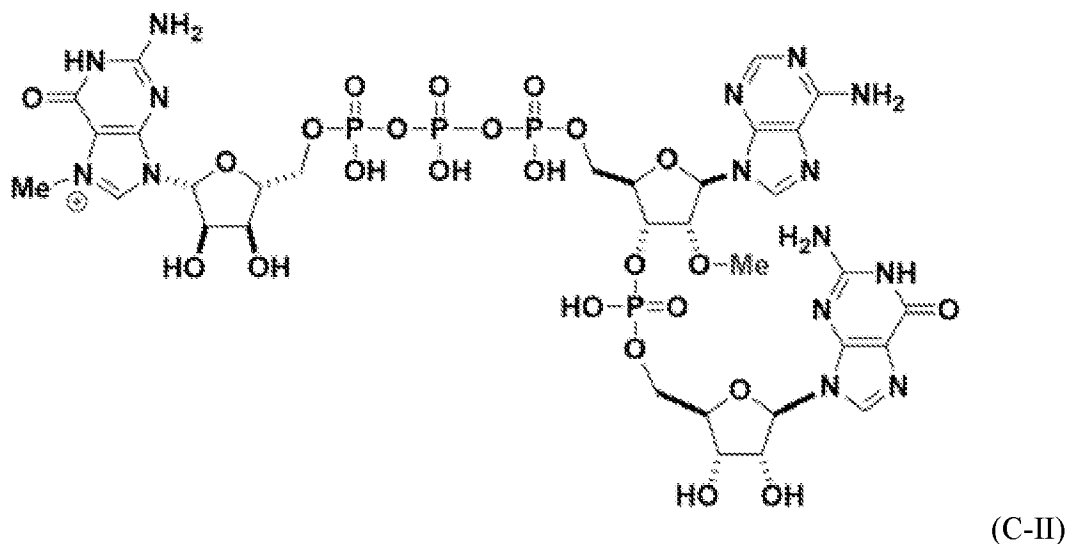
each of R₄₇ and R₄₈, independently is H, halo, C₁-C₆ alkyl, OH, SH, SeH, or BH₃⁻.

It should be understood that a cap analog, as provided herein, may include any of the cap analogs described in international publication WO 2017/066797, published on 20 April
10 2017, incorporated by reference herein in its entirety.

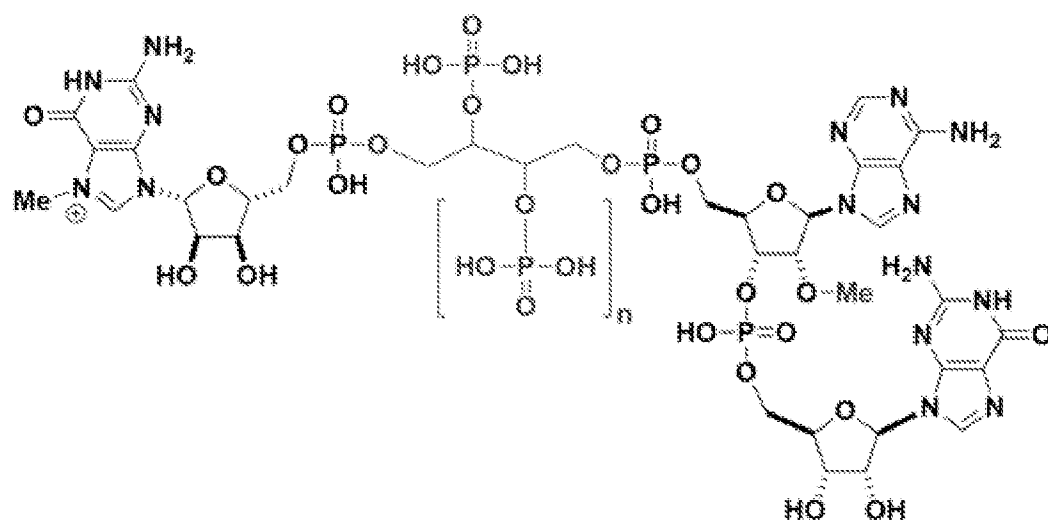
In some embodiments, the B₂ middle position can be a non-ribose molecule, such as arabinose.

In some embodiments R₂ is ethyl-based.

Thus, in some embodiments, a cap comprises the following structure:

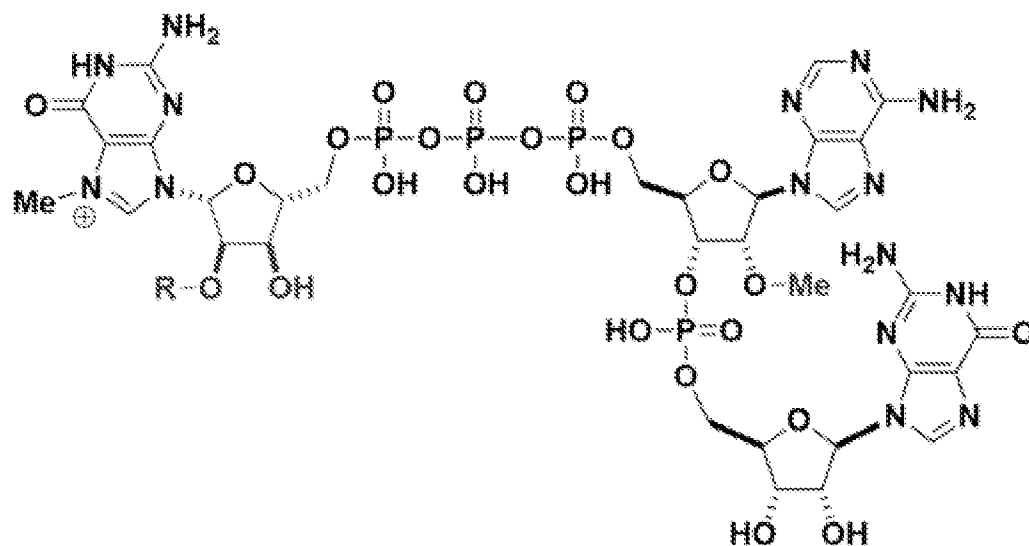


In other embodiments, a cap comprises the following structure:



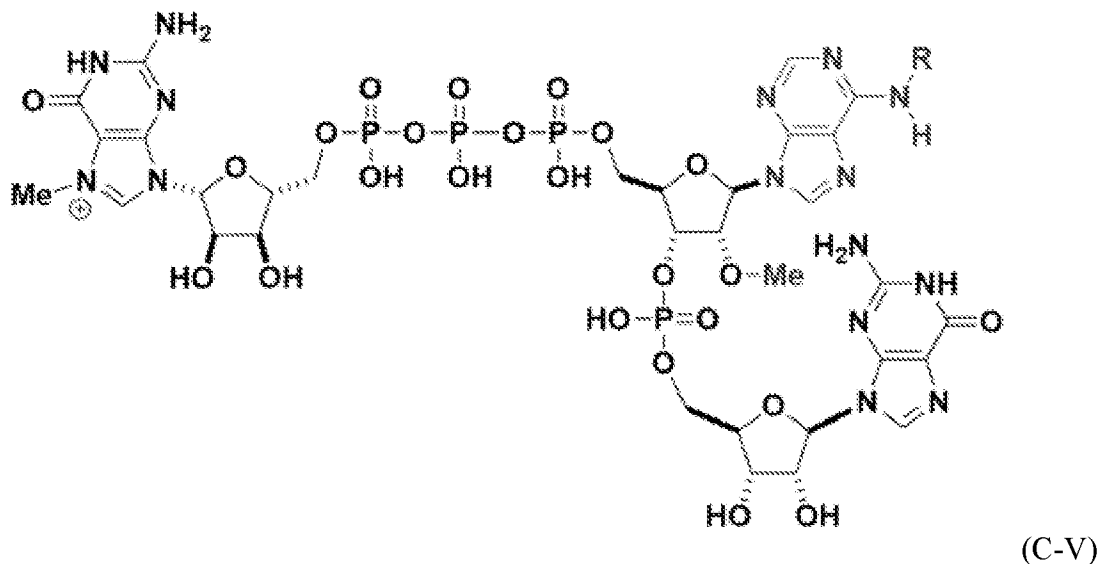
(C-III)

In yet other embodiments, a cap comprises the following structure:



(C-IV)

In still other embodiments, a cap comprises the following structure:



In some embodiments, R is an alkyl (*e.g.*, C₁-C₆ alkyl). In some embodiments, R is a methyl group (*e.g.*, C₁ alkyl). In some embodiments, R is an ethyl group (*e.g.*, C₂ alkyl).

- 5 In some embodiments, a cap comprises a sequence selected from the following sequences: GAA, GAC, GAG, GAU, GCA, GCC, GCG, GCU, GGA, GGC, GGG, GGU, GUA, GUC, GUG, and GUU. In some embodiments, a cap comprises GAA. In some embodiments, a cap comprises GAC. In some embodiments, a cap comprises GAG. In some embodiments, a cap comprises GAU. In some embodiments, a cap comprises GCA. In some
- 10 embodiments, a cap comprises GCC. In some embodiments, a cap comprises GCG. In some embodiments, a cap comprises GCU. In some embodiments, a cap comprises GGA. In some embodiments, a cap comprises GGC. In some embodiments, a cap comprises GGG. In some embodiments, a cap comprises GGU. In some embodiments, a cap comprises GUA. In some embodiments, a cap comprises GUC. In some embodiments, a cap comprises GUG. In some
- 15 embodiments, a cap comprises GUU.

In some embodiments, a cap comprises a sequence selected from the following sequences: m⁷GpppApA, m⁷GpppApC, m⁷GpppApG, m⁷GpppApU, m⁷GpppCpA, m⁷GpppCpC, m⁷GpppCpG, m⁷GpppCpU, m⁷GpppGpA, m⁷GpppGpC, m⁷GpppGpG, m⁷GpppGpU, m⁷GpppUpA, m⁷GpppUpC, m⁷GpppUpG, and m⁷GpppUpU.

- 20 In some embodiments, a cap comprises m⁷GpppApA. In some embodiments, a cap comprises m⁷GpppApC. In some embodiments, a cap comprises m⁷GpppApG. In some embodiments, a cap comprises m⁷GpppApU. In some embodiments, a cap comprises

$m^7G_{ppp}CpA$. In some embodiments, a cap comprises $m^7G_{ppp}CpC$. In some embodiments, a cap comprises $m^7G_{ppp}CpG$. In some embodiments, a cap comprises $m^7G_{ppp}CpU$. In some embodiments, a cap comprises $m^7G_{ppp}GpA$. In some embodiments, a cap comprises $m^7G_{ppp}GpC$. In some embodiments, a cap comprises $m^7G_{ppp}GpG$. In some embodiments, a cap comprises $m^7G_{ppp}GpU$. In some embodiments, a cap comprises $m^7G_{ppp}UpA$. In some embodiments, a cap comprises $m^7G_{ppp}UpC$. In some embodiments, a cap comprises $m^7G_{ppp}UpG$. In some embodiments, a cap comprises $m^7G_{ppp}UpU$.

A cap, in some embodiments, comprises a sequence selected from the following sequences: $m^7G_{3'OMe}pppApA$, $m^7G_{3'OMe}pppApC$, $m^7G_{3'OMe}pppApG$, $m^7G_{3'OMe}pppApU$, $m^7G_{3'OMe}pppCpA$, $m^7G_{3'OMe}pppCpC$, $m^7G_{3'OMe}pppCpG$, $m^7G_{3'OMe}pppCpU$, $m^7G_{3'OMe}pppGpA$, $m^7G_{3'OMe}pppGpC$, $m^7G_{3'OMe}pppGpG$, $m^7G_{3'OMe}pppGpU$, $m^7G_{3'OMe}pppUpA$, $m^7G_{3'OMe}pppUpC$, $m^7G_{3'OMe}pppUpG$, and $m^7G_{3'OMe}pppUpU$.

In some embodiments, a cap comprises $m^7G_{3'OMe}pppApA$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppApC$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppApG$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppApU$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppCpA$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppCpC$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppCpG$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppCpU$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppGpA$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppGpC$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppGpG$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppGpU$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppUpA$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppUpC$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppUpG$. In some embodiments, a cap comprises $m^7G_{3'OMe}pppUpU$.

In some embodiments, a cap comprises a sequence selected from the following sequences: $m^7G_{3'OMe}pppA_{2'OMe}pA$, $m^7G_{3'OMe}pppA_{2'OMe}pC$, $m^7G_{3'OMe}pppA_{2'OMe}pG$, $m^7G_{3'OMe}pppA_{2'OMe}pU$, $m^7G_{3'OMe}pppC_{2'OMe}pA$, $m^7G_{3'OMe}pppC_{2'OMe}pC$, $m^7G_{3'OMe}pppC_{2'OMe}pG$, $m^7G_{3'OMe}pppC_{2'OMe}pU$, $m^7G_{3'OMe}pppG_{2'OMe}pA$, $m^7G_{3'OMe}pppG_{2'OMe}pC$, $m^7G_{3'OMe}pppG_{2'OMe}pG$, $m^7G_{3'OMe}pppG_{2'OMe}pU$, $m^7G_{3'OMe}pppU_{2'OMe}pA$, $m^7G_{3'OMe}pppU_{2'OMe}pC$, $m^7G_{3'OMe}pppU_{2'OMe}pG$, and $m^7G_{3'OMe}pppU_{2'OMe}pU$.

In some embodiments, a cap comprises $m^7G_3'OMepppA_2'OMepA$. In some embodiments, a cap comprises $m^7G_3'OMepppA_2'OMepC$. In some embodiments, a cap comprises $m^7G_3'OMepppA_2'OMepG$. In some embodiments, a cap comprises $m^7G_3'OMepppA_2'OMepU$. In some embodiments, a cap comprises $m^7G_3'OMepppC_2'OMepA$. In some embodiments, a cap comprises $m^7G_3'OMepppC_2'OMepC$. In some embodiments, a cap comprises $m^7G_3'OMepppC_2'OMepG$. In some embodiments, a cap comprises $m^7G_3'OMepppC_2'OMepU$. In some embodiments, a cap comprises $m^7G_3'OMepppG_2'OMepA$. In some embodiments, a cap comprises $m^7G_3'OMepppG_2'OMepC$. In some embodiments, a cap comprises $m^7G_3'OMepppG_2'OMepG$. In some embodiments, a cap comprises $m^7G_3'OMepppG_2'OMepU$. In some embodiments, a cap comprises $m^7G_3'OMepppU_2'OMepA$. In some embodiments, a cap comprises $m^7G_3'OMepppU_2'OMepC$. In some embodiments, a cap comprises $m^7G_3'OMepppU_2'OMepG$. In some embodiments, a cap comprises $m^7G_3'OMepppU_2'OMepU$.

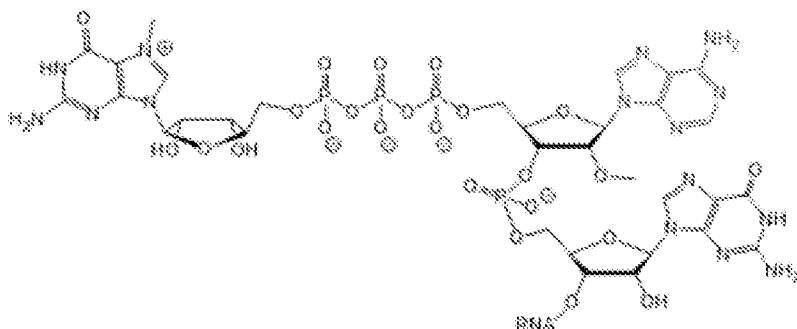
A cap, in still other embodiments, comprises a sequence selected from the following sequences: $m^7GpppA_2'OMepA$, $m^7GpppA_2'OMepC$, $m^7GpppA_2'OMepG$, $m^7GpppA_2'OMepU$, $m^7GpppC_2'OMepA$, $m^7GpppC_2'OMepC$, $m^7GpppC_2'OMepG$, $m^7GpppC_2'OMepU$, $m^7GpppG_2'OMepA$, $m^7GpppG_2'OMepC$, $m^7GpppG_2'OMepG$, $m^7GpppG_2'OMepU$, $m^7GpppU_2'OMepA$, $m^7GpppU_2'OMepC$, $m^7GpppU_2'OMepG$, and $m^7GpppU_2'OMepU$.

In some embodiments, a cap comprises $m^7GpppA_2'OMepA$. In some embodiments, a cap comprises $m^7GpppA_2'OMepC$. In some embodiments, a cap comprises $m^7GpppA_2'OMepG$. In some embodiments, a cap comprises $m^7GpppA_2'OMepU$. In some embodiments, a cap comprises $m^7GpppC_2'OMepA$. In some embodiments, a cap comprises $m^7GpppC_2'OMepC$. In some embodiments, a cap comprises $m^7GpppC_2'OMepG$. In some embodiments, a trinucleotide cap comprises $m^7GpppC_2'OMepU$. In some embodiments, a cap comprises $m^7GpppG_2'OMepA$. In some embodiments, a cap comprises $m^7GpppG_2'OMepC$. In some embodiments, a cap comprises $m^7GpppG_2'OMepG$. In some embodiments, a cap comprises $m^7GpppG_2'OMepU$. In some embodiments, a cap comprises $m^7GpppU_2'OMepA$. In some embodiments, a cap comprises $m^7GpppU_2'OMepC$. In some embodiments, a cap comprises $m^7GpppU_2'OMepG$. In some embodiments, a cap comprises $m^7GpppU_2'OMepU$.

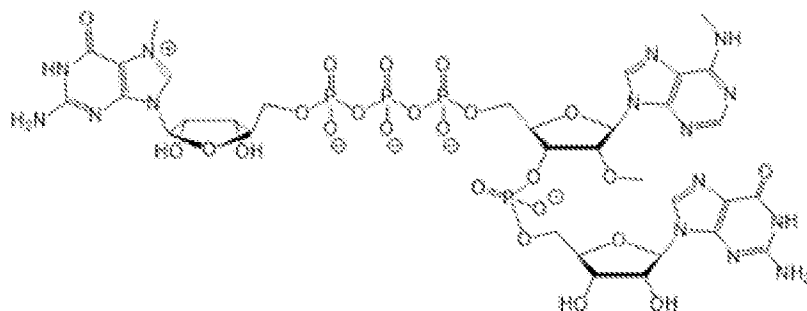
In some embodiments, a cap comprises $m^7Gpppm^6A_2'OMepG$. In some embodiments, a cap comprises $m^7Gpppe^6A_2'OMepG$.

In some embodiments, a cap comprises GAG. In some embodiments, a cap comprises GCG. In some embodiments, a cap comprises GUG. In some embodiments, a cap comprises GGG.

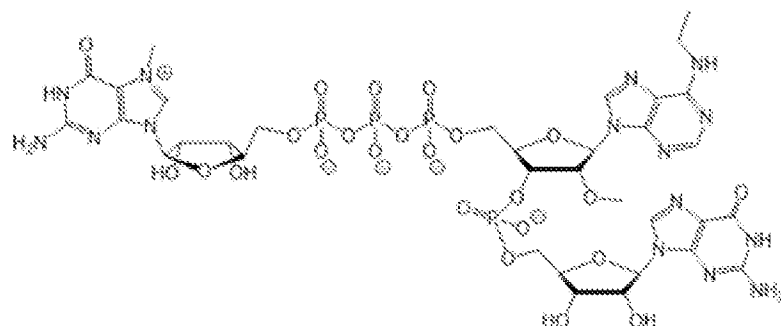
In some embodiments, a cap comprises any one of the following structures:



(C-VI)



(C-VII); or



(C-VIII).

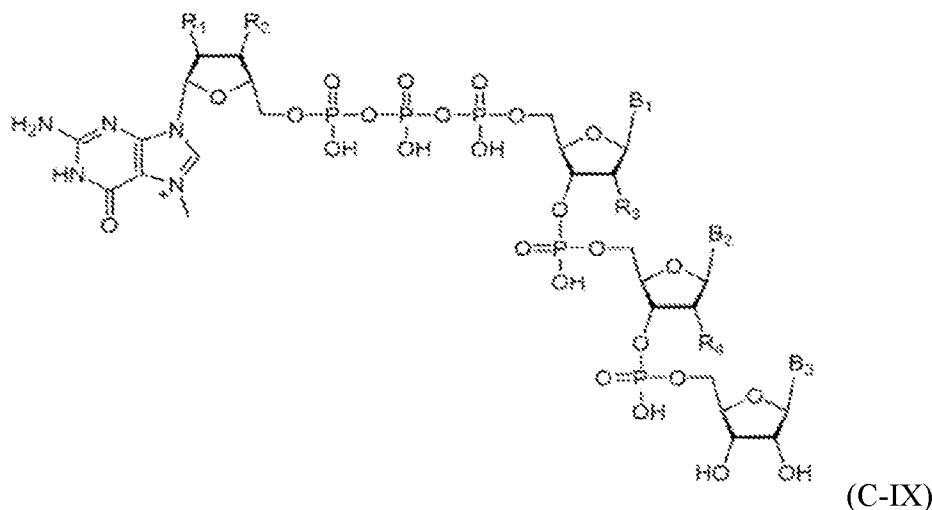
5

In some embodiments, the cap comprises $m^7GpppN_1N_2N_3$, where N_1 , N_2 , and N_3 are optional (*i.e.*, can be absent or one or more can be present) and are independently a natural, a modified, or an unnatural nucleoside base. In some embodiments, m^7G is further methylated, *e.g.*, at the 3' position. In some embodiments, the m^7G comprises an O-methyl at the 3' position. In some embodiments N_1 , N_2 , and N_3 if present, optionally, are independently an adenine, a uracil, a guanine, a thymine, or a cytosine. In some embodiments, one or more

10

(or all) of N₁, N₂, and N₃, if present, are methylated, *e.g.*, at the 2' position. In some embodiments, one or more (or all) of N₁, N₂, and N₃, if present have an O-methyl at the 2' position.

In some embodiments, the cap comprises the following structure:



wherein B₁, B₂, and B₃ are independently a natural, a modified, or an unnatural nucleoside based; and R₁, R₂, R₃, and R₄ are independently OH or O-methyl. In some embodiments, R₃ is O-methyl and R₄ is OH. In some embodiments, R₃ and R₄ are O-methyl. In some embodiments, R₄ is O-methyl. In some embodiments, R₁ is OH, R₂ is OH, R₃ is O-methyl, and R₄ is OH. In some embodiments, R₁ is OH, R₂ is OH, R₃ is O-methyl, and R₄ is O-methyl. In some embodiments, at least one of R₁ and R₂ is O-methyl, R₃ is O-methyl, and R₄ is OH. In some embodiments, at least one of R₁ and R₂ is O-methyl, R₃ is O-methyl, and R₄ is O-methyl.

In some embodiments, B₁, B₂, and B₃ are natural nucleoside bases. In some embodiments, at least one of B₁, B₂, and B₃ is a modified or unnatural base. In some embodiments, at least one of B₁, B₂, and B₃ is N⁶-methyladenine. In some embodiments, B₁ is adenine, cytosine, thymine, or uracil. In some embodiments, B₁ is adenine, B₂ is uracil, and B₃ is adenine. In some embodiments, R₁ and R₂ are OH, R₃ and R₄ are O-methyl, B₁ is adenine, B₂ is uracil, and B₃ is adenine.

In some embodiments the cap comprises a sequence selected from the following sequences: GAAA, GACA, GAGA, GAUA, GCAA, GCCA, GCGA, GCUA, GGAA, GGCA, GGGA, GGUA, GUCA, and GUUA. In some embodiments the cap comprises a

sequence selected from the following sequences: GAAG, GACG, GAGG, GAUG, GCAG, GCCG, GCGG, GCUG, GGAG, GGCG, GGGG, GGUG, GUCG, GUGG, and GUUG. In some embodiments the cap comprises a sequence selected from the following sequences: GAAU, GACU, GAGU, GAUU, GCAU, GCCU, GCGU, GCUU, GGAU, GGCU, GGGU, 5 GGUU, GUAU, GUCU, GUGU, and GUUU. In some embodiments the cap comprises a sequence selected from the following sequences: GAAC, GACC, GAGC, GAUC, GCAC, GCCC, GCGC, GCUC, GGAC, GGCC, GGGC, GGUC, GUAC, GUCC, GUGC, and GUUC.

A cap, in some embodiments, comprises a sequence selected from the following 10 sequences: $m^7G_3'OMepppApApN$, $m^7G_3'OMepppApCpN$, $m^7G_3'OMepppApGpN$, $m^7G_3'OMepppApUpN$, $m^7G_3'OMepppCpApN$, $m^7G_3'OMepppCpCpN$, $m^7G_3'OMepppCpGpN$, $m^7G_3'OMepppCpUpN$, $m^7G_3'OMepppGpApN$, $m^7G_3'OMepppGpCpN$, $m^7G_3'OMepppGpGpN$, $m^7G_3'OMepppGpUpN$, $m^7G_3'OMepppUpApN$, $m^7G_3'OMepppUpCpN$, $m^7G_3'OMepppUpGpN$, and $m^7G_3'OMepppUpUpN$, where N is a natural, a modified, or an unnatural nucleoside base.

15 A cap, in some embodiments, comprises a sequence selected from the following sequences: $m^7G_3'OMepppA_2'OMepApN$, $m^7G_3'OMepppA_2'OMepCpN$, $m^7G_3'OMepppA_2'OMepGpN$, $m^7G_3'OMepppA_2'OMepUpN$, $m^7G_3'OMepppC_2'OMepApN$, $m^7G_3'OMepppC_2'OMepCpN$, $m^7G_3'OMepppC_2'OMepGpN$, $m^7G_3'OMepppC_2'OMepUpN$, $m^7G_3'OMepppG_2'OMepApN$, $m^7G_3'OMepppG_2'OMepCpN$, $m^7G_3'OMepppG_2'OMepGpN$, $m^7G_3'OMepppG_2'OMepUpN$, 20 $m^7G_3'OMepppU_2'OMepApN$, $m^7G_3'OMepppU_2'OMepCpN$, $m^7G_3'OMepppU_2'OMepGpN$, and $m^7G_3'OMepppU_2'OMepUpN$, where N is a natural, a modified, or an unnatural nucleoside base.

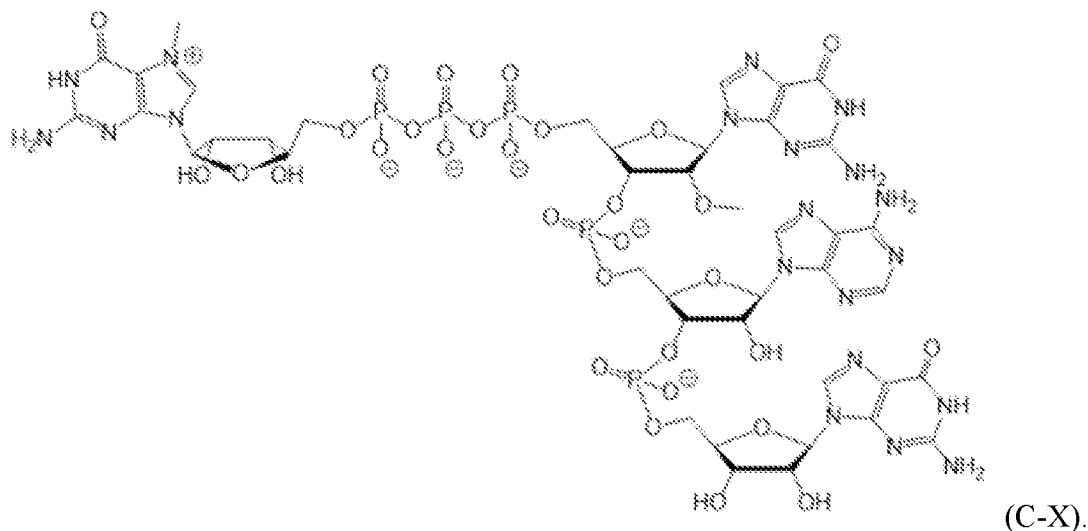
A cap, in some embodiments, comprises a sequence selected from the following sequences: $m^7GpppA_2'OMepApN$, $m^7GpppA_2'OMepCpN$, $m^7GpppA_2'OMepGpN$, $m^7GpppA_2'OMepUpN$, $m^7GpppC_2'OMepApN$, $m^7GpppC_2'OMepCpN$, $m^7GpppC_2'OMepGpN$, 25 $m^7GpppC_2'OMepUpN$, $m^7GpppG_2'OMepApN$, $m^7GpppG_2'OMepCpN$, $m^7GpppG_2'OMepGpN$, $m^7GpppG_2'OMepUpN$, $m^7GpppU_2'OMepApN$, $m^7GpppU_2'OMepCpN$, $m^7GpppU_2'OMepGpN$, and $m^7GpppU_2'OMepUpN$, where N is a natural, a modified, or an unnatural nucleoside base.

A cap, in some embodiments, comprises a sequence selected from the following sequences: $m^7G_3'OMepppA_2'OMepA_2'OMepN$, $m^7G_3'OMepppA_2'OMepC_2'OMepN$, 30 $m^7G_3'OMepppA_2'OMepG_2'OMepN$, $m^7G_3'OMepppA_2'OMepU_2'OMepN$, $m^7G_3'OMepppC_2'OMepA_2'OMepN$,

$m^7G_{3'}OMepppC_{2'}OMepC_{2'}OMepN$, $m^7G_{3'}OMepppC_{2'}OMepG_{2'}OMepN$, $m^7G_{3'}OMepppC_{2'}OMepU_{2'}OMepN$,
 $m^7G_{3'}OMepppG_{2'}OMepA_{2'}OMepN$, $m^7G_{3'}OMepppG_{2'}OMepC_{2'}OMepN$, $m^7G_{3'}OMepppG_{2'}OMepG_{2'}OMepN$,
 $m^7G_{3'}OMepppG_{2'}OMepU_{2'}OMepN$, $m^7G_{3'}OMepppU_{2'}OMepA_{2'}OMepN$, $m^7G_{3'}OMepppU_{2'}OMepC_{2'}OMepN$,
 $m^7G_{3'}OMepppU_{2'}OMepG_{2'}OMepN$, and $m^7G_{3'}OMepppU_{2'}OMepU_{2'}OMepN$, where N is a natural, a
5 modified, or an unnatural nucleoside base.

A cap, in some embodiments, comprises a sequence selected from the following
sequences: $m^7GpppA_{2'}OMepA_{2'}OMepN$, $m^7GpppA_{2'}OMepC_{2'}OMepN$, $m^7GpppA_{2'}OMepG_{2'}OMepN$,
 $m^7GpppA_{2'}OMepU_{2'}OMepN$, $m^7GpppC_{2'}OMepA_{2'}OMepN$, $m^7GpppC_{2'}OMepC_{2'}OMepN$,
 $m^7GpppC_{2'}OMepG_{2'}OMepN$, $m^7GpppC_{2'}OMepU_{2'}OMepN$, $m^7GpppG_{2'}OMepA_{2'}OMepN$,
10 $m^7GpppG_{2'}OMepC_{2'}OMepN$, $m^7GpppG_{2'}OMepG_{2'}OMepN$, $m^7GpppG_{2'}OMepU_{2'}OMepN$,
 $m^7GpppU_{2'}OMepA_{2'}OMepN$, $m^7GpppU_{2'}OMepC_{2'}OMepN$, $m^7GpppU_{2'}OMepG_{2'}OMepN$, and
 $m^7GpppU_{2'}OMepU_{2'}OMepN$, where N is a natural, a modified, or an unnatural nucleoside base.

In some embodiments, a cap comprises GGAG. In some embodiments, a cap
comprises the following structure:



13. Poly-A Tails

In some embodiments, the polynucleotides of the present disclosure (*e.g.*, a
polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide) further
20 comprise a poly-A tail. In some embodiments, terminal groups on the poly-A tail can be
incorporated for stabilization. In some embodiments, a poly-A tail comprises des-3' hydroxyl
tails.

During RNA processing, a long chain of adenine nucleotides (poly-A tail) can be added to a polynucleotide (*e.g.*, an mRNA molecule) in order to increase stability. Immediately after transcription, the 3' end of the transcript can be cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA. The process, called polyadenylation, adds a poly-A tail that can be between, for example, approximately 80 to approximately 250 residues long, including approximately 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240 or 250 residues long. In some embodiments, the poly-A tail is 100 nucleotides in length (SEQ ID NO:195).

PolyA tails can also be added after the construct is exported from the nucleus.

According to the present disclosure, terminal groups on the poly A tail can be incorporated for stabilization. Polynucleotides of the present disclosure can include des-3' hydroxyl tails. They can also include structural moieties or 2'-O-methyl modifications as taught by Junjie Li, et al. (Current Biology, vol. 15, 1501–1507, August 23, 2005), the contents of which are incorporated herein by reference in its entirety).

The polynucleotides of the present disclosure can be designed to encode transcripts with alternative polyA tail structures including histone mRNA. According to Norbury, "Terminal uridylation has also been detected on human replication-dependent histone mRNAs. The turnover of these mRNAs is thought to be important for the prevention of potentially toxic histone accumulation following the completion or inhibition of chromosomal DNA replication. These mRNAs are distinguished by their lack of a 3' poly(A) tail, the function of which is instead assumed by a stable stem-loop structure and its cognate stem-loop binding protein (SLBP); the latter carries out the same functions as those of PABP on polyadenylated mRNAs" (Norbury, "Cytoplasmic RNA: a case of the tail wagging the dog," Nature Reviews Molecular Cell Biology; AOP, published online 29 August 2013; doi:10.1038/nrm3645), the contents of which are incorporated herein by reference in its entirety.

Unique poly-A tail lengths provide certain advantages to the polynucleotides of the present disclosure. Generally, the length of a poly-A tail, when present, is greater than 30 nucleotides in length. In some embodiments, the poly-A tail is greater than 35 nucleotides in length (*e.g.*, at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140,

160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000 nucleotides).

In some embodiments, the polynucleotide or region thereof includes from about 30 to about 3,000 nucleotides (*e.g.*, from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 750, from 30 to 1,000, from 30 to 1,500, from 30 to 2,000, from 30 to 2,500, from 50 to 100, from 50 to 250, from 50 to 500, from 50 to 750, from 50 to 1,000, from 50 to 1,500, from 50 to 2,000, from 50 to 2,500, from 50 to 3,000, from 100 to 500, from 100 to 750, from 100 to 1,000, from 100 to 1,500, from 100 to 2,000, from 100 to 2,500, from 100 to 3,000, from 500 to 750, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 2,500, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 2,500, from 1,000 to 3,000, from 1,500 to 2,000, from 1,500 to 2,500, from 1,500 to 3,000, from 2,000 to 3,000, from 2,000 to 2,500, and from 2,500 to 3,000).

In some embodiments, the poly-A tail is designed relative to the length of the overall polynucleotide or the length of a particular region of the polynucleotide. This design can be based on the length of a coding region, the length of a particular feature or region or based on the length of the ultimate product expressed from the polynucleotides.

In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% greater in length than the polynucleotide or feature thereof. The poly-A tail can also be designed as a fraction of the polynucleotides to which it belongs. In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, or 90% or more of the total length of the construct, a construct region or the total length of the construct minus the poly-A tail. Further, engineered binding sites and conjugation of polynucleotides for Poly-A binding protein can enhance expression.

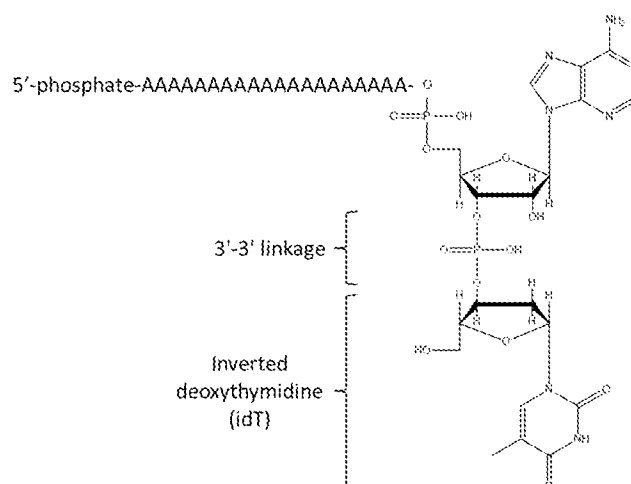
Additionally, multiple distinct polynucleotides can be linked together via the PABP (Poly-A binding protein) through the 3'-end using modified nucleotides at the 3'-terminus of the poly-A tail. Transfection experiments can be conducted in relevant cell lines at and protein production can be assayed by ELISA at 12hr, 24hr, 48hr, 72hr and day 7 post-transfection.

In some embodiments, the polynucleotides of the present disclosure are designed to include a polyA-G Quartet region. The G-quartet is a cyclic hydrogen bonded array of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. In some

emodiments, the G-quartet is incorporated at the end of the poly-A tail. The resultant polynucleotide is assayed for stability, protein production and other parameters including half-life at various time points. It has been discovered that the polyA-G quartet results in protein production from an mRNA equivalent to at least 75% of that seen using a poly-A tail of 120 nucleotides alone (SEQ ID NO:196).

In some embodiments, the polyA tail comprises an alternative nucleoside, *e.g.*, inverted thymidine. PolyA tails comprising an alternative nucleoside, *e.g.*, inverted thymidine, may be generated as described herein. For instance, mRNA constructs may be modified by ligation to stabilize the poly(A) tail. Ligation may be performed using 0.5-1.5 mg/mL mRNA (5' Cap1, 3' A100), 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM TCEP, 1000 units/mL T4 RNA Ligase 1, 1 mM ATP, 20% w/v polyethylene glycol 8000, and 5:1 molar ratio of modifying oligo to mRNA. Modifying oligo has a sequence of 5'-phosphate-AAAAAAAAAAAAAAAAAAAAA-(inverted deoxythymidine (idT) (SEQ ID NO:209)) (see below). Ligation reactions are mixed and incubated at room temperature (~22°C) for, *e.g.*, 4 hours. Stable tail mRNA are purified by, *e.g.*, dT purification, reverse phase purification, hydroxyapatite purification, ultrafiltration into water, and sterile filtration. The resulting stable tail-containing mRNAs contain the following structure at the 3' end, starting with the polyA region: A₁₀₀-UCUAGAAAAAAAAAAAAAAAAAAAAA-inverted deoxythymidine (SEQ ID NO:211).

Modifying oligo to stabilize tail (5'-phosphate-AAAAAAAAAAAAAAAAAAAAA-(inverted deoxythymidine)(SEQ ID NO:209)):



In some instances, the polyA tail comprises A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211). In some instances, the polyA tail consists of A100-UCUAG-A20-inverted deoxy-thymidine (SEQ ID NO:211).

14. Start codon region

The invention also includes a polynucleotide that comprises both a start codon region and the polynucleotide described herein (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide). In some embodiments, the polynucleotides of the present disclosure can have regions that are analogous to or function like a start codon region.

In some embodiments, the translation of a polynucleotide can initiate on a codon that is not the start codon AUG. Translation of the polynucleotide can initiate on an alternative start codon such as, but not limited to, ACG, AGG, AAG, CTG/CUG, GTG/GUG, ATA/AUA, ATT/AUU, TTG/UUG (see Touriol et al. *Biology of the Cell* 95 (2003) 169-178 and Matsuda and Mauro *PLoS ONE*, 2010 5:11; the contents of each of which are herein incorporated by reference in its entirety).

As a non-limiting example, the translation of a polynucleotide begins on the alternative start codon ACG. As another non-limiting example, polynucleotide translation begins on the alternative start codon CTG or CUG. As yet another non-limiting example, the translation of a polynucleotide begins on the alternative start codon GTG or GUG.

Nucleotides flanking a codon that initiates translation such as, but not limited to, a start codon or an alternative start codon, are known to affect the translation efficiency, the length and/or the structure of the polynucleotide. (*See, e.g.*, Matsuda and Mauro *PLoS ONE*, 2010 5:11; the contents of which are herein incorporated by reference in its entirety). Masking any of the nucleotides flanking a codon that initiates translation can be used to alter the position of translation initiation, translation efficiency, length and/or structure of a polynucleotide.

In some embodiments, a masking agent can be used near the start codon or alternative start codon in order to mask or hide the codon to reduce the probability of translation initiation at the masked start codon or alternative start codon. Non-limiting examples of masking agents include antisense locked nucleic acids (LNA) polynucleotides and exon-

junction complexes (EJCs) (*See, e.g.*, Matsuda and Mauro describing masking agents LNA polynucleotides and EJCs (PLoS ONE, 2010 5:11); the contents of which are herein incorporated by reference in its entirety).

In some embodiments, a masking agent can be used to mask a start codon of a polynucleotide in order to increase the likelihood that translation will initiate on an alternative start codon. In some embodiments, a masking agent can be used to mask a first start codon or alternative start codon in order to increase the chance that translation will initiate on a start codon or alternative start codon downstream to the masked start codon or alternative start codon.

In some embodiments, a start codon or alternative start codon can be located within a perfect complement for a miRNA binding site. The perfect complement of a miRNA binding site can help control the translation, length and/or structure of the polynucleotide similar to a masking agent. As a non-limiting example, the start codon or alternative start codon can be located in the middle of a perfect complement for a miRNA binding site. The start codon or alternative start codon can be located after the first nucleotide, second nucleotide, third nucleotide, fourth nucleotide, fifth nucleotide, sixth nucleotide, seventh nucleotide, eighth nucleotide, ninth nucleotide, tenth nucleotide, eleventh nucleotide, twelfth nucleotide, thirteenth nucleotide, fourteenth nucleotide, fifteenth nucleotide, sixteenth nucleotide, seventeenth nucleotide, eighteenth nucleotide, nineteenth nucleotide, twentieth nucleotide or twenty-first nucleotide.

In some embodiments, the start codon of a polynucleotide can be removed from the polynucleotide sequence in order to have the translation of the polynucleotide begin on a codon that is not the start codon. Translation of the polynucleotide can begin on the codon following the removed start codon or on a downstream start codon or an alternative start codon. In a non-limiting example, the start codon ATG or AUG is removed as the first 3 nucleotides of the polynucleotide sequence in order to have translation initiate on a downstream start codon or alternative start codon. The polynucleotide sequence where the start codon was removed can further comprise at least one masking agent for the downstream start codon and/or alternative start codons in order to control or attempt to control the initiation of translation, the length of the polynucleotide and/or the structure of the polynucleotide.

15. Stop Codon Region

The present disclosure also includes a polynucleotide that comprises both a stop codon region and the polynucleotide described herein (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide). In some embodiments, the polynucleotides of the present disclosure can include at least two stop codons before the 3' untranslated region (UTR). The stop codon can be selected from TGA, TAA and TAG in the case of DNA, or from UGA, UAA and UAG in the case of RNA. In some embodiments, the polynucleotides of the present disclosure include the stop codon TGA in the case of DNA, or the stop codon UGA in the case of RNA, and one additional stop codon. In a further embodiment the additional stop codon can be TAA or UAA. In some embodiments, the polynucleotides of the present disclosure include three consecutive stop codons, four stop codons, or more.

16. Identification and Ratio Determination (IDR) Sequences

An Identification and Ratio Determination (IDR) sequence is a sequence of a biological molecule (*e.g.*, nucleic acid or protein) that, when combined with the sequence of a target biological molecule, serves to identify the target biological molecule. Typically, an IDR sequence is a heterologous sequence that is incorporated within or appended to a sequence of a target biological molecule and can be used as a reference to identify the target molecule. Thus, in some embodiments, a nucleic acid (*e.g.*, mRNA) comprises (i) a target sequence of interest (*e.g.*, a coding sequence encoding a therapeutic and/or antigenic peptide or protein); and (ii) a unique IDR sequence.

An RNA species (*e.g.*, RNA having a given coding sequence) may comprise an IDR sequence that differs from the IDR sequence of other RNA species (*e.g.*, RNA(s) having different coding sequence(s)). Each IDR sequence thus identifies a particular RNA species, and so the abundance of IDR sequences may be measured to determine the abundance of each RNA species in a composition. Use of distinct IDR sequences to identify RNA species allows for analysis of multivalent RNA compositions (*e.g.*, containing multiple RNA species) containing RNA species with similar coding sequences and/or lengths, which could otherwise be difficult to distinguish using PCR- or chromatography-based analysis of full-length RNAs.

Each RNA species in a multivalent RNA composition may comprise an IDR sequence that is not a sequence isomer of an IDR sequence of another RNA species in a multivalent RNA composition (*e.g.*, the IDR sequence does not have the same number of adenosine

nucleotides, the same number of cytosine nucleotides, the same number of guanine nucleotides, and the same number of uracil nucleotides, as another IDR sequence in the composition, even if those sequences have different sequences). Having identical nucleotide compositions causes sequence isomers to have the same mass, presenting a challenge to distinguishing sequence isomers using mass-based identification methods (*e.g.*, mass spectrometry).

Each RNA species in a multivalent RNA composition may comprise an IDR sequence having a mass that differs from the mass of IDR sequences of each other RNA species in a multivalent RNA composition. For example, the mass of each IDR sequence may differ from the mass of other IDR sequences by at least 9 Da, at least 25 Da, at least 25 Da, or at least 50 Da. Use of IDR sequences with distinct masses allows RNA fragments comprising different IDR sequences to be distinguished using mass-based analysis methods (*e.g.*, mass spectrometry), which do not require reverse transcription, amplification, or sequencing of RNAs.

Each RNA species in an RNA composition may comprises an IDR sequence with a different length. For example, each IDR sequence may have a length independently selected from 0 to 25 nucleotides. The length of a nucleic acid influences the rate at which the nucleic acid traverses a chromatography column, and so the use of IDR sequences of different lengths on different RNA species allows RNA fragments having different IDR sequences to be distinguished using chromatography-based methods (*e.g.*, LC-UV).

IDR sequences may be chosen such that no IDR sequence comprises a start codon, 'AUG'. Lack of a start codon in an IDR sequence prevents undesired translation of nucleotide sequences within and/or downstream from the IDR sequence.

IDR sequences may be chosen such that no IDR sequence comprises a recognition site for a restriction enzyme. In one example, no IDR sequence comprises a recognition site for XbaI, 'UCUAG'. Lack of a recognition site for a restriction enzyme (*e.g.*, XbaI recognition site 'UCUAG') allows the restriction enzyme to be used in generating and modifying a DNA template for *in vitro* transcription, without affecting the IDR sequence or sequence of the transcribed RNA.

17. Combination of mRNA elements

Any of the polynucleotides disclosed herein can comprise one, two, three, or all of the following elements: (a) a 5'-UTR, *e.g.*, as described herein; (b) a coding region comprising a

stop element (*e.g.*, as described herein); (c) a 3'-UTR (*e.g.*, as described herein) and; optionally (d) a 3' stabilizing region, *e.g.*, as described herein. Also disclosed herein are LNP compositions comprising the same.

In some embodiments, a polynucleotide of the disclosure comprises (a) a 5' UTR described in **Table 3** or a variant or fragment thereof and (b) a coding region comprising a stop element provided herein. In some embodiments, the polynucleotide further comprises a cap structure, *e.g.*, as described herein, or a poly A tail, *e.g.*, as described herein. In some embodiments, the polynucleotide further comprises a 3' stabilizing region, *e.g.*, as described herein.

In some embodiments, a polynucleotide of the disclosure comprises (a) a 5' UTR described in **Table 3** or a variant or fragment thereof and (c) a 3' UTR described in **Table 4** or a variant or fragment thereof. In some embodiments, the polynucleotide further comprises a cap structure, *e.g.*, as described herein, or a poly A tail, *e.g.*, as described herein. In some embodiments, the polynucleotide further comprises a 3' stabilizing region, *e.g.*, as described herein.

In some embodiments, a polynucleotide of the disclosure comprises (c) a 3' UTR described in **Table 4** or a variant or fragment thereof and (b) a coding region comprising a stop element provided herein. In some embodiments, the polynucleotide comprises a sequence provided in **Table 6**. In some embodiments, the polynucleotide further comprises a cap structure, *e.g.*, as described herein, or a poly A tail, *e.g.*, as described herein. In some embodiments, the polynucleotide further comprises a 3' stabilizing region, *e.g.*, as described herein.

In some embodiments, a polynucleotide of the disclosure comprises (a) a 5' UTR described in **Table 3** or a variant or fragment thereof; (b) a coding region comprising a stop element provided herein; and (c) a 3' UTR described in **Table 4** or a variant or fragment thereof. In some embodiments, the polynucleotide further comprises a cap structure, *e.g.*, as described herein, or a poly A tail, *e.g.*, as described herein. In some embodiments, the polynucleotide further comprises a 3' stabilizing region, *e.g.*, as described herein.

Table 6: Exemplary 3' UTR and stop element sequences

SEQ ID NO	Sequence information	Sequence
-----------	----------------------	----------

121	3' UTR with stop C11 (underlined)	<u>UAGGGUUAAGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
122	3' UTR with stop C10 (underlined)	<u>UAAAGCUCCGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
123	3' UTR with stop C9 (underlined)	<u>UAAGCCCCUGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
124	3' UTR with stop C8 (underlined)	<u>UAAGCACCCGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
125	3' UTR with stop C7 (underlined)	<u>UAAGCCCCUCCGGGGGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
126	3' UTR with stop C6 (underlined)	<u>UAAGGCUAAGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
127	3' UTR with stop C5 (underlined)	<u>UAAGUCUCCGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
128	3' UTR with stop C4 (underlined)	<u>UAAAGCUAAGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
129	3' UTR with stop C3 (underlined)	<u>UAAGUCUAAGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCC</u> CUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCC GUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC
130	3' UTR with C10 stop (underlined)	<u>UAAAGCUCCAUAAAGUAGGAAACACUACAGCUGGAGCCUC</u> GGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAGUAG GAAACACUACAUCCCCCAGCCCCUCCUCCCCUCCUGCA CCCGUACCCCCUCCAUAAGUAGGAAACACUACAGUGGUC UUUGAAUAAAGUCUGAGUGGGCGGC
131	3' UTR with C7 stop (underlined)	<u>UAAGCCCCUCCGGGGUCCAUAAGUAGGAAACACUACAGC</u> CUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAG UAGGAAACACUACAUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCUCCAUAAGUAGGAAACACUACAGUG GUCUUUGAAUAAAGUCUGAGUGGGCGGC
132	3' UTR with C8 stop (underlined)	<u>UAAAGCUCCCCGGGGUCCAUAAGUAGGAAACACUACAGC</u> CUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCAUAAG UAGGAAACACUACAUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCUCCAUAAGUAGGAAACACUACAGUG GUCUUUGAAUAAAGUCUGAGUGGGCGGC

18. Polynucleotide Comprising an mRNA Encoding an ASL Polypeptide

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

5 (i) a 5' cap such as provided above;

(ii) a 5' UTR, such as the sequences provided above;

(iii) an ORF encoding a human ASL polypeptide having 100% sequence identity to the amino acid sequence according to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11;

10 (iv) at least one stop codon;

(v) a 3' UTR, such as the sequences provided above; and

(vi) a poly-A tail provided above.

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

15 (i) a 5' cap such as provided above;

(ii) a 5' UTR, such as the sequences provided above;

(iii) an ORF encoding a human ASL polypeptide, wherein the ORF has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44;

20 (iv) at least one stop codon;

(v) a 3' UTR, such as the sequences provided above; and

(vi) a poly-A tail provided above.

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

30 (i) a 5' cap such as provided above;

- (ii) a 5' UTR, such as the sequences provided above;
- (iii) an ORF encoding a human ASL polypeptide having an amino acid sequence with 100% identity to the amino acid sequence according to SEQ ID NO:2;
- (iv) at least one stop codon;
- 5 (v) a 3' UTR, such as the sequences provided above; and
- (vi) a poly-A tail provided above.

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises from 5' to 3' end:

- 10 (i) a 5' cap such as provided above;
- (ii) a 5' UTR, such as the sequences provided above;
- (iii) an ORF encoding a human ASL polypeptide, wherein the ORF has at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO:20;
- 15 (iv) at least one stop codon;
- (v) a 3' UTR, such as the sequences provided above; and
- (vi) a poly-A tail provided above.

In some embodiments, the polynucleotide further comprises a miRNA binding site, *e.g.*, a miRNA binding site that binds to miRNA-142. In some embodiments, the 5' UTR comprises the miRNA binding site. In some embodiments, the 3' UTR comprises the miRNA binding site.

In some embodiments, a polynucleotide of the present disclosure comprises a nucleotide sequence encoding a polypeptide sequence at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to the protein sequence of a wild type human ASL (SEQ ID NO:1).

In some embodiments, a polynucleotide of the present disclosure comprises a nucleotide sequence encoding a polypeptide sequence at least 70%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least

95%, at least 96% , at least 97%, at least 98%, at least 99%, or 100% identical to the protein sequence of a human ASL having the amino acid sequence of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11.

5 In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding a polypeptide, comprises (1) a 5' cap such as provided above, for example, m⁷Gp-ppGm-A or Cap 1, (2) a 5' UTR (*e.g.*, SEQ ID NO:50), (3) a nucleotide sequence ORF of SEQ ID NO:20, (3) a stop codon, (4) a 3'UTR (*e.g.*, SEQ ID NO:108), and (5) a poly-A tail provided above, for example, a
10 poly-A tail of, *e.g.*, SEQ ID NO:195 or, *e.g.*, A100-UCUAG-A20-inverted deoxy-thymidine (*e.g.*, SEQ ID NO:211).

In some embodiments, a polynucleotide of the present disclosure, for example a polynucleotide comprising an mRNA nucleotide sequence encoding an ASL polypeptide, comprises (1) a 5' cap such as provided above, for example, m⁷Gp-ppGm-A or Cap 1, (2) a 5'
15 UTR (*e.g.*, SEQ ID NO:56), (3) a nucleotide sequence ORF of SEQ ID NO:21, SEQ ID NO:22, or SEQ ID NO:23, (3) a stop codon, (4) a 3'UTR (*e.g.*, SEQ ID NO:108), and (5) a poly-A tail provided above, for example, a poly-A tail of, *e.g.*, SEQ ID NO:195 or, *e.g.*, A100-UCUAG-A20-inverted deoxy-thymidine (*e.g.*, SEQ ID NO:211).

Exemplary ASL nucleotide constructs are described below:

20 SEQ ID NO:300 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising m⁷Gp-ppGm-AG; 5' UTR of SEQ ID NO:50, ASL nucleotide ORF of SEQ ID NO:20, a 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail. In some embodiments the polyA tail may be a poly-A tail of, *e.g.*, SEQ ID NO:195 or, *e.g.*, A100-UCUAG-A20-inverted deoxy-thymidine (*e.g.*, SEQ ID NO:211). Where an inverted deoxy-
25 thymidine is present, the present disclosure describes the construct using (“+ idT”). For example, a construct according to SEQ ID NO:300 with an idT would be indicated as “SEQ ID NO:300 (+ idT).”

SEQ ID NO:301 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:50, ASL nucleotide ORF of SEQ ID NO:20, and 3'
30 UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:302 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:50, ASL nucleotide ORF of SEQ ID NO:20, and 3' UTR of SEQ ID NO:128, and a 100-nucleotide polyA tail.

5 SEQ ID NO:303 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:50, ASL nucleotide ORF of SEQ ID NO:20, and 3' UTR of SEQ ID NO:138, and a 100-nucleotide polyA tail.

SEQ ID NO:304 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:78, ASL nucleotide ORF of SEQ ID NO:20, and 3' UTR of SEQ ID NO:137, and a 100-nucleotide polyA tail.

10 SEQ ID NO:305 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:78, ASL nucleotide ORF of SEQ ID NO:20, and 3' UTR of SEQ ID NO:139, and a 100-nucleotide polyA tail.

15 SEQ ID NO:306 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:20, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:307 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:21, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

20 SEQ ID NO:308 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:22, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:309 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:23, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

25 SEQ ID NO:310 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:24, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:311 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:25, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

5 SEQ ID NO:312 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:26, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:313 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:27, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

10 SEQ ID NO:314 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:28, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

15 SEQ ID NO:315 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:29, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:330 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:55, ASL nucleotide ORF of SEQ ID NO:30, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

20 SEQ ID NO:331 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:31, and 3' UTR of SEQ ID NO:111, and a 100-nucleotide polyA tail.

SEQ ID NO:332 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:55, ASL nucleotide ORF of SEQ ID NO:32, and 3' UTR of SEQ ID NO:111, and a 100-nucleotide polyA tail.

25 SEQ ID NO:333 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:55, ASL nucleotide ORF of SEQ ID NO:33, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

SEQ ID NO:334 comprises or consists of, from 5' to 3' end: a 5' terminal cap comprising Cap 1, 5' UTR of SEQ ID NO:56, ASL nucleotide ORF of SEQ ID NO:33, and 3' UTR of SEQ ID NO:108, and a 100-nucleotide polyA tail.

5 In some embodiments, any of the polynucleotide constructs discussed above may further comprise, between the ORF and the 3'UTR, one or more stop codons, if not already present at the 5' end of the 3'UTR.

In some embodiments, in constructs denoted as having "G5" chemistry, all uracils therein are replaced by N1 methylpseudouracil. In constructs denoted as having "G6" chemistry, all uracils therein are replaced by 5-methoxyuracil.

TABLE 7. Modified mRNA constructs including ORFs encoding human ASL (wild-type or variants thereof).

Each of SEQ ID NOs:301-334 comprises a Cap 1 5' terminal cap and a 3' terminal PolyA region (100 nt). SEQ ID NO: 300 comprises a m⁷Gp-ppGm-AG 5' terminal cap and a 3' terminal Poly-A region (100 nt).

5 By "G5" is meant that all uracils (U) in the mRNA are replaced by N1-methylpseudouracils.

ASL mRNA construct	5'UTR SEQ ID NO	ASL ORF		3' UTR SEQ ID NO:
		Chemistry	SEQ ID NO	
#1A (SEQ ID NO:300)	50	G5	20	108
#1B (SEQ ID NO:301)	50	G5	20	108
#2 (SEQ ID NO:302)	50	G5	20	128
#3 (SEQ ID NO:303)	50	G5	20	138
#4 (SEQ ID NO:304)	78	G5	20	137
#5 (SEQ ID NO:305)	78	G5	20	139
#6 (SEQ ID NO:306)	56	G5	20	108
#7 (SEQ ID NO:307)	56	G5	21	108
#8 (SEQ ID NO:308)	56	G5	22	108
#9 (SEQ ID NO:309)	56	G5	23	108
#10 (SEQ ID NO:310)	56	G5	24	108
#11 (SEQ ID NO:311)	56	G5	25	108
#12 (SEQ ID NO:312)	56	G5	26	108
#13 (SEQ ID NO:313)	56	G5	27	108
#14 (SEQ ID NO:314)	56	G5	28	108
#15 (SEQ ID NO:315)	56	G5	29	108
#16 (SEQ ID NO:330)	55	G5	30	108
#17 (SEQ ID NO:331)	56	G5	31	111
#18 (SEQ ID NO:332)	55	G5	32	111
#19 (SEQ ID NO:333)	55	G5	33	108
#20 (SEQ ID NO:334)	56	G5	33	108

19. Methods of Making Polynucleotides

The present disclosure also provides methods for making a polynucleotide of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide) or a complement thereof.

In some aspects, a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) disclosed herein, and encoding an ASL polypeptide, can be constructed using *in vitro* transcription (IVT). In other aspects, a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) disclosed herein, and encoding an ASL polypeptide, can be constructed by chemical synthesis using an oligonucleotide synthesizer.

In other aspects, a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) disclosed herein, and encoding an ASL polypeptide is made by using a host cell. In certain aspects, a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) disclosed herein, and encoding an ASL polypeptide is made by one or more combination of the IVT, chemical synthesis, host cell expression, or any other methods known in the art.

Naturally occurring nucleosides, non-naturally occurring nucleosides, or combinations thereof, can totally or partially naturally replace occurring nucleosides present in the candidate nucleotide sequence and can be incorporated into a sequence-optimized nucleotide sequence (*e.g.*, a RNA, *e.g.*, an mRNA) encoding an ASL polypeptide. The resultant polynucleotides, *e.g.*, mRNAs, can then be examined for their ability to produce protein and/or produce a therapeutic outcome.

a. In Vitro Transcription / Enzymatic Synthesis

The present disclosure also provides methods for making a polynucleotide disclosed herein or a complement thereof. In some aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein can be constructed using *in vitro* transcription.

In other aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein can be constructed by chemical synthesis using an oligonucleotide synthesizer. In other aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein is made by using a host cell. In certain aspects, a polynucleotide (*e.g.*, an mRNA) disclosed herein is made by one or more combination of the IVT, chemical synthesis, host cell expression, or any other methods known in the art.

Naturally occurring nucleosides, non-naturally occurring nucleosides, or combinations thereof, can totally or partially naturally replace occurring nucleosides present in the candidate nucleotide sequence and can be incorporated into a sequence-optimized nucleotide sequence (*e.g.*, an mRNA) encoding an ASL polypeptide. The resultant mRNAs can then be examined for their ability to produce ASL and/or produce a therapeutic outcome.

While RNA can be made synthetically using methods well known in the art, In some embodiments an RNA transcript (*e.g.*, mRNA transcript) is synthesized by contacting a DNA template with a RNA polymerase (*e.g.*, a T7 RNA polymerase or a T7 RNA polymerase variant) under conditions that result in the production of RNA transcript.

In some aspects, the present disclosure provides methods of performing an IVT (*in vitro* transcription) reaction, comprising contacting a DNA template with the RNA polymerase (*e.g.*, a T7 RNA polymerase, such as a T7 RNA polymerase variant) in the presence of nucleoside triphosphates and buffer under conditions that result in the production of RNA transcripts.

Other aspects of the present disclosure provide capping methods, *e.g.*, co-transcriptional capping methods or other methods known in the art. In some embodiments, a capping method comprises reacting a polynucleotide template with a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under *in vitro* transcription reaction conditions to produce RNA transcript.

IVT conditions typically require a purified linear DNA template containing a promoter, nucleoside triphosphates, a buffer system that includes dithiothreitol (DTT) and magnesium ions, and a RNA polymerase. The exact conditions used in the transcription reaction depend on the amount of RNA needed for a specific application. Typical IVT reactions are performed by incubating a DNA template with a RNA polymerase and nucleoside triphosphates, including GTP, ATP, CTP, and UTP (or nucleotide analogs) in a transcription buffer. A RNA transcript having a 5' terminal guanosine triphosphate is produced from this reaction.

A deoxyribonucleic acid (DNA) is simply a nucleic acid template for RNA polymerase. A DNA template may include a polynucleotide encoding an ASL polypeptide. A DNA template, in some embodiments, includes a RNA polymerase promoter (*e.g.*, a T7 RNA polymerase promoter) located 5' from and operably linked to polynucleotide encoding an ASL polypeptide. A DNA template may also include a nucleotide sequence encoding a polyadenylation (polyA) tail located at the 3' end of the gene of interest.

Polypeptides of interest include, but are not limited to, biologics, antibodies, antigens (vaccines), and therapeutic proteins. The term “protein” encompasses peptides.

A RNA transcript, in some embodiments, is the product of an IVT reaction and, as will be understood by one of ordinary skill in the art, the DNA template for making an RNA molecule is known based on base complementarity. A RNA transcript, in some embodiments, is a messenger RNA (mRNA) that includes a nucleotide sequence encoding a polypeptide of

interest linked to a polyA tail. In some embodiments, the mRNA is modified mRNA (mmRNA), which includes at least one modified nucleotide.

A nucleotide includes a nitrogenous base, a five-carbon sugar (ribose or deoxyribose), and at least one phosphate group. Nucleotides include nucleoside monophosphates, nucleoside diphosphates, and nucleoside triphosphates. A nucleoside monophosphate (NMP) includes a nucleobase linked to a ribose and a single phosphate; a nucleoside diphosphate (NDP) includes a nucleobase linked to a ribose and two phosphates; and a nucleoside triphosphate (NTP) includes a nucleobase linked to a ribose and three phosphates. Nucleotide analogs are compounds that have the general structure of a nucleotide or are structurally similar to a nucleotide. Nucleotide analogs, for example, include an analog of the nucleobase, an analog of the sugar and/or an analog of the phosphate group(s) of a nucleotide.

A nucleoside includes a nitrogenous base and a 5-carbon sugar. Thus, a nucleoside plus a phosphate group yields a nucleotide. Nucleoside analogs are compounds that have the general structure of a nucleoside or are structurally similar to a nucleoside. Nucleoside analogs, for example, include an analog of the nucleobase and/or an analog of the sugar of a nucleoside.

It should be understood that the term “nucleotide” includes naturally-occurring nucleotides, synthetic nucleotides and modified nucleotides, unless indicated otherwise. Examples of naturally-occurring nucleotides used for the production of RNA, *e.g.*, in an IVT reaction, as provided herein include adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), uridine triphosphate (UTP), and 5-methyluridine triphosphate (m^5 UTP). In some embodiments, adenosine diphosphate (ADP), guanosine diphosphate (GDP), cytidine diphosphate (CDP), and/or uridine diphosphate (UDP) are used.

Examples of nucleotide analogs include, but are not limited to, antiviral nucleotide analogs, phosphate analogs (soluble or immobilized, hydrolyzable or non-hydrolyzable), dinucleotide, trinucleotide, tetranucleotide, *e.g.*, a cap analog, or a precursor/substrate for enzymatic capping (vaccinia or ligase), a nucleotide labeled with a functional group to facilitate ligation/conjugation of cap or 5' moiety (IRES), a nucleotide labeled with a 5' PO_4 to facilitate ligation of cap or 5' moiety, or a nucleotide labeled with a functional group/protecting group that can be chemically or enzymatically cleaved. Examples of

antiviral nucleotide/nucleoside analogs include, but are not limited, to Ganciclovir, Entecavir, Telbivudine, Vidarabine and Cidofovir.

Modified nucleotides may include modified nucleobases. For example, a RNA transcript (*e.g.*, mRNA transcript) of the present disclosure may include a modified nucleobase selected from pseudouridine (ψ), 1-methylpseudouridine (m1 ψ), 1-ethylpseudouridine, 2-thiouridine, 4'-thiouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-pseudouridine, 2-thio-5-aza-uridine, 2-thio-dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-pseudouridine, 4-methoxy-2-thio-pseudouridine, 4-methoxy-pseudouridine, 4-thio-1-methyl-pseudouridine, 4-thio-pseudouridine, 5-aza-uridine, dihydropseudouridine, 5-methyluridine, 5-methoxyuridine (mo5U) and 2'-O-methyl uridine. In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of the foregoing modified nucleobases.

The nucleoside triphosphates (NTPs) as provided herein may comprise unmodified or modified ATP, modified or unmodified UTP, modified or unmodified GTP, and/or modified or unmodified CTP. In some embodiments, NTPs of an IVT reaction comprise unmodified ATP. In some embodiments, NTPs of an IVT reaction comprise modified ATP. In some embodiments, NTPs of an IVT reaction comprise unmodified UTP. In some embodiments, NTPs of an IVT reaction comprise modified UTP. In some embodiments, NTPs of an IVT reaction comprise unmodified GTP. In some embodiments, NTPs of an IVT reaction comprise modified GTP. In some embodiments, NTPs of an IVT reaction comprise unmodified CTP. In some embodiments, NTPs of an IVT reaction comprise modified CTP.

The concentration of nucleoside triphosphates and cap analog present in an IVT reaction may vary. In some embodiments, NTPs and cap analog are present in the reaction at equimolar concentrations. In some embodiments, the molar ratio of cap analog (*e.g.*, trinucleotide cap) to nucleoside triphosphates in the reaction is greater than 1:1. For example, the molar ratio of cap analog to nucleoside triphosphates in the reaction may be 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 15:1, 20:1, 25:1, 50:1, or 100:1. In some embodiments, the molar ratio of cap analog (*e.g.*, trinucleotide cap) to nucleoside triphosphates in the reaction is less than 1:1. For example, the molar ratio of cap analog (*e.g.*, trinucleotide cap) to nucleoside triphosphates in the reaction may be 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, 1:25, 1:50, or 1:100.

The composition of NTPs in an IVT reaction may also vary. For example, ATP may be used in excess of GTP, CTP and UTP. As a non-limiting example, an IVT reaction may include 7.5 millimolar GTP, 7.5 millimolar CTP, 7.5 millimolar UTP, and 3.75 millimolar ATP. The same IVT reaction may include 3.75 millimolar cap analog (*e.g.*, trinucleotide cap). In some embodiments, the molar ratio of G:C:U:A:cap is 1:1:1:0.5:0.5. In some
 5 embodiments, the molar ratio of G:C:U:A:cap is 1:1:0.5:1:0.5. In some embodiments, the molar ratio of G:C:U:A:cap is 1:0.5:1:1:0.5. In some embodiments, the molar ratio of G:C:U:A:cap is 0.5:1:1:1:0.5.

In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes a modified
 10 nucleobase selected from pseudouridine (ψ), 1-methylpseudouridine ($m^1\psi$), 5-methoxyuridine (mo^5U), 5-methylcytidine (m^5C), α -thio-guanosine and α -thio-adenosine. In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes a combination of at least two (*e.g.*, 2, 3, 4 or more) of the foregoing modified nucleobases.

In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes
 15 pseudouridine (ψ). In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes 1-methylpseudouridine ($m^1\psi$). In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes 5-methoxyuridine (mo^5U). In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes 5-methylcytidine (m^5C). In some embodiments, a RNA transcript (*e.g.*, mRNA transcript) includes α -thio-guanosine. In some embodiments, a RNA transcript
 20 (*e.g.*, mRNA transcript) includes α -thio-adenosine.

In some embodiments, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) is uniformly modified (*e.g.*, fully modified, modified throughout the entire sequence) for a particular modification. For example, a polynucleotide can be uniformly modified with 1-methylpseudouridine ($m^1\psi$), meaning that all uridine residues in the mRNA
 25 sequence are replaced with 1-methylpseudouridine ($m^1\psi$). Similarly, a polynucleotide can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as any of those set forth above. Alternatively, the polynucleotide (*e.g.*, RNA polynucleotide, such as mRNA polynucleotide) may not be uniformly modified (*e.g.*, partially modified, part of the sequence is modified). Each
 30 possibility represents a separate embodiment of the present disclosure.

In some embodiments, the buffer system contains tris. The concentration of tris used in an IVT reaction, for example, may be at least 10 mM, at least 20 mM, at least 30 mM, at least 40 mM, at least 50 mM, at least 60 mM, at least 70 mM, at least 80 mM, at least 90 mM, at least 100 mM or at least 110 mM phosphate. In some embodiments, the concentration of phosphate is 20-60 mM or 10-100 mM.

In some embodiments, the buffer system contains dithiothreitol (DTT). The concentration of DTT used in an IVT reaction, for example, may be at least 1 mM, at least 5 mM, or at least 50 mM. In some embodiments, the concentration of DTT used in an IVT reaction is 1-50 mM or 5-50 mM. In some embodiments, the concentration of DTT used in an IVT reaction is 5 mM.

In some embodiments, the buffer system contains magnesium. In some embodiments, the molar ratio of NTP to magnesium ions (Mg^{2+} ; *e.g.*, $MgCl_2$) present in an IVT reaction is 1:1 to 1:5. For example, the molar ratio of NTP to magnesium ions may be 1:1, 1:2, 1:3, 1:4 or 1:5.

In some embodiments, the molar ratio of NTP plus cap analog (*e.g.*, trinucleotide cap, such as GAG) to magnesium ions (Mg^{2+} ; *e.g.*, $MgCl_2$) present in an IVT reaction is 1:1 to 1:5. For example, the molar ratio of NTP+trinucleotide cap (*e.g.*, GAG) to magnesium ions may be 1:1, 1:2, 1:3, 1:4 or 1:5.

In some embodiments, the buffer system contains Tris-HCl, spermidine (*e.g.*, at a concentration of 1-30 mM), TRITON[®] X-100 (polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether) and/or polyethylene glycol (PEG).

The addition of nucleoside triphosphates (NTPs) to the 3' end of a growing RNA strand is catalyzed by a polymerase, such as T7 RNA polymerase, for example, any one or more of the T7 RNA polymerase variants (*e.g.*, G47A) of the present disclosure. In some embodiments, the RNA polymerase (*e.g.*, T7 RNA polymerase variant) is present in a reaction (*e.g.*, an IVT reaction) at a concentration of 0.01 mg/ml to 1 mg/ml. For example, the RNA polymerase may be present in a reaction at a concentration of 0.01 mg/mL, 0.05 mg/ml, 0.1 mg/ml, 0.5 mg/ml or 1.0 mg/ml.

In some embodiments, the polynucleotide of the present disclosure is an IVT polynucleotide. Traditionally, the basic components of an mRNA molecule include at least a

coding region, a 5'UTR, a 3'UTR, a 5' cap and a poly-A tail. The IVT polynucleotides of the present disclosure can function as mRNA but are distinguished from wild-type mRNA in their functional and/or structural design features which serve, *e.g.*, to overcome existing problems of effective polypeptide production using nucleic-acid based therapeutics.

5 The primary construct of an IVT polynucleotide comprises a first region of linked nucleotides that is flanked by a first flanking region and a second flanking region. This first region can include, but is not limited to, the encoded ASL polypeptide. The first flanking region can include a sequence of linked nucleosides which function as a 5' untranslated region (UTR) such as the 5' UTR of any of the nucleic acids encoding the native 5' UTR of
10 the polypeptide or a non-native 5'UTR such as, but not limited to, a heterologous 5' UTR or a synthetic 5' UTR. The IVT encoding an ASL polypeptide can comprise at its 5 terminus a signal sequence region encoding one or more signal sequences. The flanking region can comprise a region of linked nucleotides comprising one or more complete or incomplete 5' UTRs sequences. The flanking region can also comprise a 5' terminal cap. The second
15 flanking region can comprise a region of linked nucleotides comprising one or more complete or incomplete 3' UTRs which can encode the native 3' UTR of an ASL polypeptide, or a non-native 3' UTR such as, but not limited to, a heterologous 3' UTR or a synthetic 3' UTR. The flanking region can also comprise a 3' tailing sequence. The 3' tailing sequence can be, but is not limited to, a polyA tail, a polyA-G quartet and/or a stem loop sequence.

20 Additional and exemplary features of IVT polynucleotide architecture and methods of making a polynucleotide are disclosed in PCT International application WO 2017/201325, filed on 18 May 2017, the entire contents of which are hereby incorporated by reference.

b. Chemical synthesis

25 Standard methods can be applied to synthesize an isolated polynucleotide sequence encoding an isolated polypeptide of interest, such as a polynucleotide of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide). For example, a single DNA or RNA oligomer containing a codon-optimized nucleotide sequence coding for the particular isolated polypeptide can be synthesized. In other aspects, several small oligonucleotides coding for portions of the desired polypeptide
30 can be synthesized and then ligated. In some aspects, the individual oligonucleotides typically contain 5' or 3' overhangs for complementary assembly.

A polynucleotide disclosed herein (*e.g.*, a RNA, *e.g.*, an mRNA) can be chemically synthesized using chemical synthesis methods and potential nucleobase substitutions known in the art. See, for example, International Publication Nos. WO2014093924, WO2013052523; WO2013039857, WO2012135805, WO2013151671; U.S. Publ. No. 5 US20130115272; or U.S. Pat. Nos. US8999380 or US8710200, all of which are herein incorporated by reference in their entireties.

c. Quantification of Expressed Polynucleotides Encoding ASL

In some embodiments, the polynucleotides of the present disclosure (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide), their 10 expression products, as well as degradation products and metabolites can be quantified according to methods known in the art.

In some embodiments, the polynucleotides of the present disclosure can be quantified in exosomes or when derived from one or more bodily fluid. As used herein "bodily fluids" include peripheral blood, serum, plasma, ascites, urine, cerebrospinal fluid (CSF), sputum, 15 saliva, bone marrow, synovial fluid, aqueous humor, amniotic fluid, cerumen, breast milk, bronchoalveolar lavage fluid, semen, prostatic fluid, cowper's fluid or pre-ejaculatory fluid, sweat, fecal matter, hair, tears, cyst fluid, pleural and peritoneal fluid, pericardial fluid, lymph, chyme, chyle, bile, interstitial fluid, menses, pus, sebum, vomit, vaginal secretions, mucosal secretion, stool water, pancreatic juice, lavage fluids from sinus cavities, 20 bronchopulmonary aspirates, blastocyl cavity fluid, and umbilical cord blood. Alternatively, exosomes can be retrieved from an organ selected from the group consisting of lung, heart, pancreas, stomach, intestine, bladder, kidney, ovary, testis, skin, colon, breast, prostate, brain, esophagus, liver, and placenta.

In the exosome quantification method, a sample of not more than 2 mL is obtained 25 from the subject and the exosomes isolated by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof. In the analysis, the level or concentration of a polynucleotide can be an expression level, presence, absence, truncation or alteration of the administered construct. It is 30 advantageous to correlate the level with one or more clinical phenotypes or with an assay for a human disease biomarker.

The assay can be performed using construct specific probes, cytometry, qRT-PCR, real-time PCR, PCR, flow cytometry, electrophoresis, mass spectrometry, or combinations thereof while the exosomes can be isolated using immunohistochemical methods such as enzyme linked immunosorbent assay (ELISA) methods. Exosomes can also be isolated by size exclusion chromatography, density gradient centrifugation, differential centrifugation, nanomembrane ultrafiltration, immunoabsorbent capture, affinity purification, microfluidic separation, or combinations thereof.

These methods afford the investigator the ability to monitor, in real time, the level of polynucleotides remaining or delivered. This is possible because the polynucleotides of the present disclosure differ from the endogenous forms due to the structural or chemical modifications.

In some embodiments, the polynucleotide can be quantified using methods such as, but not limited to, ultraviolet visible spectroscopy (UV/Vis). A non-limiting example of a UV/Vis spectrometer is a NANODROP® spectrometer (ThermoFisher, Waltham, MA). The quantified polynucleotide can be analyzed in order to determine if the polynucleotide can be of proper size, check that no degradation of the polynucleotide has occurred. Degradation of the polynucleotide can be checked by methods such as, but not limited to, agarose gel electrophoresis, HPLC based purification methods such as, but not limited to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), liquid chromatography-mass spectrometry (LCMS), capillary electrophoresis (CE) and capillary gel electrophoresis (CGE).

20. Pharmaceutical Compositions and Formulations

The present disclosure provides pharmaceutical compositions and formulations that comprise any of the polynucleotides described above. In some embodiments, the composition or formulation further comprises a delivery agent.

In some embodiments, the composition or formulation can contain a polynucleotide comprising a sequence optimized nucleic acid sequence disclosed herein which encodes an ASL polypeptide. In some embodiments, the composition or formulation can contain a polynucleotide (*e.g.*, a RNA, *e.g.*, an mRNA) comprising a polynucleotide (*e.g.*, an ORF) having significant sequence identity to a sequence optimized nucleic acid sequence disclosed herein which encodes an ASL polypeptide. In some embodiments, the polynucleotide further

comprises a miRNA binding site, *e.g.*, a miRNA binding site that binds miR-126, miR-142, miR-144, miR-146, miR-150, miR-155, miR-16, miR-21, miR-223, miR-24, miR-27 and miR-26a.

Pharmaceutical compositions or formulation can optionally comprise one or more additional active substances, *e.g.*, therapeutically and/or prophylactically active substances. Pharmaceutical compositions or formulation of the present disclosure can be sterile and/or pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents can be found, for example, in *Remington: The Science and Practice of Pharmacy* 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference in its entirety). In some embodiments, compositions are administered to humans, human patients or subjects. For the purposes of the present disclosure, the phrase "active ingredient" generally refers to polynucleotides to be delivered as described herein.

Formulations and pharmaceutical compositions described herein can be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of associating the active ingredient with an excipient and/or one or more other accessory ingredients, and then, if necessary and/or desirable, dividing, shaping and/or packaging the product into a desired single- or multi-dose unit.

A pharmaceutical composition or formulation in accordance with the present disclosure can be prepared, packaged, and/or sold in bulk, as a single unit dose, and/or as a plurality of single unit doses. As used herein, a "unit dose" refers to a discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. The amount of the active ingredient is generally equal to the dosage of the active ingredient that would be administered to a subject and/or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the present disclosure can vary, depending upon the identity, size, and/or condition of the subject being treated and further depending upon the route by which the composition is to be administered.

In some embodiments, the compositions and formulations described herein can contain at least one polynucleotide of the present disclosure. As a non-limiting example, the

composition or formulation can contain 1, 2, 3, 4 or 5 polynucleotides of the present disclosure. In some embodiments, the compositions or formulations described herein can comprise more than one type of polynucleotide. In some embodiments, the composition or formulation can comprise a polynucleotide in linear and circular form. In some embodiments, the composition or formulation can comprise a circular polynucleotide and an *in vitro* transcribed (IVT) polynucleotide. In yet another embodiment, the composition or formulation can comprise an IVT polynucleotide, a chimeric polynucleotide and a circular polynucleotide.

Although the descriptions of pharmaceutical compositions and formulations provided herein are principally directed to pharmaceutical compositions and formulations that are suitable for administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to any other animal, *e.g.*, to non-human animals, *e.g.* non-human mammals.

The present disclosure provides pharmaceutical formulations that comprise a polynucleotide described herein (*e.g.*, a polynucleotide comprising a nucleotide sequence encoding an ASL polypeptide). The polynucleotides described herein can be Formulated using one or more excipients to: (1) increase stability; (2) increase cell transfection; (3) permit the sustained or delayed release (*e.g.*, from a depot formulation of the polynucleotide); (4) alter the biodistribution (*e.g.*, target the polynucleotide to specific tissues or cell types); (5) increase the translation of encoded protein *in vivo*; and/or (6) alter the release profile of encoded protein *in vivo*.

In some embodiments of the pharmaceutical formulations disclosed herein, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) disclosed herein is Formulated with a delivery agent comprising LNP-1A, LNP 1-B, LNP-2A, LNP-2B, LNP-3A, or LNP-3B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-1A. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-1B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-2A. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-2B. In some embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-3A. In some

embodiments, the polynucleotide (*e.g.*, a RNA, *e.g.*, a mRNA) of the present disclosure is Formulated with LNP-3B.

A pharmaceutically acceptable excipient, as used herein, includes, but are not limited to, any and all solvents, dispersion media, or other liquid vehicles, dispersion or suspension
5 aids, diluents, granulating and/or dispersing agents, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, binders, lubricants or oil, coloring, sweetening or flavoring agents, stabilizers, antioxidants, antimicrobial or antifungal agents, osmolality adjusting agents, pH adjusting agents, buffers, chelants, cyoprotectants, and/or bulking agents, as suited to the particular dosage form desired. Various excipients for
10 Formulating pharmaceutical compositions and techniques for preparing the composition are known in the art (see Remington: The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro (Lippincott, Williams & Wilkins, Baltimore, MD, 2006; incorporated herein by reference in its entirety).

Exemplary diluents include, but are not limited to, calcium or sodium carbonate,
15 calcium phosphate, calcium hydrogen phosphate, sodium phosphate, lactose, sucrose, cellulose, microcrystalline cellulose, kaolin, mannitol, sorbitol, etc., and/or combinations thereof.

Exemplary surface active agents and/or emulsifiers include, but are not limited to, natural emulsifiers (*e.g.*, acacia, agar, alginic acid, sodium alginate, tragacanth, chondrux,
20 cholesterol, xanthan, pectin, gelatin, egg yolk, casein, wool fat, cholesterol, wax, and lecithin), sorbitan fatty acid esters (*e.g.*, polyoxyethylene sorbitan monooleate [TWEEN®80], sorbitan monopalmitate [SPAN®40], glyceryl monooleate, polyoxyethylene esters, polyethylene glycol fatty acid esters (*e.g.*, CREMOPHOR®), polyoxyethylene ethers (*e.g.*, polyoxyethylene lauryl ether [BRIJ®30]), PLUORINC®F 68, POLOXAMER®188, etc.
25 and/or combinations thereof.

Exemplary binding agents include, but are not limited to, starch, gelatin, sugars (*e.g.*, sucrose, glucose, dextrose, dextrin, molasses, lactose, lactitol, mannitol), amino acids (*e.g.*, glycine), natural and synthetic gums (*e.g.*, acacia, sodium alginate), ethylcellulose, hydroxyethylcellulose, hydroxypropyl methylcellulose, etc., and combinations thereof.

30 Oxidation is a potential degradation pathway for mRNA, especially for liquid mRNA formulations. In order to prevent oxidation, antioxidants can be added to the formulations.

Exemplary antioxidants include, but are not limited to, alpha tocopherol, ascorbic acid, ascorbyl palmitate, benzyl alcohol, butylated hydroxyanisole, m-cresol, methionine, butylated hydroxytoluene, monothioglycerol, sodium or potassium metabisulfite, propionic acid, propyl gallate, sodium ascorbate, etc., and combinations thereof.

5 Exemplary chelating agents include, but are not limited to, ethylenediaminetetraacetic acid (EDTA), citric acid monohydrate, disodium edetate, fumaric acid, malic acid, phosphoric acid, sodium edetate, tartaric acid, trisodium edetate, etc., and combinations thereof.

10 Exemplary antimicrobial or antifungal agents include, but are not limited to, benzalkonium chloride, benzethonium chloride, methyl paraben, ethyl paraben, propyl paraben, butyl paraben, benzoic acid, hydroxybenzoic acid, potassium or sodium benzoate, potassium or sodium sorbate, sodium propionate, sorbic acid, etc., and combinations thereof.

15 Exemplary preservatives include, but are not limited to, vitamin A, vitamin C, vitamin E, beta-carotene, citric acid, ascorbic acid, butylated hydroxyanisol, ethylenediamine, sodium lauryl sulfate (SLS), sodium lauryl ether sulfate (SLES), etc., and combinations thereof.

In some embodiments, the pH of polynucleotide solutions is maintained between pH 5 and pH 8 to improve stability. Exemplary buffers to control pH can include, but are not limited to sodium phosphate, sodium citrate, sodium succinate, histidine (or histidine-HCl), sodium malate, sodium carbonate, etc., and/or combinations thereof.

20 Exemplary lubricating agents include, but are not limited to, magnesium stearate, calcium stearate, stearic acid, silica, talc, malt, hydrogenated vegetable oils, polyethylene glycol, sodium benzoate, sodium or magnesium lauryl sulfate, etc., and combinations thereof.

25 The pharmaceutical composition or formulation described here can contain a cryoprotectant to stabilize a polynucleotide described herein during freezing. Exemplary cryoprotectants include, but are not limited to mannitol, sucrose, trehalose, lactose, glycerol, dextrose, etc., and combinations thereof.

The pharmaceutical composition or formulation described here can contain a bulking agent in lyophilized polynucleotide formulations to yield a "pharmaceutically elegant" cake, stabilize the lyophilized polynucleotides during long term (*e.g.*, 36 month) storage.

Exemplary bulking agents of the present disclosure can include, but are not limited to sucrose, trehalose, mannitol, glycine, lactose, raffinose, and combinations thereof.

In some embodiments, the pharmaceutical composition or formulation further comprises a delivery agent. The delivery agent of the present disclosure can include, without
5 limitation, liposomes, lipid nanoparticles, lipidoids, polymers, lipoplexes, microvesicles, exosomes, peptides, proteins, cells transfected with polynucleotides, hyaluronidase, nanoparticle mimics, nanotubes, conjugates, and combinations thereof.

21. Methods of Use

10 The polynucleotides, pharmaceutical compositions and formulations described above are used in the preparation, manufacture, and therapeutic use of to treat and/or prevent ASL-related diseases, disorders or conditions (*e.g.*, argininosuccinic aciduria). In some embodiments, the polynucleotides, compositions and formulations of the present disclosure are used to treat and/or prevent argininosuccinic aciduria.

15 In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of treating or delaying the onset and/or progression of argininosuccinic aciduria in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described
20 above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of reducing plasma ammonia levels in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical
25 compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of reducing plasma argininosuccinic acid (ASA) levels in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides,
30 polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of reducing plasma citrulline levels in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of increasing ASL levels in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of increasing ASL activity in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of reducing orotate levels or orotic acid levels in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the polynucleotides, polypeptides, pharmaceutical compositions, and formulations of the present disclosure are used in a method of reducing ureagenesis in a subject (*e.g.*, a human subject), comprising: administering to the subject an effective amount of any of the the polynucleotides, polypeptides, pharmaceutical compositions, and formulations described above.

In some embodiments, the present disclosure relates to use of any of the polynucleotides, polypeptides, or pharmaceutical compositions or formulations discussed above in the manufacture of a medicament for treating and/or preventing an ASL-related disease, treating and/or preventing argininosuccinic aciduria, treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL

levels, increasing ASL activity, increasing glutathione levels (e.g., liver glutathione levels), reducing orotate levels or orotic acid levels, reducing ureagenesis, reducing liver xCT antiporter levels, or reducing liver xCT antiporter activity, in a subject.

5 In some embodiments, the polynucleotides, pharmaceutical compositions and formulations of the present disclosure are used in methods for reducing the levels of ammonia in a subject in need thereof, *e.g.*, a subject with hyperammonemia. For instance, one aspect of the present disclosure provides a method of alleviating the signs and symptoms of argininosuccinic aciduria in a subject comprising the administration of a composition or formulation comprising a polynucleotide encoding ASL to that subject (e.g., an mRNA
10 encoding an ASL polypeptide).

In some embodiments, the administration of an effective amount of a polynucleotide, pharmaceutical composition or formulation of the present disclosure reduces the levels of a biomarker of argininosuccinic aciduria, *e.g.*, ammonia, ASA, citrulline, orotate, and/or any combination thereof. In some embodiments, the administration of the polynucleotide,
15 pharmaceutical composition or formulation of the present disclosure results in reduction in the level of one or more biomarkers of argininosuccinic aciduria, *e.g.*, ammonia, ASA, citrulline, and/or orotate, within a short period of time (*e.g.*, within about 6 hours, within about 8 hours, within about 12 hours, within about 16 hours, within about 20 hours, or within about 24 hours) after administration of the polynucleotide, pharmaceutical composition or
20 formulation of the present disclosure.

In some embodiments, the administration of an effective amount of a polynucleotide, pharmaceutical composition or formulation of the present disclosure increases body weight of a subject (*e.g.*, a human subject). In some embodiments, the administration of the polynucleotide, polypeptide, pharmaceutical composition, or formulation of the present
25 disclosure results in an increase in body weight within a short period of time (*e.g.*, within about 12 hours, 24 hours, 48 hours, 72 hours, 96 hours, 5 days, 7 days, 14 days, 24 days, 48 days, or 60 days) after administration of the polynucleotide, pharmaceutical composition or formulation of the present disclosure. In some embodiments, the administration of an effective amount of a polynucleotide, pharmaceutical composition or formulation of the
30 present disclosure maintains body weight of a subject (*e.g.*, human subject).

Replacement therapy is a potential treatment for ASL. Thus, in certain aspects of the present disclosure, the polynucleotides, *e.g.*, mRNA, disclosed herein comprise one or more sequences encoding an ASL polypeptide that is suitable for use in gene replacement therapy for argininosuccinic aciduria. In some embodiments, the present disclosure treats a lack of ASL or ASL activity, or decreased or abnormal ASL activity in a subject by providing a polynucleotide, *e.g.*, mRNA, that encodes an ASL polypeptide to the subject. In some embodiments, the polynucleotide is sequence-optimized. In some embodiments, the polynucleotide (*e.g.*, an mRNA) comprises a nucleic acid sequence (*e.g.*, an ORF) encoding an ASL polypeptide, wherein the nucleic acid is sequence-optimized, *e.g.*, by modifying its G/C, uridine, or thymidine content, and/or the polynucleotide comprises at least one chemically modified nucleoside. In some embodiments, the polynucleotide comprises a miRNA binding site, *e.g.*, a miRNA binding site that binds miRNA-142.

In some embodiments, the administration of a composition or formulation comprising polynucleotide, pharmaceutical composition or formulation of the present disclosure to a subject results in a decrease in ammonia (*e.g.*, in plasma or in cells) to a level at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or to 100% lower than the level observed prior to the administration of the composition or formulation.

In some embodiments, the administration of the polynucleotide, pharmaceutical composition or formulation of the present disclosure results in expression of ASL in cells of the subject. In some embodiments, administering the polynucleotide, pharmaceutical composition or formulation of the present disclosure results in an increase of ASL enzymatic activity in the subject. For example, in some embodiments, the polynucleotides of the present disclosure are used in methods of administering a composition or formulation comprising an mRNA encoding an ASL polypeptide to a subject, wherein the method results in an increase of ASL enzymatic activity in at least some cells of a subject.

In some embodiments, the administration of a composition or formulation comprising an mRNA encoding an ASL polypeptide to a subject results in an increase of ASL enzymatic activity in cells subject to a level at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or

to 100% or more of the activity level expected in a normal subject, *e.g.*, a human not suffering from argininosuccinic aciduria.

In some embodiments, the administration of the polynucleotide, pharmaceutical composition or formulation of the present disclosure results in expression of ASL protein in
5 at least some of the cells of a subject that persists for a period of time sufficient to allow significant chloride channel activity to occur.

In some embodiments, the polynucleotides, pharmaceutical compositions, or formulations of the present disclosure can be repeatedly administered such that ASL protein is expressed at a therapeutic level for a period of time sufficient to have a beneficial
10 biological effect as described herein. In some embodiments, the repeat administration comprises more than one dose (*e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18, 20, 24, or more). In some embodiments, the interval between any two doses of the repeat administration is about 4 hours, about 6 hours, about 8 hours, about 12 hours, about 18 hours, about 20 hours, about 24 hours, about 30 hours, about 36 hours, about 40 hours, about 48 hours, about 50 hours,
15 about 60 hours, about 70 hours, about 80 hours, about 90 hours, about 100 hours, about 120 hours, about 1 day, about 2 days, about 3 days, about 4 days, about 5 days, about 6 days, about 7 days, about 8 days, about 9 days, about 10 days, about 12 days, about 14 days, about 15 days, about 18 days, about 20 days, about 21 days, about 24 days, about 25 days, about 28 days, about 30 days, about 40 days, about 50 days, about 60 days, about 90 days, about 120
20 days, about 150 days, about 180 days, about 1 week, about 2 weeks, about 3 weeks, about 4 weeks, about 5 weeks, about 6 weeks, about 7 weeks, about 8 weeks, about 9 weeks, about 10 weeks, about 11 weeks, about 12 weeks, about 15 weeks, about 20 weeks, about 25 weeks, about 30 weeks, about 40 weeks, about 50 weeks, about 52 weeks, about 1 month, about 2 months, about 3 months, about 4 months, about 5 months, about 6 months, about 7 months,
25 about 8 months, about 9 months, about 10 months, about 11 months, about 12 months, about 1 year, or any range or value therein between. In some embodiments including three or more doses, the interval between any two doses does not need to be the same (*e.g.*, one week between the first and second dose; one month between any two subsequent doses).

In some embodiments, the expression of the encoded polypeptide is increased. In
30 some embodiments, the polynucleotide increases ASL expression levels in cells when introduced into those cells, *e.g.*, by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at

least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or to 100% with respect to the ASL expression level in the cells before the polypeptide is introduced in the cells.

In some embodiments, the method or use comprises administering a polynucleotide, *e.g.*, mRNA, comprising a nucleotide sequence comprising a polynucleotide of any one of SEQ ID NOs: 20–33 or 41–44, wherein the polynucleotide sequence encodes an ASL polypeptide (*e.g.*, according to any one of SEQ ID NOs: 1–11).

Other aspects of the present disclosure relate to transplantation of cells containing polynucleotides to a mammalian subject. Administration of cells to mammalian subjects is known to those of ordinary skill in the art, and includes, but is not limited to, local implantation (*e.g.*, topical or subcutaneous administration), organ delivery or systemic injection (*e.g.*, intravenous injection or inhalation), and the formulation of cells in pharmaceutically acceptable carriers.

The present disclosure also provides methods to increase ASL activity in a subject in need thereof, *e.g.*, a subject with argininosuccinic aciduria, comprising administering to the subject a therapeutically effective amount of a composition or formulation comprising mRNA encoding an ASL polypeptide disclosed herein, *e.g.*, a human ASL polypeptide, a mutant thereof, or a fusion protein comprising a human ASL.

In some aspects, the ASL activity measured after administration to a subject in need thereof, *e.g.*, a subject with argininosuccinic aciduria, is at least the normal ASL activity level observed in healthy human subjects. In some aspects, the ASL activity measured after administration is at higher than the ASL activity level observed in argininosuccinic aciduria patients, *e.g.*, untreated argininosuccinic aciduria patients. In some aspects, the increase in ASL activity in a subject in need thereof, *e.g.*, a subject with argininosuccinic aciduria, after administering to the subject a therapeutically effective amount of a composition or formulation comprising mRNA encoding an ASL polypeptide disclosed herein is at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 55, at least about 60, at least about 65, at least about 70, at least about 75, at least about 80, at least about 85, at least about 90, at least about 95, at least about 100, or greater than 100 percent of the normal ASL activity level observed in healthy human subjects. In some aspects, the

increase in ASL activity above the ASL activity level observed in argininosuccinic aciduria patients after administering to the subject a composition or formulation comprising an mRNA encoding an ASL polypeptide disclosed herein (*e.g.*, after a single dose administration or repeated administration) is maintained for at least 1 day, at least 2 days, at least 3 days, at least 4 days, at least 5 days, at least 6 days, at least 7 days, at least 8 days, at least 9 days, at least 10 days, at least 12 days, at least 14 days, at least 21 days, or at least 28 days.

The present disclosure also provides a method to treat, prevent, or ameliorate the symptoms of argininosuccinic aciduria (*e.g.*, high levels of ammonia in the blood, hyperammonemia, high levels of orotic acid, refusal to eat, vomiting, lethargy, irritability, seizures, hypotonia, hepatomegaly, respiratory abnormalities, and edema within the brain) in an argininosuccinic aciduria patient comprising administering to the subject a therapeutically effective amount of a composition or formulation comprising mRNA encoding an ASL polypeptide disclosed herein. In some aspects, the administration of a therapeutically effective amount of a composition or formulation comprising mRNA encoding an ASL polypeptide disclosed herein to subject in need of treatment for argininosuccinic aciduria results in reducing the symptoms of argininosuccinic aciduria.

The skilled artisan will appreciate that the therapeutic effectiveness of a drug or a treatment of the instant invention can be characterized or determined by measuring the level of expression of an encoded protein (*e.g.*, enzyme) in a sample or in samples taken from a subject (*e.g.*, from a preclinical test subject (rodent, primate, etc.) or from a clinical subject (human)). Likewise, the therapeutic effectiveness of a drug or a treatment of the instant invention can be characterized or determined by measuring the level of activity of an encoded protein (*e.g.*, enzyme) in a sample or in samples taken from a subject (*e.g.*, from a preclinical test subject (rodent, primate, etc.) or from a clinical subject (human)). Furthermore, the therapeutic effectiveness of a drug or a treatment of the instant invention can be characterized or determined by measuring the level of an appropriate biomarker in sample(s) taken from a subject. Levels of protein and/or biomarkers can be determined post-administration with a single dose of an mRNA therapeutic of the present disclosure or can be determined and/or monitored at several time points following administration with a single dose or can be determined and/or monitored throughout a course of treatment, *e.g.*, a multi-dose treatment.

ASL Protein Expression Levels

Certain aspects of the present disclosure feature measurement, determination and/or monitoring of the expression level or levels of ASL protein in a subject, for example, in an animal (*e.g.*, rodents, primates, and the like) or in a human subject. Animals include normal, healthy or wild type animals, as well as animal models for use in understanding argininosuccinic aciduria and treatments thereof. Exemplary animal models include rodent models, for example, ASL deficient mice also referred to as ASL mice.

ASL protein expression levels can be measured or determined by any art-recognized method for determining protein levels in biological samples, *e.g.*, from blood samples or a needle biopsy. The term "level" or "level of a protein" as used herein, preferably means the weight, mass or concentration of the protein within a sample or a subject. It will be understood by the skilled artisan that in certain embodiments the sample may be subjected, *e.g.*, to any of the following: purification, precipitation, separation, *e.g.* centrifugation and/or HPLC, and subsequently subjected to determining the level of the protein, *e.g.*, using mass and/or spectrometric analysis. In exemplary embodiments, enzyme-linked immunosorbent assay (ELISA) can be used to determine protein expression levels. In other exemplary embodiments, protein purification, separation and LC-MS can be used as a means for determining the level of a protein according to the invention. In some embodiments, an mRNA therapy of the present disclosure (*e.g.*, a single intravenous dose) results in increased ASL protein expression levels in the tissue (*e.g.*, heart, liver, brain, or skeletal muscle) of the subject (*e.g.*, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold increase and/or increased to at least 50%, at least 60%, at least 70%, at least 75%, 80%, at least 85%, at least 90%, at least 95%, or at least 100% of normal levels) for at least 6 hours, at least 12 hours, at least 24 hours, at least 36 hours, at least 48 hours, at least 60 hours, at least 72 hours, at least 84 hours, at least 96 hours, at least 108 hours, at least 122 hours after administration of a single dose or multiple doses of the mRNA therapy.

ASL Protein Activity

In argininosuccinic aciduria patients, ASL enzymatic activity is reduced compared to a normal physiological activity level. Further aspects of the present disclosure feature measurement, determination and/or monitoring of the activity level(s) (*i.e.*, enzymatic activity level(s)) of ASL protein in a subject, for example, in an animal (*e.g.*, rodent, primate, and the like) or in a human subject. Activity levels can be measured or determined by any art-

recognized method for determining enzymatic activity levels in biological samples. The term "activity level" or "enzymatic activity level" as used herein, preferably means the activity of the enzyme per volume, mass or weight of sample or total protein within a sample. In exemplary embodiments, the "activity level" or "enzymatic activity level" is described in terms of units per milliliter of fluid (*e.g.*, bodily fluid, *e.g.*, serum, plasma, urine and the like) or is described in terms of units per weight of tissue or per weight of protein (*e.g.*, total protein) within a sample. Units ("U") of enzyme activity can be described in terms of weight or mass of substrate hydrolyzed per unit time. In certain embodiments of the present disclosure feature ASL activity described in terms of U/ml plasma or U/mg protein (tissue), where units ("U") are described in terms of nmol substrate hydrolyzed per hour (or nmol/hr).

In certain embodiments, an mRNA therapy of the present disclosure features a pharmaceutical composition comprising a dose of mRNA effective to result in at least 5 U/mg, at least 10 U/mg, at least 20 U/mg, at least 30 U/mg, at least 40 U/mg, at least 50 U/mg, at least 60 U/mg, at least 70 U/mg, at least 80 U/mg, at least 90 U/mg, at least 100 U/mg, or at least 150 U/mg of ASL activity in tissue (*e.g.*, liver) between 6 and 12 hours, or between 12 and 24, between 24 and 48, or between 48 and 72 hours post administration (*e.g.*, at 48 or at 72 hours post administration).

In some embodiments, an mRNA therapy of the present disclosure (*e.g.*, a single intravenous dose) results in increased ASL activity levels in the liver tissue of the subject (*e.g.*, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold increase and/or increased to at least 50%, at least 60%, at least 70%, at least 75%, 80%, at least 85%, at least 90%, at least 95%, or at least 100% of normal levels) for at least 6 hours, at least 12 hours, at least 24 hours, or at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 or more days after administration of a single dose or multiple doses of the mRNA therapy.

In exemplary embodiments, an mRNA therapy of the present disclosure features a pharmaceutical composition comprising a single intravenous dose of mRNA that results in the above-described levels of activity. In some embodiments, an mRNA therapy of the present disclosure features a pharmaceutical composition which can be administered in multiple single unit intravenous doses of mRNA that maintain the above-described levels of activity.

ASL Biomarkers

In some embodiments, the administration of an effective amount of a polynucleotide, pharmaceutical composition or formulation of the present disclosure reduces the levels of a biomarker of ASL, *e.g.*, ammonia, ASA, citrulline, or orotate levels. In some embodiments, the administration of the polynucleotide, pharmaceutical composition or formulation of the present disclosure results in reduction in the level of one or more biomarkers of ASL, *e.g.*, ammonia, ASA, citrulline, or orotate, within a short period of time after administration of the polynucleotide, pharmaceutical composition or formulation of the present disclosure.

In some embodiments, a polynucleotide of the present disclosure (or lipid nanoparticle comprising a polynucleotide according to the present disclosure), when administered as a single dose or, as multiple doses, to a subject (*e.g.*, a human subject), is sufficient to:

(i) reduce plasma ammonia levels in the subject to a level of less than or equal to about 500 μM , less than or equal to about 450 μM , less than or equal to about 400 μM , less than or equal to about 350 μM , less than or equal to about 300 μM , less than or equal to about 250 μM , less than or equal to about 200 μM , less than or equal to about 150 μM , less than or equal to about 100 μM , less than or equal to about 75 μM , less than or equal to about 50 μM , less than or equal to about 40 μM , less than or equal to about 30 μM , less than or equal to about 25 μM , or less than or equal to about 10 μM , for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(ii) reduce plasma ammonia levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(iii) reduce plasma citrulline levels in the subject to less than or equal to about 500 μM , less than or equal to about 450 μM , less than or equal to about 400 μM , less than or equal to about 350 μM , less than or equal to about 300 μM , less than or equal to about 250 μM , less than or equal to about 200 μM , less than or equal to about 150 μM , less than or

equal to about 100 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

5 (iv) reduce plasma citrulline levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at
10 least about 95%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

 (v) reduce plasma argininosuccinic acid (ASA) levels in the subject to less than or equal to about 150 μ M, less than or equal to about less than or equal to about 2000 μ M, less than or equal to about 1500 μ M, less than or equal to about 1000 μ M, less than or equal to
15 about 900 μ M, less than or equal to about 800 μ M, less than or equal to about 700 μ M, less than or equal to about 600 μ M, less than or equal to about 500 μ M, less than or equal to about 450 μ M, less than or equal to about 400 μ M, less than or equal to about 350 μ M, less than or equal to about 300 μ M, less than or equal to about 250 μ M, less than or equal to about 200 μ M, less than or equal to about 150 μ M, less than or equal to about 100 μ M, less than or
20 equal to about 75 μ M, less than or equal to about 50 μ M, less than or equal to about 40 μ M, less than or equal to about 30 μ M, less than or equal to about 25 μ M, or less than or equal to about 10 μ M, less than or equal to about 5 μ M, or less, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

25 (vi) reduce plasma argininosuccinic acid (ASA) levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at
30 least about 94%, at least about 95%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

 (vii) increase cellular ASL levels in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least

about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, compared to the subject's baseline cellular ASL levels for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120 hours, or at least one week post-administration; or

(viii) increase ASL activity in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline ASL activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(ix) reduce plasma orotate levels to less than or equal to about 150 $\mu\text{mol}/\text{mmol}$ creatinine, less than or equal to about 100 $\mu\text{mol}/\text{mmol}$ creatinine, less than or equal to about 50 $\mu\text{mol}/\text{mmol}$ creatinine, less than or equal to about 25 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 10 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 5 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 4 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 3 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 2 $\mu\text{mol}/\text{mmol}$ creatinine, or less than or equal to about 1 $\mu\text{mol}/\text{mmol}$ creatinine, or less, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(x) reduce plasma orotate levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(xi) increase liver ASL activity in the subject to a level of at least about 0.4, at least about 0.5, at least about 0.6, at least about 0.7, at least about 0.8, at least about 0.9, at least about 1.0, at least about 1.1, at least about 1.2, or at least about 1.5, relative to wildtype liver ASL activity;

(xii) increase glutathione levels (e.g., liver glutathione levels) in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least

about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

5 (xiii) reduce liver xCT antiporter levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at
10 least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(xiv) reduce liver xCT antiporter activity in the subject in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least
15 about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline ASL activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(xv) reduce total plasma homocysteine (HCyS) levels in the subject by at least
20 about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline total plasma HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

25 (xvi) reduce total liver homocysteine (HCyS) levels in the subject at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline total liver HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96
30 hours, at least 120, or at least one week post-administration.

Further aspects of the present disclosure feature determining the level (or levels) of a biomarker determined in a sample as compared to a level (*e.g.*, a reference level) of the same or another biomarker in another sample, *e.g.*, from the same patient, from another patient,

from a control and/or from the same or different time points, and/or a physiologic level, and/or an elevated level, and/or a supraphysiologic level, and/or a level of a control. The skilled artisan will be familiar with physiologic levels of biomarkers, for example, levels in normal or wild type animals, normal or healthy subjects, and the like, in particular, the level or levels characteristic of subjects who are healthy and/or normal functioning. As used herein, the phrase “elevated level” means amounts greater than normally found in a normal or wild type preclinical animal or in a normal or healthy subject, *e.g.* a human subject. As used herein, the term “supraphysiologic” means amounts greater than normally found in a normal or wild type preclinical animal or in a normal or healthy subject, *e.g.* a human subject, optionally producing a significantly enhanced physiologic response. As used herein, the term “comparing” or “compared to” preferably means the mathematical comparison of the two or more values, *e.g.*, of the levels of the biomarker(s). It will thus be readily apparent to the skilled artisan whether one of the values is higher, lower or identical to another value or group of values if at least two of such values are compared with each other. Comparing or comparison to can be in the context, for example, of comparing to a control value, *e.g.*, as compared to a reference blood, serum, plasma, and/or tissue (*e.g.*, liver) ammonia, ASA, citrulline, and/or orotate level, in said subject prior to administration (*e.g.*, in a person suffering from argininosuccinic aciduria) or in a normal or healthy subject. Comparing or comparison to can also be in the context, for example, of comparing to a control value, *e.g.*, as compared to a reference blood, serum, plasma and/or tissue (*e.g.*, liver) ammonia, ASA, citrulline, and/or orotate level in said subject prior to administration (*e.g.*, in a person suffering from argininosuccinic aciduria) or in a normal or healthy subject.

As used herein, a “control” is preferably a sample from a subject wherein the argininosuccinic aciduria status of said subject is known. In some embodiments, a control is a sample of a healthy patient. In some embodiments, the control is a sample from at least one subject having a known argininosuccinic aciduria status, for example, a severe, mild, or healthy argininosuccinic aciduria status, *e.g.* a control patient. In some embodiments, the control is a sample from a subject not being treated for argininosuccinic aciduria. In a still further embodiment, the control is a sample from a single subject or a pool of samples from different subjects and/or samples taken from the subject(s) at different time points.

The term “level” or “level of a biomarker” as used herein, preferably means the mass, weight or concentration of a biomarker of the present disclosure within a sample or a subject.

It will be understood by the skilled artisan that in certain embodiments the sample may be subjected to, *e.g.*, one or more of the following: substance purification, precipitation, separation, *e.g.* centrifugation and/or HPLC and subsequently subjected to determining the level of the biomarker, *e.g.* using mass spectrometric analysis. In certain embodiments, LC-
5 MS can be used as a means for determining the level of a biomarker according to the invention.

The term "determining the level" of a biomarker as used herein can mean methods which include quantifying an amount of at least one substance in a sample from a subject, for example, in a bodily fluid from the subject (*e.g.*, serum, plasma, urine, lymph, etc.) or in a
10 tissue of the subject (*e.g.*, liver, etc.).

The term "reference level" as used herein can refer to levels (*e.g.*, of a biomarker) in a subject prior to administration of an mRNA therapy of the present disclosure (*e.g.*, in a person suffering from argininosuccinic aciduria) or in a normal or healthy subject.

As used herein, the term "normal subject" or "healthy subject" refers to a subject not
15 suffering from symptoms associated with argininosuccinic aciduria. Moreover, a subject will be considered to be normal (or healthy) if it has no mutation of the functional portions or domains of the ASL gene and/or no mutation of the ASL gene resulting in a reduction of or deficiency of the enzyme ASL or the activity thereof, resulting in symptoms associated with argininosuccinic aciduria. Said mutations will be detected if a sample from the subject is
20 subjected to a genetic testing for such ASL mutations. In certain embodiments of the present disclosure, a sample from a healthy subject is used as a control sample, or the known or standardized value for the level of biomarker from samples of healthy or normal subjects is used as a control.

In some embodiments, comparing the level of the biomarker in a sample from a
25 subject in need of treatment for argininosuccinic aciduria or in a subject being treated for argininosuccinic aciduria to a control level of the biomarker comprises comparing the level of the biomarker in the sample from the subject (in need of treatment or being treated for argininosuccinic aciduria) to a baseline or reference level, wherein if a level of the biomarker in the sample from the subject (in need of treatment or being treated for argininosuccinic
30 aciduria) is elevated, increased or higher compared to the baseline or reference level, this is indicative that the subject is suffering from argininosuccinic aciduria and/or is in need of

treatment; and/or wherein if a level of the biomarker in the sample from the subject (in need of treatment or being treated for argininosuccinic aciduria) is decreased or lower compared to the baseline level this is indicative that the subject is not suffering from, is successfully being treated for argininosuccinic aciduria, or is not in need of treatment for argininosuccinic aciduria. The stronger the reduction (*e.g.*, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 10-fold, at least 20-fold, at least 30-fold, at least 40-fold, at least 50-fold reduction and/or at least 10%, at least 20%, at least 30% at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100% reduction) of the level of a biomarker, within a certain time period, *e.g.*, within 6 hours, within 12 hours, 24 hours, 36 hours, 48 hours, 60 hours, or 72 hours, and/or for a certain duration of time, *e.g.*, 48 hours, 72 hours, 96 hours, 120 hours, 144 hours, 1 week, 2 weeks, 3 weeks, 4 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 12 months, 18 months, 24 months, etc. the more successful is a therapy, such as for example an mRNA therapy of the present disclosure (*e.g.*, a single dose or a multiple-dose regimen).

A reduction of at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least 100% or more of the level of biomarker, in particular, in bodily fluid (*e.g.*, plasma, serum, urine, *e.g.*, urinary sediment) or in tissue(s) in a subject (*e.g.*, liver), within 1, 2, 3, 4, 5, 6 or more days following administration is indicative of a dose suitable for successful treatment argininosuccinic aciduria, wherein reduction as used herein, preferably means that the level of biomarker determined at the end of a specified time period (*e.g.*, post-administration, for example, of a single intravenous dose) is compared to the level of the same biomarker determined at the beginning of said time period (*e.g.*, pre-administration of said dose). Exemplary time periods include 12, 24, 48, 72, 96, 120 or 144 hours post administration, in particular 24, 48, 72 or 96 hours post administration.

By way of non-limiting example, plasma ASA or citrulline levels may be reduced by at least about 50% (*e.g.*, 50% to 90%, 50% to 80%, 50% to 75%, 50% to 70%, 50% to 65%, 50 to 60%, 55% to 70%, 55% to 65%,) within 12 hours, 14 hours, 16 hours, 18 hours, 20 hours, 24 hours, 30 hours, 36 hours, 48 hours, 60 hours, or 72 hours post-administration, relative to the plasma ASA or plasma citrulline level before administration.

A sustained reduction in substrate levels (*e.g.*, biomarkers) is particularly indicative of mRNA therapeutic dosing and/or administration regimens successful for treatment of argininosuccinic aciduria. Such sustained reduction can be referred to herein as “duration” of effect. In exemplary embodiments, a reduction of at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% at least about 96%, at least about 97%, at least about 98%, at least about 99%, at least about 100% or more of the level of biomarker, in particular, in a bodily fluid (*e.g.*, plasma, serum, urine, *e.g.*, urinary sediment) or in tissue(s) in a subject (*e.g.*, liver), within 1, 2, 3, 4, 5, 6, 7, 8 or more days following administration is indicative of a successful therapeutic approach. In exemplary embodiments, sustained reduction in substrate (*e.g.*, biomarker) levels in one or more samples (*e.g.*, fluids and/or tissues) is preferred. For example, mRNA therapies resulting in sustained reduction in a biomarker, optionally in combination with sustained reduction of said biomarker in at least one tissue, preferably two, three, four, five or more tissues, is indicative of successful treatment.

In some embodiments, a single dose of an mRNA therapy of the present disclosure is less than or equal to about 1.5 mpk (mgs/kg), less than or equal to about 1.4 mpk, less than or equal to about 1.3 mpk, less than or equal to about 1.2 mpk, less than or equal to about 1.1 mpk, less than or equal to about 1.0 mpk, less than or equal to about 0.9 mpk, less than or equal to about 0.8 mpk, less than or equal to about 0.75 mpk, less than or equal to about 0.7 mpk, less than or equal to about 0.6 mpk, less than or equal to about 0.5 mpk, less than or equal to about 0.4 mpk, less than or equal to about 0.2 mpk, or any range or value therein between. In some embodiments, a single dose of an mRNA therapy of the present disclosure is about 0.2 mpk to about 1.5 mpk, about 0.2 mpk to about 1.4 mpk, about 0.2 mpk to about 1.3 mpk, about 0.2 mpk to about 1.2 mpk, about 0.2 mpk to about 1.1 mpk, about 0.2 mpk to about 1.0 mpk, about 0.2 mpk to about 0.9 mpk, about 0.2 to about 0.8 mpk, about 0.2 mpk to about 0.7 mpk, about 0.2 mpk to about 0.6 mpk, about 0.2 mpk to about 0.5 mpk, about 0.3 mpk to about 0.7 mpk, about 0.4 mpk to about 0.8 mpk, about 0.3 mpk to about 1.5 mpk, about 0.4 mpk to about 1.5 mpk, about 0.5 mpk to about 1.5 mpk to about 1.5 mpk, about 0.6 mpk to about 1.5 mpk, about 0.7 mpk to about 1.5 mpk, about 0.8 mpk to about 1.5 mpk, about 0.9 mpk to about 1.5 mpk, about 1.0 mpk to about 1.5 mpk, about 1.2 mpk to about 1.5 mpk, or any range or value therein between.

In some embodiments including more than one dose, any two doses in a multiple-dose regimen may have the same or different dose (*e.g.*, first dose is 1.5 mpk; subsequent doses are 0.5 mpk).

22. Forms of Administration

The polynucleotides, pharmaceutical compositions and formulations of the present disclosure described above can be administered by any route that results in a therapeutically effective outcome, such as intravenous (into a vein) administration. These also include, but are not limited to enteral (into the intestine), gastroenteral, epidural (into the dura matter), oral (by way of the mouth), transdermal, peridural, intracerebral (into the cerebrum), intracerebroventricular (into the cerebral ventricles), epicutaneous (application onto the skin), intradermal, (into the skin itself), subcutaneous (under the skin), nasal administration (through the nose), intravenous bolus, intravenous drip, intraarterial (into an artery), intramuscular (into a muscle), intracardiac (into the heart), intraosseous infusion (into the bone marrow), intrathecal (into the spinal canal), intraperitoneal, (infusion or injection into the peritoneum), intravesical infusion, intravitreal, (through the eye), intracavernous injection (into a pathologic cavity) intracavitary (into the base of the penis), intravaginal administration, intrauterine, extra-amniotic administration, transdermal (diffusion through the intact skin for systemic distribution), transmucosal (diffusion through a mucous membrane), transvaginal, insufflation (snorting), sublingual, sublabial, enema, eye drops (onto the conjunctiva), in ear drops, auricular (in or by way of the ear), buccal (directed toward the cheek), conjunctival, cutaneous, dental (to a tooth or teeth), electro-osmosis, endocervical, endosinusial, endotracheal, extracorporeal, hemodialysis, infiltration, interstitial, intra-abdominal, intra-amniotic, intra-articular, intrabiliary, intrabronchial, intrabursal, intracartilaginous (within a cartilage), intracaudal (within the cauda equine), intracisternal (within the cisterna magna cerebellomedullaris), intracorneal (within the cornea), dental intracornal, intracoronary (within the coronary arteries), intracorporus cavernosum (within the dilatable spaces of the corporus cavernosa of the penis), intradiscal (within a disc), intraductal (within a duct of a gland), intraduodenal (within the duodenum), intradural (within or beneath the dura), intraepidermal (to the epidermis), intraesophageal (to the esophagus), intragastric (within the stomach), intragingival (within the gingivae), intraileal (within the distal portion of the small intestine), intralesional (within or introduced directly to a localized lesion), intraluminal (within a lumen of a tube), intralymphatic (within the lymph),

intramedullary (within the marrow cavity of a bone), intrameningeal (within the meninges), intraocular (within the eye), intraovarian (within the ovary), intrapericardial (within the pericardium), intrapleural (within the pleura), intraprostatic (within the prostate gland), intrapulmonary (within the lungs or its bronchi), intrasinal (within the nasal or periorbital
5 sinuses), intraspinal (within the vertebral column), intrasynovial (within the synovial cavity of a joint), intratendinous (within a tendon), intratesticular (within the testicle), intrathecal (within the cerebrospinal fluid at any level of the cerebrospinal axis), intrathoracic (within the thorax), intratubular (within the tubules of an organ), intratympanic (within the aurus media), intravascular (within a vessel or vessels), intraventricular (within a ventricle), iontophoresis
10 (by means of electric current where ions of soluble salts migrate into the tissues of the body), irrigation (to bathe or flush open wounds or body cavities), laryngeal (directly upon the larynx), nasogastric (through the nose and into the stomach), occlusive dressing technique (topical route administration that is then covered by a dressing that occludes the area), ophthalmic (to the external eye), oropharyngeal (directly to the mouth and pharynx),
15 parenteral, percutaneous, periarticular, peridural, perineural, periodontal, rectal, respiratory (within the respiratory tract by inhaling orally or nasally for local or systemic effect), retrobulbar (behind the pons or behind the eyeball), intramyocardial (entering the myocardium), soft tissue, subarachnoid, subconjunctival, submucosal, topical, transplacental (through or across the placenta), transtracheal (through the wall of the trachea), transtympanic
20 (across or through the tympanic cavity), ureteral (to the ureter), urethral (to the urethra), vaginal, caudal block, diagnostic, nerve block, biliary perfusion, cardiac perfusion, photopheresis or spinal. In specific embodiments, compositions can be administered in a way that allows them cross the blood-brain barrier, vascular barrier, or other epithelial barrier. In some embodiments, a formulation for a route of administration can include at least one
25 inactive ingredient.

23. Definitions

In order that the present disclosure can be more readily understood, certain terms are first defined. As used in this application, except as otherwise expressly provided herein, each
30 of the following terms shall have the meaning set forth below. Additional definitions are set forth throughout the application.

The present disclosure includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The

invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

In this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the context clearly dictates otherwise. The terms "a" (or "an"), as well as the terms "one or more," and "at least one" can be used interchangeably herein. In certain aspects, the term "a" or "an" means "single." In other aspects, the term "a" or "an" includes "two or more" or "multiple."

Furthermore, "and/or" where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. Thus, the term "and/or" as used in a phrase such as "A and/or B" herein is intended to include "A and B," "A or B," "A" (alone), and "B" (alone). Likewise, the term "and/or" as used in a phrase such as "A, B, and/or C" is intended to encompass each of the following aspects: A, B, and C; A, B, or C; A or C; A or B; B or C; A and C; A and B; B and C; A (alone); B (alone); and C (alone).

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure is related. For example, the Concise Dictionary of Biomedicine and Molecular Biology, Juo, Pei-Show, 2nd ed., 2002, CRC Press; The Dictionary of Cell and Molecular Biology, 3rd ed., 1999, Academic Press; and the Oxford Dictionary Of Biochemistry And Molecular Biology, Revised, 2000, Oxford University Press, provide one of skill with a general dictionary of many of the terms used in this disclosure.

Wherever aspects are described herein with the language "comprising," otherwise analogous aspects described in terms of "consisting of" and/or "consisting essentially of" are also provided.

Units, prefixes, and symbols are denoted in their Système International de Unites (SI) accepted form. Numeric ranges are inclusive of the numbers defining the range. Where a range of values is recited, it is to be understood that each intervening integer value, and each fraction thereof, between the recited upper and lower limits of that range is also specifically disclosed, along with each subrange between such values. The upper and lower limits of any range can independently be included in or excluded from the range, and each range where either, neither or both limits are included is also encompassed within the invention. Where a

value is explicitly recited, it is to be understood that values which are about the same quantity or amount as the recited value are also within the scope of the present disclosure. Where a combination is disclosed, each subcombination of the elements of that combination is also specifically disclosed and is within the scope of the present disclosure. Conversely, where different elements or groups of elements are individually disclosed, combinations thereof are also disclosed. Where any element of an invention is disclosed as having a plurality of alternatives, examples of that invention in which each alternative is excluded singly or in any combination with the other alternatives are also hereby disclosed; more than one element of an invention can have such exclusions, and all combinations of elements having such exclusions are hereby disclosed.

Nucleotides are referred to by their commonly accepted single-letter codes. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation. Nucleobases are referred to herein by their commonly known one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Accordingly, A represents adenine, C represents cytosine, G represents guanine, T represents thymine, U represents uracil.

Amino acids are referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Unless otherwise indicated, amino acid sequences are written left to right in amino to carboxy orientation.

About: The term "about" as used in connection with a numerical value throughout the specification and the claims denotes an interval of accuracy, familiar and acceptable to a person skilled in the art, such interval of accuracy is $\pm 10\%$.

Where ranges are given, endpoints are included. Furthermore, unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the present disclosure, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

Administered in combination: As used herein, the term "administered in combination" or "combined administration" means that two or more agents are administered to a subject at the same time or within an interval such that there can be an overlap of an effect of each agent on the patient. In some embodiments, they are administered within about 60, 30, 15, 10,

5, or 1 minute of one another. In some embodiments, the administrations of the agents are spaced sufficiently closely together such that a combinatorial (*e.g.*, a synergistic) effect is achieved.

Amino acid substitution: The term "amino acid substitution" refers to replacing an amino acid residue present in a parent or reference sequence (*e.g.*, a wild type ASL sequence) with another amino acid residue. An amino acid can be substituted in a parent or reference sequence (*e.g.*, a wild type ASL polypeptide sequence), for example, via chemical peptide synthesis or through recombinant methods known in the art. Accordingly, a reference to a "substitution at position X" refers to the substitution of an amino acid present at position X with an alternative amino acid residue. In some aspects, substitution patterns can be described according to the schema AnY, wherein A is the single letter code corresponding to the amino acid naturally or originally present at position n, and Y is the substituting amino acid residue. In other aspects, substitution patterns can be described according to the schema An(YZ), wherein A is the single letter code corresponding to the amino acid residue substituting the amino acid naturally or originally present at position X, and Y and Z are alternative substituting amino acid residue.

In the context of the present disclosure, substitutions (even when they referred to as amino acid substitution) are conducted at the nucleic acid level, *i.e.*, substituting an amino acid residue with an alternative amino acid residue is conducted by substituting the codon encoding the first amino acid with a codon encoding the second amino acid.

Animal: As used herein, the term "animal" refers to any member of the animal kingdom. In some embodiments, "animal" refers to humans at any stage of development. In some embodiments, "animal" refers to non-human animals at any stage of development. In certain embodiments, the non-human animal is a mammal (*e.g.*, a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, and worms. In some embodiments, the animal is a transgenic animal, genetically-engineered animal, or a clone.

Approximately: As used herein, the term "approximately," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" refers to a range of values that fall within 25%, 20%,

19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

5 *Associated with:* As used herein with respect to a disease, the term "associated with" means that the symptom, measurement, characteristic, or status in question is linked to the diagnosis, development, presence, or progression of that disease. As association can, but need not, be causatively linked to the disease. For example, symptoms, sequelae, or any effects causing a decrease in the quality of life of a patient of argininosuccinic aciduria are
10 considered associated with argininosuccinic aciduria and in some embodiments of the present disclosure can be treated, ameliorated, or prevented by administering the polynucleotides of the present disclosure to a subject in need thereof.

When used with respect to two or more moieties, the terms "associated with," "conjugated," "linked," "attached," and "tethered," when used with respect to two or more
15 moieties, means that the moieties are physically associated or connected with one another, either directly or via one or more additional moieties that serves as a linking agent, to form a structure that is sufficiently stable so that the moieties remain physically associated under the conditions in which the structure is used, *e.g.*, physiological conditions. An "association" need not be strictly through direct covalent chemical bonding. It can also suggest ionic or
20 hydrogen bonding or a hybridization based connectivity sufficiently stable such that the "associated" entities remain physically associated.

Biocompatible: As used herein, the term "biocompatible" means compatible with living cells, tissues, organs or systems posing little to no risk of injury, toxicity or rejection by the immune system.

25 *Biodegradable:* As used herein, the term "biodegradable" means capable of being broken down into innocuous products by the action of living things.

Biologically active: As used herein, the phrase "biologically active" refers to a characteristic of any substance that has activity in a biological system and/or organism. For instance, a substance that, when administered to an organism, has a biological effect on that
30 organism, is considered to be biologically active. In particular embodiments, a polynucleotide

of the present disclosure can be considered biologically active if even a portion of the polynucleotide is biologically active or mimics an activity considered biologically relevant.

Chimera: As used herein, "chimera" is an entity having two or more incongruous or heterogeneous parts or regions. For example, a chimeric molecule can comprise a first part comprising an ASL polypeptide, and a second part (*e.g.*, genetically fused to the first part) comprising a second therapeutic protein (*e.g.*, a protein with a distinct enzymatic activity, an antigen binding moiety, or a moiety capable of extending the plasma half life of ASL, for example, an Fc region of an antibody).

Sequence Optimization: The term "sequence optimization" refers to a process or series of processes by which nucleobases in a reference nucleic acid sequence are replaced with alternative nucleobases, resulting in a nucleic acid sequence with improved properties, *e.g.*, improved protein expression or decreased immunogenicity.

In general, the goal in sequence optimization is to produce a synonymous nucleotide sequence than encodes the same polypeptide sequence encoded by the reference nucleotide sequence. Thus, there are no amino acid substitutions (as a result of codon optimization) in the polypeptide encoded by the codon optimized nucleotide sequence with respect to the polypeptide encoded by the reference nucleotide sequence.

Codon substitution: The terms "codon substitution" or "codon replacement" in the context of sequence optimization refer to replacing a codon present in a reference nucleic acid sequence with another codon. A codon can be substituted in a reference nucleic acid sequence, for example, via chemical peptide synthesis or through recombinant methods known in the art. Accordingly, references to a "substitution" or "replacement" at a certain location in a nucleic acid sequence (*e.g.*, an mRNA) or within a certain region or subsequence of a nucleic acid sequence (*e.g.*, an mRNA) refer to the substitution of a codon at such location or region with an alternative codon.

As used herein, the terms "coding region" and "region encoding" and grammatical variants thereof, refer to an Open Reading Frame (ORF) in a polynucleotide that upon expression yields a polypeptide or protein.

Compound: As used herein, the term "compound," is meant to include all stereoisomers and isotopes of the structure depicted. As used herein, the term "stereoisomer"

means any geometric isomer (*e.g.*, *cis*- and *trans*- isomer), enantiomer, or diastereomer of a compound. The present disclosure encompasses any and all stereoisomers of the compounds described herein, including stereomerically pure forms (*e.g.*, geometrically pure, enantiomerically pure, or diastereomerically pure) and enantiomeric and stereoisomeric mixtures, *e.g.*, racemates. Enantiomeric and stereomeric mixtures of compounds and means of resolving them into their component enantiomers or stereoisomers are well-known. “Isotopes” refers to atoms having the same atomic number but different mass numbers resulting from a different number of neutrons in the nuclei. For example, isotopes of hydrogen include tritium and deuterium. Further, a compound, salt, or complex of the present disclosure can be prepared in combination with solvent or water molecules to form solvates and hydrates by routine methods.

Contacting: As used herein, the term “contacting” means establishing a physical connection between two or more entities. For example, contacting a mammalian cell with a nanoparticle composition means that the mammalian cell and a nanoparticle are made to share a physical connection. Methods of contacting cells with external entities both *in vivo* and *ex vivo* are well known in the biological arts. For example, contacting a nanoparticle composition and a mammalian cell disposed within a mammal can be performed by varied routes of administration (*e.g.*, intravenous, intramuscular, intradermal, and subcutaneous) and can involve varied amounts of nanoparticle compositions. Moreover, more than one mammalian cell can be contacted by a nanoparticle composition.

Conservative amino acid substitution: A “conservative amino acid substitution” is one in which the amino acid residue in a protein sequence is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, including basic side chains (*e.g.*, lysine, arginine, or histidine), acidic side chains (*e.g.*, aspartic acid or glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, or cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, or tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, or histidine). Thus, if an amino acid in a polypeptide is replaced with another amino acid from the same side chain family, the amino acid substitution is considered to be conservative. In another aspect, a string of amino acids can be

conservatively replaced with a structurally similar string that differs in order and/or composition of side chain family members.

Non-conservative amino acid substitution: Non-conservative amino acid substitutions include those in which (i) a residue having an electropositive side chain (*e.g.*, Arg, His or Lys) is substituted for, or by, an electronegative residue (*e.g.*, Glu or Asp), (ii) a hydrophilic residue (*e.g.*, Ser or Thr) is substituted for, or by, a hydrophobic residue (*e.g.*, Ala, Leu, Ile, Phe or Val), (iii) a cysteine or proline is substituted for, or by, any other residue, or (iv) a residue having a bulky hydrophobic or aromatic side chain (*e.g.*, Val, His, Ile or Trp) is substituted for, or by, one having a smaller side chain (*e.g.*, Ala or Ser) or no side chain (*e.g.*, Gly).

Other amino acid substitutions can be readily identified by workers of ordinary skill. For example, for the amino acid alanine, a substitution can be taken from any one of D-alanine, glycine, beta-alanine, L-cysteine and D-cysteine. For lysine, a replacement can be any one of D-lysine, arginine, D-arginine, homo-arginine, methionine, D-methionine, ornithine, or D- ornithine. Generally, substitutions in functionally important regions that can be expected to induce changes in the properties of isolated polypeptides are those in which (i) a polar residue, *e.g.*, serine or threonine, is substituted for (or by) a hydrophobic residue, *e.g.*, leucine, isoleucine, phenylalanine, or alanine; (ii) a cysteine residue is substituted for (or by) any other residue; (iii) a residue having an electropositive side chain, *e.g.*, lysine, arginine or histidine, is substituted for (or by) a residue having an electronegative side chain, *e.g.*, glutamic acid or aspartic acid; or (iv) a residue having a bulky side chain, *e.g.*, phenylalanine, is substituted for (or by) one not having such a side chain, *e.g.*, glycine. The likelihood that one of the foregoing non-conservative substitutions can alter functional properties of the protein is also correlated to the position of the substitution with respect to functionally important regions of the protein: some non-conservative substitutions can accordingly have little or no effect on biological properties.

Conserved: As used herein, the term "conserved" refers to nucleotides or amino acid residues of a polynucleotide sequence or polypeptide sequence, respectively, that are those that occur unaltered in the same position of two or more sequences being compared.

Nucleotides or amino acids that are relatively conserved are those that are conserved amongst more related sequences than nucleotides or amino acids appearing elsewhere in the sequences.

In some embodiments, two or more sequences are said to be "completely conserved" if they are 100% identical to one another. In some embodiments, two or more sequences are said to be "highly conserved" if they are at least 70% identical, at least 80% identical, at least 90% identical, or at least 95% identical to one another. In some embodiments, two or more sequences are said to be "highly conserved" if they are about 70% identical, about 80% identical, about 90% identical, about 95%, about 98%, or about 99% identical to one another. In some embodiments, two or more sequences are said to be "conserved" if they are at least 30% identical, at least 40% identical, at least 50% identical, at least 60% identical, at least 70% identical, at least 80% identical, at least 90% identical, or at least 95% identical to one another. In some embodiments, two or more sequences are said to be "conserved" if they are about 30% identical, about 40% identical, about 50% identical, about 60% identical, about 70% identical, about 80% identical, about 90% identical, about 95% identical, about 98% identical, or about 99% identical to one another. Conservation of sequence can apply to the entire length of an polynucleotide or polypeptide or can apply to a portion, region or feature thereof.

Controlled Release: As used herein, the term "controlled release" refers to a pharmaceutical composition or compound release profile that conforms to a particular pattern of release to effect a therapeutic outcome.

Cyclic or Cyclized: As used herein, the term "cyclic" refers to the presence of a continuous loop. Cyclic molecules need not be circular, only joined to form an unbroken chain of subunits. Cyclic molecules such as the engineered RNA or mRNA of the present disclosure can be single units or multimers or comprise one or more components of a complex or higher order structure.

Delivering: As used herein, the term "delivering" means providing an entity to a destination. For example, delivering a polynucleotide to a subject can involve administering a nanoparticle composition including the polynucleotide to the subject (*e.g.*, by an intravenous, intramuscular, intradermal, or subcutaneous route). Administration of a nanoparticle composition to a mammal or mammalian cell can involve contacting one or more cells with the nanoparticle composition.

Delivery Agent: As used herein, "delivery agent" refers to any substance that facilitates, at least in part, the *in vivo*, *in vitro*, or *ex vivo* delivery of a polynucleotide to targeted cells.

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

10 *Dosing regimen:* As used herein, a "dosing regimen" or a "dosing regimen" is a schedule of administration or physician determined regimen of treatment, prophylaxis, or palliative care.

15 *Effective Amount:* As used herein, the term "effective amount" of an agent is that amount sufficient to effect beneficial or desired results, for example, clinical results, and, as such, an "effective amount" depends upon the context in which it is being applied. For example, in the context of administering an agent that treats a protein deficiency (*e.g.*, an ASL deficiency), an effective amount of an agent is, for example, an amount of mRNA expressing sufficient ASL to ameliorate, reduce, eliminate, or prevent the symptoms associated with the ASL deficiency, as compared to the severity of the symptom observed without administration of the agent. The term "effective amount" can be used interchangeably with "effective dose," "therapeutically effective amount," or "therapeutically effective dose."

20 *Encapsulate:* As used herein, the term "encapsulate" means to enclose, surround or encase.

25 *Encapsulation Efficiency:* As used herein, "encapsulation efficiency" refers to the amount of a polynucleotide that becomes part of a nanoparticle composition, relative to the initial total amount of polynucleotide used in the preparation of a nanoparticle composition. For example, if 97 mg of polynucleotide are encapsulated in a nanoparticle composition out of a total 100 mg of polynucleotide initially provided to the composition, the encapsulation efficiency can be given as 97%. As used herein, "encapsulation" can refer to complete, substantial, or partial enclosure, confinement, surrounding, or encasement.

30 *Enhanced Delivery:* As used herein, the term "enhanced delivery" means delivery of more (*e.g.*, at least 1.5 fold more, at least 2-fold more, at least 3-fold more, at least 4-fold

more, at least 5-fold more, at least 6-fold more, at least 7-fold more, at least 8-fold more, at least 9-fold more, at least 10-fold more) of a polynucleotide by a nanoparticle to a target tissue of interest (*e.g.*, mammalian liver) compared to the level of delivery of a polynucleotide by a control nanoparticle to a target tissue of interest (*e.g.*, MC3, KC2, or DLinDMA). The level of delivery of a nanoparticle to a particular tissue can be measured by comparing the amount of protein produced in a tissue to the weight of said tissue, comparing the amount of polynucleotide in a tissue to the weight of said tissue, comparing the amount of protein produced in a tissue to the amount of total protein in said tissue, or comparing the amount of polynucleotide in a tissue to the amount of total polynucleotide in said tissue. It will be understood that the enhanced delivery of a nanoparticle to a target tissue need not be determined in a subject being treated, it can be determined in a surrogate such as an animal model (*e.g.*, a rat model).

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

Formulation: As used herein, a "formulation" includes at least a polynucleotide and one or more of a carrier, an excipient, and a delivery agent.

Fragment: A "fragment," as used herein, refers to a portion. For example, fragments of proteins can comprise polypeptides obtained by digesting full-length protein isolated from cultured cells. In some embodiments, a fragment is a subsequences of a full length protein (*e.g.*, ASL) wherein N-terminal, and/or C-terminal, and/or internal subsequences have been deleted. In some preferred aspects of the present disclosure, the fragments of a protein of the present disclosure are functional fragments.

Functional: As used herein, a "functional" biological molecule is a biological molecule in a form in which it exhibits a property and/or activity by which it is characterized. Thus, a functional fragment of a polynucleotide of the present disclosure is a polynucleotide capable of expressing a functional ASL fragment. As used herein, a functional fragment of ASL refers to a fragment of wild type ASL (*i.e.*, a fragment of any of its naturally occurring isoforms), or a mutant or variant thereof, wherein the fragment retains a least about 10%, at

least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95% of the biological activity of the corresponding full length protein.

The terms "ASL enzymatic activity" and "ASL activity," are used interchangeably in the present disclosure and refer to ASL's ability to catalyze the reversible breakdown of argininosuccinate (ASA) to produce the arginine and dicarboxylic acid fumarate.

Accordingly, a fragment or variant retaining or having ASL enzymatic activity or ASL activity refers to a fragment or variant that has measurable enzymatic activity in catalyzing the reversible breakdown of argininosuccinate (ASA) to produce the arginine and dicarboxylic acid fumarate. Therefore, a fragment or variant retaining or having ASL enzymatic activity or ASL activity refers to a fragment or variant that has measurable enzymatic activity in converting ammonia to urea.

Helper Lipid: As used herein, the term "helper lipid" refers to a compound or molecule that includes a lipidic moiety (for insertion into a lipid layer, *e.g.*, lipid bilayer) and a polar moiety (for interaction with physiologic solution at the surface of the lipid layer). Typically the helper lipid is a phospholipid. A function of the helper lipid is to "complement" the amino lipid and increase the fusogenicity of the bilayer and/or to help facilitate endosomal escape, *e.g.*, of nucleic acid delivered to cells. Helper lipids are also believed to be a key structural component to the surface of the LNP.

Homology: As used herein, the term "homology" refers to the overall relatedness between polymeric molecules, *e.g.* between nucleic acid molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Generally, the term "homology" implies an evolutionary relationship between two molecules. Thus, two molecules that are homologous will have a common evolutionary ancestor. In the context of the present disclosure, the term homology encompasses both to identity and similarity.

In some embodiments, polymeric molecules are considered to be "homologous" to one another if at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% of the monomers in the molecule are identical (exactly the same monomer) or are similar (conservative substitutions). The term "homologous" necessarily

refers to a comparison between at least two sequences (polynucleotide or polypeptide sequences).

Identity: As used herein, the term "identity" refers to the overall monomer conservation between polymeric molecules, *e.g.*, between polynucleotide molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of the percent identity of two polynucleotide sequences, for example, can be performed by aligning the two sequences for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second nucleic acid sequences for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In certain embodiments, the length of a sequence aligned for comparison purposes is at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or 100% of the length of the reference sequence. The nucleotides at corresponding nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which needs to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. When comparing DNA and RNA, thymine (T) and uracil (U) can be considered equivalent.

Suitable software programs are available from various sources, and for alignment of both protein and nucleotide sequences. One suitable program to determine percent sequence identity is *bl2seq*, part of the BLAST suite of program available from the U.S. government's National Center for Biotechnology Information BLAST web site (blast.ncbi.nlm.nih.gov). *Bl2seq* performs a comparison between two sequences using either the BLASTN or BLASTP algorithm. BLASTN is used to compare nucleic acid sequences, while BLASTP is used to compare amino acid sequences. Other suitable programs are, *e.g.*, Needle, Stretcher, Water, or Matcher, part of the EMBOSS suite of bioinformatics programs and also available from the European Bioinformatics Institute (EBI) at www.ebi.ac.uk/Tools/psa.

Sequence alignments can be conducted using methods known in the art such as MAFFT, Clustal (ClustalW, Clustal X or Clustal Omega), MUSCLE, etc.

Different regions within a single polynucleotide or polypeptide target sequence that aligns with a polynucleotide or polypeptide reference sequence can each have their own percent sequence identity. It is noted that the percent sequence identity value is rounded to the nearest tenth. For example, 80.11, 80.12, 80.13, and 80.14 are rounded down to 80.1, while 80.15, 80.16, 80.17, 80.18, and 80.19 are rounded up to 80.2. It also is noted that the length value will always be an integer.

In certain aspects, the percentage identity "%ID" of a first amino acid sequence (or nucleic acid sequence) to a second amino acid sequence (or nucleic acid sequence) is calculated as $\%ID = 100 \times (Y/Z)$, where Y is the number of amino acid residues (or nucleobases) scored as identical matches in the alignment of the first and second sequences (as aligned by visual inspection or a particular sequence alignment program) and Z is the total number of residues in the second sequence. If the length of a first sequence is longer than the second sequence, the percent identity of the first sequence to the second sequence will be higher than the percent identity of the second sequence to the first sequence.

One skilled in the art will appreciate that the generation of a sequence alignment for the calculation of a percent sequence identity is not limited to binary sequence-sequence comparisons exclusively driven by primary sequence data. It will also be appreciated that sequence alignments can be generated by integrating sequence data with data from heterogeneous sources such as structural data (*e.g.*, crystallographic protein structures), functional data (*e.g.*, location of mutations), or phylogenetic data. A suitable program that integrates heterogeneous data to generate a multiple sequence alignment is T-Coffee, available at www.tcoffee.org, and alternatively available, *e.g.*, from the EBI. It will also be appreciated that the final alignment used to calculate percent sequence identity can be curated either automatically or manually.

Insertional and deletional variants: "Insertional variants" when referring to polypeptides are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in a native or starting sequence. "Immediately adjacent" to an amino acid means connected to either the alpha-carboxy or alpha-amino functional group of the amino acid. "Deletional variants" when referring to polypeptides are those with one or more amino acids in the native or starting amino acid sequence removed. Ordinarily, deletional variants will have one or more amino acids deleted in a particular region of the molecule.

Intact: As used herein, in the context of a polypeptide, the term "intact" means retaining an amino acid corresponding to the wild type protein, *e.g.*, not mutating or substituting the wild type amino acid. Conversely, in the context of a nucleic acid, the term "intact" means retaining a nucleobase corresponding to the wild type nucleic acid, *e.g.*, not mutating or substituting the wild type nucleobase.

Ionizable amino lipid: The term "ionizable amino lipid" includes those lipids having one, two, three, or more fatty acid or fatty alkyl chains and a pH-titratable amino head group (*e.g.*, an alkylamino or dialkylamino head group). An ionizable amino lipid is typically protonated (*i.e.*, positively charged) at a pH below the pKa of the amino head group and is substantially not charged at a pH above the pKa. Such ionizable amino lipids include, but are not limited to DLin-MC3-DMA (MC3), (13Z,165Z)-N,N-dimethyl-3-nonydocosa-13-16-dien-1-amine (L608), and a compound of any one of Formula I, II, and II described herein (*e.g.*, any one of Compound I-1, Compound I-2, Compound I-3, or Compound I-VI).

Linker: As used herein, a "linker" refers to a group of atoms, *e.g.*, 10-1,000 atoms, and can be comprised of the atoms or groups such as, but not limited to, carbon, amino, alkylamino, oxygen, sulfur, sulfoxide, sulfonyl, carbonyl, and imine. The linker can be attached to a modified nucleoside or nucleotide on the nucleobase or sugar moiety at a first end, and to a payload, *e.g.*, a detectable or therapeutic agent, at a second end. The linker can be of sufficient length as to not interfere with incorporation into a nucleic acid sequence. The linker can be used for any useful purpose, such as to form polynucleotide multimers (*e.g.*, through linkage of two or more chimeric polynucleotides molecules or IVT polynucleotides) or polynucleotides conjugates, as well as to administer a payload, as described herein. Examples of chemical groups that can be incorporated into the linker include, but are not limited to, alkyl, alkenyl, alkynyl, amido, amino, ether, thioether, ester, alkylene, heteroalkylene, aryl, or heterocyclyl, each of which can be optionally substituted, as described herein. Examples of linkers include, but are not limited to, unsaturated alkanes, polyethylene glycols (*e.g.*, ethylene or propylene glycol monomeric units, *e.g.*, diethylene glycol, dipropylene glycol, triethylene glycol, tripropylene glycol, tetraethylene glycol, or tetraethylene glycol), and dextran polymers and derivatives thereof. Other examples include, but are not limited to, cleavable moieties within the linker, such as, for example, a disulfide bond (-S-S-) or an azo bond (-N=N-), which can be cleaved using a reducing agent or photolysis. Non-limiting examples of a selectively cleavable bond include an amido bond can

be cleaved for example by the use of tris(2-carboxyethyl)phosphine (TCEP), or other reducing agents, and/or photolysis, as well as an ester bond can be cleaved for example by acidic or basic hydrolysis.

Methods of Administration: As used herein, “methods of administration” can include
5 intravenous, intramuscular, intradermal, subcutaneous, or other methods of delivering a composition to a subject. A method of administration can be selected to target delivery (*e.g.*, to specifically deliver) to a specific region or system of a body.

Modified: As used herein "modified" refers to a changed state or structure of a molecule of the present disclosure. Molecules can be modified in many ways including
10 chemically, structurally, and functionally. In some embodiments, the mRNA molecules of the present disclosure are modified by the introduction of non-natural nucleosides and/or nucleotides, *e.g.*, as it relates to the natural ribonucleotides A, U, G, and C. Noncanonical nucleotides such as the cap structures are not considered "modified" although they differ from the chemical structure of the A, C, G, U ribonucleotides.

Nanoparticle Composition: As used herein, a “nanoparticle composition” is a composition comprising one or more lipids. Nanoparticle compositions are typically sized on the order of micrometers or smaller and can include a lipid bilayer. Nanoparticle compositions encompass lipid nanoparticles (LNPs), liposomes (*e.g.*, lipid vesicles), and lipoplexes. For example, a nanoparticle composition can be a liposome having a lipid bilayer
15 with a diameter of 500 nm or less.
20

Naturally occurring: As used herein, "naturally occurring" means existing in nature without artificial aid.

Nucleic acid sequence: The terms "nucleic acid sequence," "nucleotide sequence," or "polynucleotide sequence" are used interchangeably and refer to a contiguous nucleic acid
25 sequence. The sequence can be either single stranded or double stranded DNA or RNA, *e.g.*, an mRNA.

The term "nucleic acid," in its broadest sense, includes any compound and/or substance that comprises a polymer of nucleotides. These polymers are often referred to as polynucleotides. Exemplary nucleic acids or polynucleotides of the present disclosure
30 include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs),

threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs),
locked nucleic acids (LNAs, including LNA having a β -D-ribo configuration, α -LNA having
an α -L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino
functionalization, and 2'-amino- α -LNA having a 2'-amino functionalization), ethylene
5 nucleic acids (ENA), cyclohexenyl nucleic acids (CeNA) or hybrids or combinations thereof.

The phrase "nucleotide sequence encoding" refers to the nucleic acid (*e.g.*, an mRNA
or DNA molecule) coding sequence which encodes a polypeptide. The coding sequence can
further include initiation and termination signals operably linked to regulatory elements
including a promoter and polyadenylation signal capable of directing expression in the cells
10 of an individual or mammal to which the nucleic acid is administered. The coding sequence
can further include sequences that encode signal peptides.

Operably linked: As used herein, the phrase "operably linked" refers to a functional
connection between two or more molecules, constructs, transcripts, entities, moieties or the
like.

15 *Optionally substituted:* Herein a phrase of the form "optionally substituted X" (*e.g.*,
optionally substituted alkyl) is intended to be equivalent to "X, wherein X is optionally
substituted" (*e.g.*, "alkyl, wherein said alkyl is optionally substituted"). It is not intended to
mean that the feature "X" (*e.g.*, alkyl) *per se* is optional.

Part: As used herein, a "part" or "region" of a polynucleotide is defined as any portion
20 of the polynucleotide that is less than the entire length of the polynucleotide.

Patient: As used herein, "patient" refers to a subject who can seek or be in need of
treatment, requires treatment, is receiving treatment, will receive treatment, or a subject who
is under care by a trained professional for a particular disease or condition. In some
embodiments, the treatment is needed, required, or received to prevent or decrease the risk of
25 developing acute disease, *i.e.*, it is a prophylactic treatment.

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

Pharmaceutically acceptable excipients: The phrase "pharmaceutically acceptable excipient," as used herein, refers any ingredient other than the compounds described herein (for example, a vehicle capable of suspending or dissolving the active compound) and having the properties of being substantially nontoxic and non-inflammatory in a patient. Excipients

5 can include, for example: antiadherents, 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, suspension or dispersing agents, sweeteners, and waters of hydration.

Exemplary excipients include, but are not limited to: butylated hydroxytoluene (BHT),

10 calcium carbonate, calcium phosphate (dibasic), calcium stearate, croscarmellose, crosslinked polyvinyl pyrrolidone, citric acid, crospovidone, cysteine, ethylcellulose, gelatin, hydroxypropyl cellulose, hydroxypropyl methylcellulose, lactose, magnesium stearate, maltitol, mannitol, methionine, methylcellulose, methyl paraben, microcrystalline cellulose, polyethylene glycol, polyvinyl pyrrolidone, povidone, pregelatinized starch, propyl paraben,

15 retinyl palmitate, shellac, silicon dioxide, sodium carboxymethyl cellulose, sodium citrate, sodium starch glycolate, sorbitol, starch (corn), stearic acid, sucrose, talc, titanium dioxide, vitamin A, vitamin E, vitamin C, and xylitol.

Pharmaceutically acceptable salts: The present disclosure also includes pharmaceutically acceptable salts of the compounds described herein. As used herein,

20 "pharmaceutically acceptable salts" refers to derivatives of the disclosed compounds wherein the parent compound is modified by converting an existing acid or base moiety to its salt form (e.g., by reacting the free base group with a suitable organic acid). Examples of pharmaceutically acceptable salts include, but are not limited to, mineral or organic acid salts of basic residues such as amines; alkali or organic salts of acidic residues such as carboxylic

25 acids; and the like. Representative acid addition salts include acetate, acetic acid, adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzene sulfonic acid, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, fumarate, glucoheptonate, glycerophosphate, hemisulfate, heptonate, hexanoate, hydrobromide, hydrochloride, hydroiodide, 2-hydroxy-ethanesulfonate,

30 lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3-phenylpropionate, phosphate, picrate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, toluenesulfonate, undecanoate, valerate salts, and the like.

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

Pharmaceutically acceptable solvate: The term "pharmaceutically acceptable solvate," as used herein, means a compound of the present disclosure wherein molecules of a suitable solvent are incorporated in the crystal lattice. A suitable solvent is physiologically tolerable at the dosage administered. For example, solvates can be prepared by crystallization, recrystallization, or precipitation from a solution that includes organic solvents, water, or a mixture thereof. Examples of suitable solvents are ethanol, water (for example, mono-, di-, and tri-hydrates), *N*-methylpyrrolidinone (NMP), dimethyl sulfoxide (DMSO), *N,N'*-dimethylformamide (DMF), *N,N'*-dimethylacetamide (DMAC), 1,3-dimethyl-2-imidazolidinone (DMEU), 1,3-dimethyl-3,4,5,6-tetrahydro-2-(1H)-pyrimidinone (DMPU), acetonitrile (ACN), propylene glycol, ethyl acetate, benzyl alcohol, 2-pyrrolidone, benzyl benzoate, and the like. When water is the solvent, the solvate is referred to as a "hydrate."

Pharmacokinetic: As used herein, "pharmacokinetic" refers to any one or more properties of a molecule or compound as it relates to the determination of the fate of substances administered to a living organism. Pharmacokinetics is divided into several areas including the extent and rate of absorption, distribution, metabolism and excretion. This is

commonly referred to as ADME where: (A) Absorption is the process of a substance entering the blood circulation; (D) Distribution is the dispersion or dissemination of substances throughout the fluids and tissues of the body; (M) Metabolism (or Biotransformation) is the irreversible transformation of parent compounds into daughter metabolites; and (E) Excretion (or Elimination) refers to the elimination of the substances from the body. In rare cases, some drugs irreversibly accumulate in body tissue.

Polynucleotide: The term "polynucleotide" as used herein refers to polymers of nucleotides of any length, including ribonucleotides, deoxyribonucleotides, analogs thereof, or mixtures thereof. This term refers to the primary structure of the molecule. Thus, the term includes triple-, double- and single-stranded deoxyribonucleic acid ("DNA"), as well as triple-, double- and single-stranded ribonucleic acid ("RNA"). It also includes modified, for example by alkylation, and/or by capping, and unmodified forms of the polynucleotide. More particularly, the term "polynucleotide" includes polydeoxyribonucleotides (containing 2-deoxy-D-ribose), polyribonucleotides (containing D-ribose), including tRNA, rRNA, hRNA, siRNA and mRNA, whether spliced or unspliced, any other type of polynucleotide which is an N- or C-glycoside of a purine or pyrimidine base, and other polymers containing nonnucleotidic backbones, for example, polyamide (*e.g.*, peptide nucleic acids "PNAs") and polymorpholino polymers, and other synthetic sequence-specific nucleic acid polymers providing that the polymers contain nucleobases in a configuration which allows for base pairing and base stacking, such as is found in DNA and RNA. In particular aspects, the polynucleotide comprises an mRNA. In other aspect, the mRNA is a synthetic mRNA. In some aspects, the synthetic mRNA comprises at least one unnatural nucleobase. In some aspects, all nucleobases of a certain class have been replaced with unnatural nucleobases (*e.g.*, all uridines in a polynucleotide disclosed herein can be replaced with an unnatural nucleobase, *e.g.*, 5-methoxyuridine). In some aspects, the polynucleotide (*e.g.*, a synthetic RNA or a synthetic DNA) comprises only natural nucleobases, *i.e.*, A (adenosine), G (guanosine), C (cytidine), and T (thymidine) in the case of a synthetic DNA, or A, C, G, and U (uridine) in the case of a synthetic RNA.

The skilled artisan will appreciate that the T bases in the codon maps disclosed herein are present in DNA, whereas the T bases would be replaced by U bases in corresponding RNAs. For example, a codon-nucleotide sequence disclosed herein in DNA form, *e.g.*, a vector or an in-vitro translation (IVT) template, would have its T bases transcribed as U

based in its corresponding transcribed mRNA. In this respect, both codon-optimized DNA sequences (comprising T) and their corresponding mRNA sequences (comprising U) are considered codon-optimized nucleotide sequence of the present disclosure. A skilled artisan would also understand that equivalent codon-maps can be generated by replaced one or more
5 bases with non-natural bases. Thus, *e.g.*, a TTC codon (DNA map) would correspond to a UUC codon (RNA map), which in turn would correspond to a ΨΨC codon (RNA map in which U has been replaced with pseudouridine).

Standard A-T and G-C base pairs form under conditions which allow the formation of hydrogen bonds between the N3-H and C4-oxy of thymidine and the N1 and C6-NH₂,
10 respectively, of adenosine and between the C2-oxy, N3 and C4-NH₂, of cytidine and the C2-NH₂, N'—H and C6-oxy, respectively, of guanosine. Thus, for example, guanosine (2-amino-6-oxy-9-β-D-ribofuranosyl-purine) can be modified to form isoguanosine (2-oxy-6-amino-9-β-D-ribofuranosyl-purine). Such modification results in a nucleoside base which will no longer effectively form a standard base pair with cytosine. However, modification of
15 cytosine (1-β-D-ribofuranosyl-2-oxy-4-amino-pyrimidine) to form isocytosine (1-β-D-ribofuranosyl-2-amino-4-oxy-pyrimidine-) results in a modified nucleotide which will not effectively base pair with guanosine but will form a base pair with isoguanosine (U.S. Pat. No. 5,681,702 to Collins et al.). Isocytosine is available from Sigma Chemical Co. (St. Louis, Mo.); isocytidine can be prepared by the method described by Switzer et al. (1993)
20 Biochemistry 32:10489-10496 and references cited therein; 2'-deoxy-5-methyl-isocytidine can be prepared by the method of Tor et al., 1993, J. Am. Chem. Soc. 115:4461-4467 and references cited therein; and isoguanine nucleotides can be prepared using the method described by Switzer et al., 1993, *supra*, and Mantsch et al., 1993, Biochem. 14:5593-5601, or by the method described in U.S. Pat. No. 5,780,610 to Collins et al. Other nonnatural base
25 pairs can be synthesized by the method described in Piccirilli et al., 1990, Nature 343:33-37, for the synthesis of 2,6-diaminopyrimidine and its complement (1-methylpyrazolo-[4,3]pyrimidine-5,7-(4H,6H)-dione. Other such modified nucleotide units which form unique base pairs are known, such as those described in Leach et al. (1992) J. Am. Chem. Soc. 114:3675-3683 and Switzer et al., *supra*.

30 *Polypeptide*: The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer can comprise modified amino acids. The terms also encompass an amino acid polymer that has

been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid
5 (including, for example, unnatural amino acids such as homocysteine, ornithine, p-acetylphenylalanine, D-amino acids, and creatine), as well as other modifications known in the art.

The term, as used herein, refers to proteins, polypeptides, and peptides of any size, structure, or function. Polypeptides include encoded polynucleotide products, naturally
10 occurring polypeptides, synthetic polypeptides, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing. A polypeptide can be a monomer or can be a multi-molecular complex such as a dimer, trimer or tetramer. They can also comprise single chain or multichain polypeptides. Most commonly disulfide linkages are found in multichain polypeptides. The term polypeptide can also apply to amino acid
15 polymers in which one or more amino acid residues are an artificial chemical analogue of a corresponding naturally occurring amino acid. In some embodiments, a "peptide" can be less than or equal to 50 amino acids long, *e.g.*, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 amino acids long.

Polypeptide variant: As used herein, the term "polypeptide variant" refers to
20 molecules that differ in their amino acid sequence from a native or reference sequence. The amino acid sequence variants can possess substitutions, deletions, and/or insertions at certain positions within the amino acid sequence, as compared to a native or reference sequence. Ordinarily, variants will possess at least about 50% identity, at least about 60% identity, at least about 70% identity, at least about 80% identity, at least about 90% identity, at least
25 about 95% identity, at least about 99% identity to a native or reference sequence. In some embodiments, they will be at least about 80%, or at least about 90% identical to a native sequence

Preventing: As used herein, the term "preventing" refers to partially or completely delaying onset of an infection, disease, disorder and/or condition; partially or completely
30 delaying onset of one or more symptoms, features, or clinical manifestations of a particular infection, disease, disorder, and/or condition; partially or completely delaying onset of one or more symptoms, features, or manifestations of a particular infection, disease, disorder, and/or

condition; partially or completely delaying progression from an infection, a particular disease, disorder and/or condition; and/or decreasing the risk of developing pathology associated with the infection, the disease, disorder, and/or condition.

5 *Prophylactic*: As used herein, "prophylactic" refers to a therapeutic or course of action used to prevent the spread of disease.

Prophylaxis: As used herein, a "prophylaxis" refers to a measure taken to maintain health and prevent the spread of disease. An "immune prophylaxis" refers to a measure to produce active or passive immunity to prevent the spread of disease.

10 *Pseudouridine*: As used herein, pseudouridine (ψ) refers to the C-glycoside isomer of the nucleoside uridine. A "pseudouridine analog" is any modification, variant, isoform or derivative of pseudouridine. For example, pseudouridine analogs include but are not limited to 1-carboxymethyl-pseudouridine, 1-propynyl-pseudouridine, 1-taurinomethyl-pseudouridine, 1-taurinomethyl-4-thio-pseudouridine, 1-methylpseudouridine ($m^1\psi$) (also known as N1-methyl-pseudouridine), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydropseudouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\psi$), and 2'-O-methyl-pseudouridine (ψm).

20 *Purified*: As used herein, "purify," "purified," "purification" means to make substantially pure or clear from unwanted components, material defilement, admixture or imperfection.

25 *Reference Nucleic Acid Sequence*: The term "reference nucleic acid sequence" or "reference nucleic acid" or "reference nucleotide sequence" or "reference sequence" refers to a starting nucleic acid sequence (*e.g.*, a RNA, *e.g.*, an mRNA sequence) that can be sequence optimized. In some embodiments, the reference nucleic acid sequence is a wild type nucleic acid sequence, a fragment or a variant thereof. In some embodiments, the reference nucleic acid sequence is a previously sequence optimized nucleic acid sequence.

30 *Salts*: In some aspects, the pharmaceutical composition for delivery disclosed herein and comprises salts of some of their lipid constituents. The term "salt" includes any anionic

and cationic complex. Non-limiting examples of anions include inorganic and organic anions, *e.g.*, fluoride, chloride, bromide, iodide, oxalate (*e.g.*, hemioxalate), phosphate, phosphonate, hydrogen phosphate, dihydrogen phosphate, oxide, carbonate, bicarbonate, nitrate, nitrite, nitride, bisulfite, sulfide, sulfite, bisulfate, sulfate, thiosulfate, hydrogen sulfate, borate, formate, acetate, benzoate, citrate, tartrate, lactate, acrylate, polyacrylate, fumarate, maleate, itaconate, glycolate, gluconate, malate, mandelate, tiglate, ascorbate, salicylate, polymethacrylate, perchlorate, chlorate, chlorite, hypochlorite, bromate, hypobromite, iodate, an alkylsulfonate, an arylsulfonate, arsenate, arsenite, chromate, dichromate, cyanide, cyanate, thiocyanate, hydroxide, peroxide, permanganate, and mixtures thereof.

Sample: As used herein, the term "sample" or "biological sample" refers to a subset of its tissues, cells or component parts (*e.g.*, body fluids, including but not limited to blood, mucus, lymphatic fluid, synovial fluid, cerebrospinal fluid, saliva, amniotic fluid, amniotic cord blood, urine, vaginal fluid and semen). A sample further can include a homogenate, lysate or extract prepared from a whole organism or a subset of its tissues, cells or component parts, or a fraction or portion thereof, including but not limited to, for example, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, blood cells, tumors, organs. A sample further refers to a medium, such as a nutrient broth or gel, which can contain cellular components, such as proteins or nucleic acid molecule.

Signal Sequence: As used herein, the phrases "signal sequence," "signal peptide," and "transit peptide" are used interchangeably and refer to a sequence that can direct the transport or localization of a protein to a certain organelle, cell compartment, or extracellular export. The term encompasses both the signal sequence polypeptide and the nucleic acid sequence encoding the signal sequence. Thus, references to a signal sequence in the context of a nucleic acid refer in fact to the nucleic acid sequence encoding the signal sequence polypeptide.

Similarity: As used herein, the term "similarity" refers to the overall relatedness between polymeric molecules, *e.g.* between polynucleotide molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of percent similarity of polymeric molecules to one another can be performed in the same manner as a calculation of percent identity, except that calculation of percent similarity takes into account conservative substitutions as is understood in the art.

Single unit dose: As used herein, a "single unit dose" is a dose of any therapeutic administered in one dose/at one time/single route/single point of contact, *i.e.*, single administration event.

Specific delivery: As used herein, the term "specific delivery," "specifically deliver," or "specifically delivering" means delivery of more (*e.g.*, at least 1.5 fold more, at least 2-fold more, at least 3-fold more, at least 4-fold more, at least 5-fold more, at least 6-fold more, at least 7-fold more, at least 8-fold more, at least 9-fold more, at least 10-fold more) of a polynucleotide by a nanoparticle to a target tissue of interest (*e.g.*, mammalian liver) compared to an off-target tissue (*e.g.*, mammalian spleen). The level of delivery of a nanoparticle to a particular tissue can be measured by comparing the amount of protein produced in a tissue to the weight of said tissue, comparing the amount of polynucleotide in a tissue to the weight of said tissue, comparing the amount of protein produced in a tissue to the amount of total protein in said tissue, or comparing the amount of polynucleotide in a tissue to the amount of total polynucleotide in said tissue. For example, for renovascular targeting, a polynucleotide is specifically provided to a mammalian kidney as compared to the liver and spleen if 1.5, 2-fold, 3-fold, 5-fold, 10-fold, 15 fold, or 20 fold more polynucleotide per 1 g of tissue is delivered to a kidney compared to that delivered to the liver or spleen following systemic administration of the polynucleotide. It will be understood that the ability of a nanoparticle to specifically deliver to a target tissue need not be determined in a subject being treated, it can be determined in a surrogate such as an animal model (*e.g.*, a rat model).

Stable: As used herein "stable" refers to a compound that is sufficiently robust to survive isolation to a useful degree of purity from a reaction mixture, and in some cases capable of formulation into an efficacious therapeutic agent.

Stabilized: As used herein, the term "stabilize," "stabilized," "stabilized region" means to make or become stable.

Subject: By "subject" or "individual" or "animal" or "patient" or "mammal," is meant any subject, particularly a mammalian subject, for whom diagnosis, prognosis, or therapy is desired. Mammalian subjects include, but are not limited to, humans, domestic animals, farm animals, zoo animals, sport animals, pet animals such as dogs, cats, guinea pigs, rabbits, rats, mice, horses, cattle, cows; primates such as apes, monkeys, orangutans, and chimpanzees; canids such as dogs and wolves; felids such as cats, lions, and tigers; equids such as horses,

donkeys, and zebras; bears, food animals such as cows, pigs, and sheep; ungulates such as deer and giraffes; rodents such as mice, rats, hamsters and guinea pigs; and so on. In certain embodiments, the mammal is a human subject. In other embodiments, a subject is a human patient. In a particular embodiment, a subject is a human patient in need of treatment.

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

Substantially equal: As used herein as it relates to time differences between doses, the term means plus/minus 2%.

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

Susceptible to: An individual who is "susceptible to" a disease, disorder, and/or condition has not been diagnosed with and/or cannot exhibit symptoms of the disease, disorder, and/or condition but harbors a propensity to develop a disease or its symptoms. In
20 some embodiments, an individual who is susceptible to a disease, disorder, and/or condition (for example, argininosuccinic aciduria) can be characterized by one or more of the following: (1) a genetic mutation associated with development of the disease, disorder, and/or condition; (2) a genetic polymorphism associated with development of the disease, disorder, and/or condition; (3) increased and/or decreased expression and/or activity of a protein and/or
25 nucleic acid associated with the disease, disorder, and/or condition; (4) habits and/or lifestyles associated with development of the disease, disorder, and/or condition; (5) a family history of the disease, disorder, and/or condition; and (6) exposure to and/or infection with a microbe associated with development of the disease, disorder, and/or condition. In some
30 embodiments, an individual who is susceptible to a disease, disorder, and/or condition will develop the disease, disorder, and/or condition. In some embodiments, an individual who is

susceptible to a disease, disorder, and/or condition will not develop the disease, disorder, and/or condition.

Sustained release: As used herein, the term "sustained release" refers to a pharmaceutical composition or compound release profile that conforms to a release rate over a specific period of time.

Synthetic: The term "synthetic" means produced, prepared, and/or manufactured by the hand of man. Synthesis of polynucleotides or other molecules of the present disclosure can be chemical or enzymatic.

Targeted Cells: As used herein, "targeted cells" refers to any one or more cells of interest. The cells can be found *in vitro*, *in vivo*, *in situ* or in the tissue or organ of an organism. The organism can be an animal, for example a mammal, a human, a subject or a patient.

Target tissue: As used herein "target tissue" refers to any one or more tissue types of interest in which the delivery of a polynucleotide would result in a desired biological and/or pharmacological effect. Examples of target tissues of interest include specific tissues, organs, and systems or groups thereof. In particular applications, a target tissue can be a liver, a kidney, a lung, a spleen, or a vascular endothelium in vessels (*e.g.*, intra-coronary or intra-femoral). An "off-target tissue" refers to any one or more tissue types in which the expression of the encoded protein does not result in a desired biological and/or pharmacological effect.

The presence of a therapeutic agent in an off-target issue can be the result of: (i) leakage of a polynucleotide from the administration site to peripheral tissue or distant off-target tissue via diffusion or through the bloodstream (*e.g.*, a polynucleotide intended to express a polypeptide in a certain tissue would reach the off-target tissue and the polypeptide would be expressed in the off-target tissue); or (ii) leakage of an polypeptide after administration of a polynucleotide encoding such polypeptide to peripheral tissue or distant off-target tissue via diffusion or through the bloodstream (*e.g.*, a polynucleotide would expressed a polypeptide in the target tissue, and the polypeptide would diffuse to peripheral tissue).

Targeting sequence: As used herein, the phrase "targeting sequence" refers to a sequence that can direct the transport or localization of a protein or polypeptide.

Therapeutic Agent: The term "therapeutic agent" refers to an agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect. For example, in some embodiments, an mRNA encoding an ASL polypeptide can be a therapeutic agent.

Therapeutically effective amount: As used herein, the term "therapeutically effective amount" means an amount of an agent to be delivered (*e.g.*, nucleic acid, drug, therapeutic agent, diagnostic agent, prophylactic agent, etc.) that is sufficient, when administered to a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition.

Therapeutically effective outcome: As used herein, the term "therapeutically effective outcome" means an outcome that is sufficient in a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition.

Transcription: As used herein, the term "transcription" refers to methods to produce mRNA (*e.g.*, an mRNA sequence or template) from DNA (*e.g.*, a DNA template or sequence).

Transfection: As used herein, "transfection" refers to the introduction of a polynucleotide (*e.g.*, exogenous nucleic acids) into a cell wherein a polypeptide encoded by the polynucleotide is expressed (*e.g.*, mRNA) or the polypeptide modulates a cellular function (*e.g.*, siRNA, miRNA). As used herein, "expression" of a nucleic acid sequence refers to translation of a polynucleotide (*e.g.*, an mRNA) into a polypeptide or protein and/or post-translational modification of a polypeptide or protein. Methods of transfection include, but are not limited to, chemical methods, physical treatments and cationic lipids or mixtures.

Treating, treatment, therapy: As used herein, the term "treating" or "treatment" or "therapy" refers to partially or completely alleviating, ameliorating, improving, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of a disease, *e.g.*, argininosuccinic aciduria. For

example, "treating" argininosuccinic aciduria can refer to diminishing symptoms associate with the disease, prolong the lifespan (increase the survival rate) of patients, reducing the severity of the disease, preventing or delaying the onset of the disease, etc. Treatment can be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition and/or to a subject who exhibits only early signs of a disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

Unmodified: As used herein, "unmodified" refers to any substance, compound or molecule prior to being changed in some way. Unmodified can, but does not always, refer to the wild type or native form of a biomolecule. Molecules can undergo a series of modifications whereby each modified molecule can serve as the "unmodified" starting molecule for a subsequent modification.

Uracil: Uracil is one of the four nucleobases in the nucleic acid of RNA, and it is represented by the letter U. Uracil can be attached to a ribose ring, or more specifically, a ribofuranose via a β -N₁-glycosidic bond to yield the nucleoside uridine. The nucleoside uridine is also commonly abbreviated according to the one letter code of its nucleobase, *i.e.*, U. Thus, in the context of the present disclosure, when a monomer in a polynucleotide sequence is U, such U is designated interchangeably as a "uracil" or a "uridine."

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

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

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

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

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

Variant: The term variant as used in present disclosure refers to both natural variants (*e.g.*, polymorphisms, isoforms, etc.) and artificial variants in which at least one amino acid residue in a native or starting sequence (*e.g.*, a wild type sequence) has been removed and a different amino acid inserted in its place at the same position. These variants can be described

as "substitutional variants." The substitutions can be single, where only one amino acid in the molecule has been substituted, or they can be multiple, where two or more amino acids have been substituted in the same molecule. If amino acids are inserted or deleted, the resulting variant would be an "insertional variant" or a "deletional variant" respectively.

5 *Initiation Codon:* As used herein, the term "initiation codon", used interchangeably with the term "start codon", refers to the first codon of an open reading frame that is translated by the ribosome and is comprised of a triplet of linked adenine-uracil-guanine nucleobases. The initiation codon is depicted by the first letter codes of adenine (A), uracil (U), and guanine (G) and is often written simply as "AUG". Although natural mRNAs may
10 use codons other than AUG as the initiation codon, which are referred to herein as "alternative initiation codons", the initiation codons of polynucleotides described herein use the AUG codon. During the process of translation initiation, the sequence comprising the initiation codon is recognized via complementary base-pairing to the anticodon of an initiator tRNA (Met-tRNA^{Met}) bound by the ribosome. Open reading frames may contain more than
15 one AUG initiation codon, which are referred to herein as "alternate initiation codons".

The initiation codon plays a critical role in translation initiation. The initiation codon is the first codon of an open reading frame that is translated by the ribosome. Typically, the initiation codon comprises the nucleotide triplet AUG, however, in some instances translation initiation can occur at other codons comprised of distinct nucleotides. The initiation of
20 translation in eukaryotes is a multistep biochemical process that involves numerous protein-protein, protein-RNA, and RNA-RNA interactions between messenger RNA molecules (mRNAs), the 40S ribosomal subunit, other components of the translation machinery (*e.g.*, eukaryotic initiation factors; eIFs). The current model of mRNA translation initiation postulates that the pre-initiation complex (alternatively "43S pre-initiation complex";
25 abbreviated as "PIC") translocates from the site of recruitment on the mRNA (typically the 5' cap) to the initiation codon by scanning nucleotides in a 5' to 3' direction until the first AUG codon that resides within a specific translation-promotive nucleotide context (the Kozak sequence) is encountered (Kozak (1989) J Cell Biol 108:229-241). Scanning by the PIC ends upon complementary base-pairing between nucleotides comprising the anticodon of the
30 initiator Met-tRNA^{Met} transfer RNA and nucleotides comprising the initiation codon of the mRNA. Productive base-pairing between the AUG codon and the Met-tRNA^{Met} anticodon elicits a series of structural and biochemical events that culminate in the joining of the large

60S ribosomal subunit to the PIC to form an active ribosome that is competent for translation elongation.

Kozak Sequence: The term “Kozak sequence” (also referred to as “Kozak consensus sequence”) refers to a translation initiation enhancer element to enhance expression of a gene or open reading frame, and which in eukaryotes, is located in the 5' UTR. The Kozak consensus sequence was originally defined as the sequence GCCRCC (SEQ ID NO:79), where R = a purine, following an analysis of the effects of single mutations surrounding the initiation codon (AUG) on translation of the preproinsulin gene (Kozak (1986) Cell 44:283-292). Polynucleotides disclosed herein comprise a Kozak consensus sequence, or a derivative or modification thereof. (Examples of translational enhancer compositions and methods of use thereof, see U.S. Pat. No. 5,807,707 to Andrews et al., incorporated herein by reference in its entirety; U.S. Pat. No. 5,723,332 to Chernajovsky, incorporated herein by reference in its entirety; U.S. Pat. No. 5,891,665 to Wilson, incorporated herein by reference in its entirety.)

Modified: As used herein “modified” or “modification” refers to a changed state or a change in composition or structure of a polynucleotide (e.g., mRNA). Polynucleotides may be modified in various ways including chemically, structurally, and/or functionally. For example, polynucleotides may be structurally modified by the incorporation of one or more RNA elements, wherein the RNA element comprises a sequence and/or an RNA secondary structure(s) that provides one or more functions (e.g., translational regulatory activity). Accordingly, polynucleotides of the disclosure may be comprised of one or more modifications (e.g., may include one or more chemical, structural, or functional modifications, including any combination thereof).

Nucleobase: As used herein, the term “nucleobase” (alternatively “nucleotide base” or “nitrogenous base”) refers to a purine or pyrimidine heterocyclic compound found in nucleic acids, including any derivatives or analogs of the naturally occurring purines and pyrimidines that confer improved properties (e.g., binding affinity, nuclease resistance, chemical stability) to a nucleic acid or a portion or segment thereof. Adenine, cytosine, guanine, thymine, and uracil are the nucleobases predominately found in natural nucleic acids. Other natural, non-natural, and/or synthetic nucleobases, as known in the art and/or described herein, can be incorporated into nucleic acids. Unless otherwise specified, the nucleobase sequence of a SEQ ID NO described herein encompasses both natural nucleobases and chemically modified

nucleobases (e.g., a “U” designation in a SEQ ID NO encompasses both uracil and chemically modified uracil).

Nucleoside/Nucleotide: As used herein, the term “nucleoside” refers to a compound containing a sugar molecule (e.g., a ribose in RNA or a deoxyribose in DNA), or derivative or analog thereof, covalently linked to a nucleobase (e.g., a purine or pyrimidine), or a derivative or analog thereof (also referred to herein as “nucleobase”), but lacking an internucleoside linking group (e.g., a phosphate group). As used herein, the term “nucleotide” refers to a nucleoside covalently bonded to an internucleoside linking group (e.g., a phosphate group), or any derivative, analog, or modification thereof that confers improved chemical and/or functional properties (e.g., binding affinity, nuclease resistance, chemical stability) to a nucleic acid or a portion or segment thereof.

Nucleic acid: As used herein, the term “nucleic acid” is used in its broadest sense and encompasses any compound and/or substance that includes a polymer of nucleotides, or derivatives or analogs thereof. These polymers are often referred to as “polynucleotides”.

Accordingly, as used herein the terms “nucleic acid” and “polynucleotide” are equivalent and are used interchangeably. Exemplary nucleic acids or polynucleotides of the disclosure include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), DNA-RNA hybrids, RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, mRNAs, modified mRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, RNAs that induce triple helix formation, threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β -D-ribo configuration, α -LNA having an α -L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino functionalization, and 2'-amino- α -LNA having a 2'-amino functionalization) or hybrids thereof.

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

dimensional and/or three dimensional state of an RNA molecule. Nucleic acid structure can be further demarcated into four organizational categories referred to herein as “molecular structure”, “primary structure”, “secondary structure”, and “tertiary structure” based on increasing organizational complexity.

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

10 *Pre-Initiation Complex (PIC)*: As used herein, the term “pre-initiation complex” (alternatively “43S pre-initiation complex”; abbreviated as “PIC”) refers to a ribonucleoprotein complex comprising a 40S ribosomal subunit, eukaryotic initiation factors (eIF1, eIF1A, eIF3, eIF5), and the eIF2-GTP-Met-tRNA^{Met} ternary complex, that is intrinsically capable of attachment to the 5' cap of an mRNA molecule and, after attachment, of performing ribosome scanning of the 5' UTR.

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

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

Residence time: As used herein, the term “residence time” refers to the time of occupancy of a pre-initiation complex (PIC) or a ribosome at a discrete position or location along an mRNA molecule.

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

24. Equivalents and Scope

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments in accordance with the invention described herein. The scope of the present disclosure is not intended to be limited to the above Description, but rather is as set forth in the appended claims.

In the claims, articles such as "a," "an," and "the" can mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

It is also noted that the term "comprising" is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term "comprising" is used herein, the term "consisting of" is thus also encompassed and disclosed.

Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the present disclosure, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

In addition, it is to be understood that any particular embodiment of the present disclosure that falls within the prior art can be explicitly excluded from any one or more of the claims. Since such embodiments are deemed to be known to one of ordinary skill in the art, they can be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the compositions of the present disclosure (*e.g.*, any nucleic acid or protein encoded thereby; any method of production; any method of use; etc.) can be excluded from any one or more claims, for any reason, whether or not related to the existence of prior art.

All cited sources, for example, references, publications, databases, database entries, and art cited herein, are incorporated into this application by reference, even if not expressly stated in the citation. In case of conflicting statements of a cited source and the instant application, the statement in the instant application shall control.

Section and table headings are not intended to be limiting.

CONSTRUCT SEQUENCES

TABLE 8. Polynucleotide Construct Sequences and Chemistries

*Whereas the text corresponding to Table 3 above indicates that any 5'UTR may begin with an A (not shown in Table 3) or G (not shown in Table 3), Table 8 below shows the starting nucleotide (G or A) for the 5'UTRs used in the disclosed constructs.

SEQ ID NO:	ORF Sequence (Amino Acid)	ORF Sequence (Nucleotide)		5' UTR*	3' UTR		Chemistry
		2	20		50	108	
300	MASEGKLWGRFVGAVDPIME		AUGGCCAGCGAGAGCGGCAAGCUGUGGGGGCGGAGAUUCGUGGGCGCCGUGGACCCCA	AGGAAA	UGAUA	UGAUA	#1A
	KFNASIAYDRHLWEVDVQSKA		UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACUUGGGAGGUGGACGU	UCGCAA	AUAGG	AUAGG	Cap:
	YSRGLERAGLLTKAEMDQILHG		GCAGGGCAGCAAGGCCUACAGCAGAGGCCUUGGAGAGAGCCGGCCUGUGACCAAGGCC	AAUUUG	CUGGA	CUGGA	m ⁷ Gp-
	LDKVAEEWAQGTFKLNSNDEDI		GAGUUGGACCAAGUCCUGCACGGCCUGGACAAGGUGGGCGAGGAGUGGGCCAGGGCA	CUCUUC	GCCUC	GCCUC	ppGm-AG
	HTANERRLKELIGATAGKLHTG		CCUUAAGCUAAGCAACGACGAGGAGCAUCCACACCGCCACGAGAGAAACGACUGAA	GCGUUA	GGUGG	GGUGG	
	RSRNDQVVTDLRLWMRQTCSTL		GGAGCUAUCGGCGCCACCGCGGGUAGCUUACACCGGAGAGAGCGAGAAACGACCCAG	GAUUUC	CCUAG	CCUAG	Poly A tail:
	SGLLWELIRTMVDRAEAERDVL		GUGGUGACCGACCUAGAGACUUGUGGAGAGACAGACCCUGCAGCACCCUGAGCGGCCUUC	UUUUAG	CUUCU	CUUCU	100 nt
	FPGYTHLQRAQPIRWSHWILSH		UUUGGAAUUGAUCAGAACCAUGGUGGACAGAGCCGAGGCGAGAGAGACGUGUGUUU	UUUUUCU	UGCCC	UGCCC	Chemistry:
	AVALTRDSERLLEVRKRINVL		CCCCGGCUACACCCACCUAGCAGAGAGCCAGCCCAUCCGCUUGGAGCCACUGGAUCCUG	CGCAAC	CUUGG	CUUGG	G5
	LGSAGIAGNPLGVDRELLRAEL		AGCCACGCCUGUGGCCUUGACCAAGAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA	UAGCAA	GCCUC	GCCUC	
	NFGAITLNSMDATSERDFVAEF		UCAACGUGCUGCCCCUGGGCAGCGCGGCCAUAGCAGGUAACCCUUUAGGCGUUGACAG	GCUUUU	CCCCC	CCCCC	
	LFWASLCMTHLSRMAEDLILYC		AGAGCUGCUGCGGGCCGAGCUGAAACUUCGGUGGCCAUACACCCUGAAUUCUUGGACGCC	UGUUCU	AGCCC	AGCCC	
	TKEFSFVQLSDAYSTGSSLMPQ		ACUAGCGAGCGCGAUUUCUGGGCAGAGUUCUUGUUGGGCCAGCCUGUGCAUGACCC	CGCC	CUCU	CUCU	
	KKNPDSLELIRSKAGRVFGRCA		AUCUGAGCAGAAUGGCCGAAGACCUGAUCUGUAUCUGACCAAGGAGUUCAGCUUCGU		CCCCU	CCCCU	
	GLLMTLKGLPSTYNKDLQEDKE		GCAGCUGAGCGCAGCAUAVAGCACCGGAAGCAGCCUAGUAGCCCCAGAAAGAACCCC		UCCUG	UCCUG	
	AVFEVSDTMSAVLQVATGVIST		GACAGCCUGGAACUUAUCAGAAAGUAGGCGAGGCGAGAGUUCGGCCGGUGCGCGGUC		CACCC	CACCC	
	LQIHQENMGQALS PDMLATDLA		UGCUAUGACCCUGAAGGGCCUGCCAGCACCUAACAACAAGGACCUUCAGGAGGACAA		GUACC	GUACC	
	YYLVRKGMFRRQAEASGKAVF		GGAGGCCGUGUUCGAGGUGAGCGCACCAUGAGCGCCGUGCUGCAGGUGGCCACCCGC		CCCGU	CCCGU	
	MAETKGVNLNQLSLQELQTI SP		GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCCGACA		GGUCU	GGUCU	
	LFSGDVICVWDYGHVSVEQY GAL		UGUUAGCCACAGAUUGGCCUACUACCUUAGUCAGAAAGGGCAUGCCAUUCCGGCAGGC		UUGAA	UUGAA	
	GGTARSSVDWQIRQVRALLQAA		CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCGAAACGAGGGAGUUGCGCUGAAC		UAAAG	UAAAG	
	QA		CAGUUGUCUCCAGGAACUGCAGACCAUUCUCCUUCUUGUACGCGGGCGACGUGAUCU		UCUGA	UCUGA	
			GCGUGUGGGAUACGGCCACAGCGUGGAGGAGUACGGCGGCGUUGGUGGCACCGCACG		GUGGG	GUGGG	
			AUCCAGCGUAGACUGGCGAGAUCCGCCAGGUGAGAGGCCUUGCUGCAAGCCAGCAGGCC		CGGC	CGGC	

	2	20	50	108	Chemistry
	MASEGKLWGGRFVGAVDPIME KFNASIAYDRHLWEVDVQGSKA YSRGLERAGLLTKAEMDQILHG LDKVAEEWAQGTFKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQVVTDLRLWMRQTCSL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVLP LGS GAIAGNPLGV DRELLRAEL NFGAITINSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGRVFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALS P DMLATDLA YYLVRKGMPPFROAHEASGKAVF MAETKGVALNQLSLQELQTI SP LFSGDVICWVDYGH SVEQY GAL GGTARSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCACAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCGGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUAAACCCUUUAGGCGUUGACAG AGAGCUGCUGCGGGCCGAGCUGAAACUUCGGUGCCAUCACCCUGAAUUCUUGGACGCC ACUAGCGAGCGGAUUUCUGGGCAGAUCCUGUUCUGGGCCAGCCUGUGCAUGACCC AUCUGAGCAGAAUGGCCGGAAGACCUGAUCCUGUACUGCAACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGAUGCCCCAGAAAGAAACCCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAGGCAGAGUUCGGCCGUGCGCGGGUIC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACAACAAGGACCUUCAGAGGACAA GGAGCCGUGUUCGAGGUGAGCGCACCAUGAGCGCCGUGUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAAUUGGGCCAGGCCUUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCGAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCGAAACGAAGGAGUUGCGCUGAAC CAGUUGUCUCCAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGGCAGCUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCUGGUGGCACCGCACG AUCCAGCGUAGACUGGCAGAUCCGCCAGGUGAGAGCCUUGCUGCAAGCCAGCAGGCC	GGGAA AUCGC AAAAU UUGCU CUUCG CGUUA GAUUU CUUUU AGUUU UCUCG CAACU AGCAA GCUUU UUGUU CUCGC C UGAAUA AAGUCU GAGUGG GCGGC	#1B Cap: Cap1 Poly A tail: 100 nt Chemistry: G5	

	2	20	50	128	Chemistry
302	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLERAGLLT KAEMDQ I LHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGHVSVEQYGAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUCACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGCUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGCACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGGAAGACCUGAUCCUGUACUGCACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAAACUUAUCAGAAAGUAAGGCAAGGCAAGUUCGGCCGUGCGCGGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCAUACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGAUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUCGC AAAAU UUGCU CUUCG CGUUA GAUUU CUUUU AGUUU UCUCG CAACU AGCAA GCUUU UUGUU CUCGC C	UAAAGC UAAAGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#2 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	2	20	50	138	Chemistry
	MASEGKLWGGRFVGAVDPIME KFNASIAYDRHLWEVDVQGSKA YSRGLERAGLLTKAEMDQILHG LDKVAEEWAQGTFKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQVVTDLRLWMRQTCS TL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGVDR ELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGRVFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPRQAHEASGKAVF MAETKGVALNQLSLQELQTI SP LFSGDVICWVDYGHVSVEQYGAL GGTARSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCACAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCGCCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAAACCAUGGUGGACAGAGCCGAGGCGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUGCUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUAAACCCUUUAGGCGUUGACAG AGAGCUGCUGCGGGCCGAGCUUGAACUUCGGUGCCAUACCCUGAAUCCAUGGACGCC ACUAGCGAGCGGAUUUCUGGGCAGAUCCUGUUCUGGGCCAGCCUGUGCAUGACCC AUCUGAGCAGAAUGGCCGAAGACCUUGAUCCUGUACUGCACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUGAUGCCCCAGAAAGAAACCCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAGGCAGAGUGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAAAGGACCUUCAGGAGACAA GGAGCCGUGUUCGAGGUGAGCGCACCAUGAGCGCCGUGUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAAUUGGGCCAGGCCUUGAGCCCCGACA UGUUAAGCCACAGAUUCUGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCGAAACGAAAGGAGUUGCGCUGAAC CAGUUGUCUCCAGGAACUGCAGACCAUUCUCCCCUCUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCUGGUGGCACCGCACG AUCCAGCGUAGACUGGCAGAUCCGCCAGGUGAGAGCCUUGCUGCAAGCCAGCAGGCC	GGGAA AUCGC AAAAU UUGCU CUUCG CGUUA GAUUU CUUUU AGUUU UCUCG AACAGC UCCUUA CUCUGA GGAAGU UGGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#3 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5	

	2	20	78	137	Chemistry
304	MASESGKLWGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLERAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTI SP LFSGDVICWWDYGH SVEQY GAL GGTARSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGAU GGACCAAGUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCU G CAGAGAGCCCAAGCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUUUAUCGGUGCCAUCCUUGUACUACCCUUGAUAUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCUGUGGACAGUUCUUGUACUUGUACUAGCCAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUUGCAGCCAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUUGCCCAAGAGAAAGAACCC GACAGCCUGGAAACUUAUCAGAAAGUAAGGCAGGACAGAGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUAGGCCCGACA UGUUA GCCACAGAU CUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCAUUGGCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUC CAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCGCAGGCC	GGGAA AUCUC CCUGA GCUUC AGGGA GUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UAAUAG UAAACC UCACUC ACGGCC ACAUUG AGUGCC AGGUC CGGCU GGUUA UAGUAG UGUAGA GCAUUG CAGCAC UUAGAC UGGGGU GCUGUA GUCUUU AUUGUA GUCUUU CCACAU ACCUGA UAAUUC UUAGAU AAUUC UUUUUU UAAUUA AAUCUA UUCUUA GGCU	#4 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	2	20	78	139	Chemistry
305	MASEGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLERAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQVTDLR LWMRQTCSTL SGLLWELIRTMVDR AEAERDVL FPGYTHLQRAQPIR WSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVI ST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGAU GGACCAAGUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCCAG GUGGUGACCGACCUAGAGACUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUGAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCU G CAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUUUAUCGUGGAGAGUUCUUGUGCCAUACCCUUAUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCGUGGAGAGUUCUUGUACUGACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUUGCCCAAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAGGAGAGUGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACAACAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUC CAGGAACUGCAGACCAUUCUCCCUUGUUCAGCGCGGACGUGAUCU GCGUGUGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCGCAGGCC	GGGAA AUCUC CCUGA GCUUC AGGGA GUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UAAAGC UAAACC UCACUC ACGGCC ACAUUG AGUGCC AGGUC CGGCU GGUUA UAGUAG UGUAGA GCAUUG CAGCAC UUAGAC UGGGGU GCUGUA GUCUUU AUUGUA GUCUUU CCACAU ACCUGA UAAUUC UUAGAU AAUUUC UUAUUU UAAUUA AAUCUA UUCUUA GGCU	#5 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	2	20	56	108	Chemistry
306	MASESGKLWGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLERAGLLT KAEMDQ I LHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTISP LFSGDVICVWDYGHVSVEQYGAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGUGGGCGGCAGAUUCUGUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGAGGCCUGGAGAGAGCCGGCCUGCUGACCAAGGCC GAGUUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCACAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACCCACCUGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAAGGCAAGUUGUGCCCGGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAACAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUGAGCCCCGACA UGUUAAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#6 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	3	21	56	108	Chemistry
307	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLTRAEMDQILHG LDKVAEEWAQGTFLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQVTDRLRWMRQTCSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTISP LFSGDVICWWDYGH SVEQYGAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUCACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCCAGAGCC GAGAUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCCGAGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCCAG GUGGUGACCGACCUAGACUCUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGCUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUCUGUGGCGAGUUCUUGUACUGACUUGGCCAGCCUUGAUACCC AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACUAGCAGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCGCAGAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAAGGCAAGUUCGGCCGUGGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCCGACA UGUUAAGCCACAGAUUCUGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGAUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#7 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	4	22	56	108	Chemistry
308	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGTFR LNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPIR WSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUCACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUCAGACUUAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUCACACCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAUGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGCUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGCUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUCUGUGGCAGAUUCUUGUACUGACUUGGCCAGCCUUGAUACCC AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACUAGCAGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAGGCGAGAGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#8 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	5	23	56	108	Chemistry
309	MASESGKLWGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRRL ELIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRD SERLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFGAITLNSMDAT SERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQYGAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAG AGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAUGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAGGAGGAGGAGUGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGAUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#9 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	6	24	56	108	Chemistry
310	MASESGKLWGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATGGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQYGAL GGTARSSVDWQIRQVRALLQAAQ	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGGGCGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCGUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACACCAAGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACACCAAGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAAGGCAAGUUGCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAACAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#10 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	7	25	56	108	Chemistry
311	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRLA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAAGGCAAGUGUUCGGCCGCGUGGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAACAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCCCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#11 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	8	26	56	108	Chemistry
312	MASEGKLGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFDAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTI SP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGGCCGCCUGCUGACCAAGGCC GAGUUGGACCAAGUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGACGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGCACCCUUGAACUUGGACGCC AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGCACCAAGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAGGAGGAGGAGUGUUCGGCCGUGCGGGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACGAGGAGAAACAUGGGCCAGGCCUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#12 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	9	27	56	108	Chemistry
313	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFGAITLNSMDAVSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPRQAHEASGKAVF MAETKGV ALNQLSLQELQTISP LFSGDVICWWDYGH SVEQYGAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGAUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAUGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC GUGAGCGAGCGCGAUUUCGUGGCGAGUUCUUGUACUGCAACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGGAAGACCUGAUCCUGUACUGCAACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAAACUUAUCAGAAAGUAAGGCAAGGCAAGUGUUCGGCCGUGCGGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUAGAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUCUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGAUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#13 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	10	28	56	108	Chemistry
314	MASEGKLGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLTKAEMDQILHG LDKVAEEWAQGTFKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQVTDRLWMRQTCSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL LGSAGIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TPEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRC GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTI SP LFSGDVICVWDYGHVSVEQY GAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCGGCCUGCUGACCAAGGCC GAGUUGGACCAAGUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUAGACCAAGAGUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACUUGGCCAGCCUUGAUUCGU AUCUGAGCAGAAUGGCCGAGACCUAGUCCUGUACUGACUUGGCCAGCCUUGAUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCGCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAGGAGGAGGAGUUGUGCCCGGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUACAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCUUCUUGUUCAGCGGCCGAGUUCU GCGUGUGGAGCUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCUUGCUGCAAGCCCGCAGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#14 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	11	29	56	108	Chemistry
315	MASEGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQVTDLR LWMRQTCSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRC A GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTISP LFSGDVICWWDYGHVSVEQYNAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGCCCGCCUGCUGACCAAGGCC GAGUUGGACCAGAUCCUGCAGCGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGCGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACCCACCUAGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCGUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAACUCCAUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGACACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAGACCUUGAUCCUGUACUGACACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUGCCCCAGAAAGAACCC GACAGCCUGGAACUUAUCAGAAAGUAAGGCAAGGCAAGUGUUCGGCCGUGCGCGGUC UGCUAAUGACCCUGAAGGGCCUGCCAGCACCUACAACAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUGAGCCCCGACA UGUUAAGCCACAGAUUCGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCGGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGCGCAGCUGAUCU GCGUGUGGGAUACGGCCACAGCGUGGAGCAGUACAACGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#15 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	1	30	55	108	Chemistry
	MASESGKLWGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQTCSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFGAITLNSMDAT SERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTI SP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGCGGAGCGAGAGCGGUA AAUUGGGGAGGACGGUUCUGGGUGCCGUGGACCCUA UAAUGGAGAAAUUCAAACGCCAGUAUUGCAUACGAUCGUAUCUGUGGGAAGUGGACGU GCAAGGAAGCAAAGCUUACUCCAGAGGCCUCGAGAAAGCUGGCCUGCUGACCAAGGCC GAGUUGGACCAAGUCCUGCAGCGGACUGGACAAAGUGGCCGAGGAGUGGGCACAGGGAA CUUCAAACUGAAUUCAAACGACGAGGACAUCAUACCCGCCAACGAGCGGAGACUGAA GGAACUGAUCGGCGCCACAGCUGGCAAGCUCCACACAGGACGAUCAAGGAACGACCCAG GUGGUGACAGAU CUGAGGCUUGUGAUGAGGACAGACCUGUCGACCCUCUCAGGUCUGC UGUGGGAGUUAUCCGCCACCAUGGUCGAUGCGGCCGAGCCGAGCGGACGUAUCUGU CCCAGGGUACACUACUUGCAGAGGGUCAGCCUAUCCGUGGUCGCACUGGAUCCUC AGUCACGCCGUGGGGUGACCGGGGACUCUGAGCGGCCUGCUGGGAAGUGCGUAACGCA UCAACGUGUGCCACUGGGUCUCCGGGCUAUUGUGGCAUCCACUAGGCGUGGACCG GGAGCUUCUGAGAGCCGAACUGAACUUCGGGGCGAUCACCUUAAUAGCAUGGACGCC ACUCCGAACGCGAUUUCGUUGCAGAGUUCUGUUUUGGGCAUCCUGUGCAUGACUC ACCUCAGUCGU AUGGCCGAAGACCUAUCCUCUAUUGCACAAAGAGAUUUCUUUGU GCAGCUUUCUGACGCCUACAGCACAGGUUCCUCACUCAUGCCACAGAGAAGAAUCCU GACAGCCUAGAGCUAAUCAGAUCAUAGGCCGCGCAGAGUCUUUGGCCGUGUGCAGGGC UGCUGAUGACACUUAAGGACUGGCCUCUACUUAUAAUAAAGGACCUGCAAGAGGAUAA GGAGGCCGUGUUGAGGUUAGCGACACCAUGUCAGCCGUCUCCAGGUAGCAACGGGC GUGAUCAGCACUCUGCAGAUCCAUCAAGAGAAACAUGGGUCAGGCCUGUCACCCAGACA UGCUGGCCACAGAUUUGGCCUAUUAACUGGUGCGCAAGGGGAUGCCCUUUGGGCAAGC CCACGAAGCAUCAGGCAAGGCUUGUUAUGGCAGAGACAAAGGAGUGGGCCCUCAAC CAAUUGUCUCUGCAGGAUUGCAAACCAUCUCUCCUGUUUCCGGGGACGUGAUCU GCGUGUGGGACUACGGGCAUUCUGUGGAACAGUACGGUGCGCGUGGGAGGGACCGCUAG GUCCUCCGUCGACUGGCAGAUCCGGGCAAGUGCGGGGCCCCUGCUGCAAGCGCAGCGGCU	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG AGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#16 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	1	31	56	111	Chemistry
331	MASESGKLWGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQVTDLR LWMRQTCSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTISP LFSGDVICVWDYGHVSVEQYGAL GGTARSSVDWQIRQVRALLQAA QA	AUGGCAAGCGAGUCCGGGAAACU UUGGGGAGGUGCGGUUUGUGGGCGCUGUGGACCCCA UUAUGGAGAAAGUUAACCGGUCCAUCGCAUACGACCGACACACCUCUGGGAGGUGGACGU GCAGGGGCUAAAGCCUAUAGCAGGGGUUCUGGAGAAAGCCGGGCGUCUCACUAAGGCU GAAUGGACCAGAUACUGCACGGGCGUGGAUAAAGGUGGCUAGAGUGGGCCCAAGGCA CGUUAAGCUGAAUAGCAACGACGAGGACAUAACACACAGCCACGAGAGAAGACUGAA GGAGUUAUUGGGCGACAGCCGGGAAGUUGCACACUGGCAGGUCCCGGAACGACCAA GUGGUAAACCGACUUGCGGUUGUGGAUGCGUCAGACGUGUCCACUCUGAGCGGCUUGC UGUGGGAGCUAAUCCGUACA AUGGUGGACAGGGGCUAGCCGAGAGACGUGGCUUUG CCCAGGAUACACCCACCGCAGCGGACAGCCUAUACGGUGGUGGCGACUGGAUUUG AGCCACGGGUGCGGCUCAUAGAGAUAGCGAGCGCCU UUGGAAUGUGCGCAACCGGA UCAACGUUCUGCGGUCGGGUCCGGGCAUUGCCGGAAUCCCCUGGGCGUAGACAG GGAGCUCCUAAGACGAGUUGGAGAAUUCGGUGCCAUAAACCUGAACAGCAUGGACGCC ACAUCGGAGCGGGAUUCGUGGCGAAGAUUUUAUUGUAUUGAACAGAGAUUUUUGU AUCUGAGCCGU AUGGCGGAAGAUUUUAUUGUAUUGUAUUGAACAGAGAUUUUUGU CCAGCUGUCAGACGCUUACUCCACAGGGUCAUCACUGAUGCCACAGAGAAGAACCCU GACAGCCUGGAAUUAUCCCGCUCCAGGGCGGCAUUGGGCGUUGUGCGCGCC UGCUGAUGACUCUGAAGGGUUUUAACCUUACCUAUAACAAGGAUCUCCAGGAGGACAA GGAAAGCUGUUUCGAGGUGUGCCGAUACCAUGUCCCGCGUGUUGCAGGUGGCAACAGGU GUUAUAGCACCCUCCAGAUUACCAAGAGAAACAUGGGCCAGGCACUGCGCCGAUA UGCUGGCUACUGAUCUGGCAUUAUUCUGCCGGAAGGGCAUGCCAUUCGGCGAGGC GCACGAGGCCUCAGGGGAAGGCUUGUUAUGGCGGAAACCAAGGGCGUGGCAUCUAAAC CAGUUGAGCUUGCAAGAGCUACAGACUAUCAGCCCCUCUUAUUGGAGCGUGAUCU GUGUGGGGAUUAACGGCCACUCCGUGGAGCAGUACGGUGCCCCUCGGUGGCAACGCCG AAGCUCUGCAUUGGCAGAUUCGGCAGGUGCGGGCGCUACUGCAAGCCCCAGCAGGCA	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCUU UGGGCC UCCCCC CAGCCC CUCUUC CCCUUC CUGCAC CCGUAC CCCCUC CAUAAA GUAGGA AACACU ACAGUG GUCUUU GAUAAA AGUCUG AGUGGG CGGC	#17 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	1	32	55	111	Chemistry
	MASESGKLWGGRFVGA VDPIME KFNASIA YDRHLWEVDVQGSKA YSRGLEKAGLLTKAEMDQILHG LDKVAEEWAQGTFKLNSNDEDI HTANERRLKELIGATAGKLHTG RSRNDQVTDRLRWMRQTCSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL LGS GAIAGNPLGVDRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRC A GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGVALNQLSLQELQTISP LFSGDVICVWDYGHVSVEQYGAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGCGGAGUGAAAGUGGGAAAUUGGGGUGGGCGCUUCGUAGGGGCGGUGGACCCUA UCAUGGAGAAAUUCAAACGCAAGCAUCGCCUACGAUAGGCAUCUUGGGAGGUUGACGU UCAAGGUUCUAAAGCGGUUUAAGAGGGCUAGAGAAAGCGGGACUGCUUACGAAAGCC GAGUUGGACCAAGUCCUACACGGGCUAGACAAGGUCGAGAAAGAGUGGGCUCAGGGGA CUUCAAACUCAAACAGCAACGACGAAGAAUUAUACUACUACGAGAGGAGACUGAA GGAAUCGAUAGGGGCCACCGCUGGCAAGCUGCACACAGGGCGCAGCCGCAACGACCCAG GUCGUGACUGACCUAGAGACUUGGAUGCGCCAGACGUGCAGUACACUGCAGGCCUCC UUUGGGAGCUGAUCGGGACUUGGUGGAUAGGGCCGAGGCGGACGCUUUGUU UCCAGGCUACACUACCGUGCAGCGGCUCAACCUAUUCGUUGGUCACACUGGGAUACUG AGUCACGGCGUUGCCUACUAGAGACUACAGACGGCUCUUGGAAGUACGGAGCGGUA UCAACGUACUACCCUUGGAUCUGGAGCCUACGAGGGACCCACUAGGGCGUAGACAG AGAAUUGUUGCGAGCCGAAUCUGAAUUCGGCGCAUAAACACUGAAUUCUUGGACGCC ACGAGCGAGCGGGAUUUGUCGAGAGUUUCUUGGGCAAGUUUGUGUAUGACGC ACCUUCCAGGAUGGCGUAGAACCCUGAUCCUGAUUUGUAUAGGAGUUUCCUUCGU CCAAAUUAGCGACGCAUACAGCACCGGCGAGCUCUUAUUGCCCGAGAGAAAUUCC GAUUCGCCUGGAGCUCUACAGAAAGCGGCUUGGGCGGUGUUGGUGGCGCGGCC UCUUGAUGACACUGAAAGGCCUUGCCUCCACCUACAUAUAAAGAUUUGCAGGAAGAUAA AGAGGCCGUCUUCGAGGUUUCUGACACAAUGUCUGCCGUGUUAACAAGUGGCUACAGGU GUGAUUAGUACCUUGCAGAUCCACAGGAGAAACAUGGGUCAAGCCUGCGCCGGAUA UGCUCGCCACCGACUUGGCUUACUACUGGUGCGAAAGGGUAGCCUUCGGCGAGGC CCACGAGGCCAGUGGAAAGGCCGUGUUAUUGGGCGAAACUAAAGGAGUAGCUCUGAAC CAACUAGCCUUCAGGAGUUAACAACUACUCCCGCUGUUCAGUGGUGACGUGAUCU GCGUGGGGACUACGGCCAUUCUGUGGAACAGUACGGAGCGCUGGGUGGAACCGCCCG UUCCUCAGUGGACUUGGCAGAUUCGCCAGGUACGGGGCCUUGCUGCAGGGCGCAGCGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG AGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCUU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCUC CAUAAA GUAGGA AACACU ACAGUG GUCUUU GAUAAA AGUCUG AGUGGG CGGC	#18 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

332

	1	33	55	108	Chemistry
333	MASESGKLWGGRFVGAVDPI ME KFNASIA YDRHLWEVDVQGS KA YSRGLEKAGLLT KAEMDQILHG LDKVAEEWAQGT FKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQV TDLRLWMRQT CSTL SGLLWELIR TMVDRAEAERDVL FPGYTHLQRAQPI RWSHWILSH AVALTRDSE RLLLEVRKRINVL P LGS GAIAGNPLGV DRELLRAEL NFGAITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLILYC TKEFSFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGR VFGRCA GLLMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALSPDMLATDLA YYLVRKGMPPFRQAHEASGKAVF MAETKGV ALNQLSLQELQTI SP LFSGDVICWWDYGH SVEQY GAL GGTARSSVDWQIRQVRALLQAAQ QA	AUGGCCAGCGAGAGCGGCAAGCTUGGGCGGCGAGAUUCUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUGGGAGGUGGACGU GCAGGGCAGCAAGGCCUACAGCAGAGGCCUGGAGAAAGGCCGCCUGCUGACCAAGGCC GAGUUGGACCAGAUCCUGCAGCGGCCUGGACAAGGUGGCCGAGGAGUGGGCCAGGGCA CCUUAAGCUGAAACAGCAACGACGAGGACAUCACACCCGCCAACGAGAGAAGACUGAA GGAGCUGAUCGGGCCACCGCGGGUAAGCUUACACCCGGCAGAAAGCAGAAACGACCAG GUGGUGACCGACCUAGACUUGUGGAGAGACAGACCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUAUCAGAACCAUGGUGGACAGAGCCGAGGAGAGAGACGUGCUGUU CCCCGGCUACCCACCUGCAGAGAGCCCAAGCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGCCUUGACCAAGAUAGCGAGAGACUUGUGGAGGUGAGAAAGAGAA UCAACGUGUGCCCCUGGGCAGCGGCCCAUAGCAGGUACCCUUUAGGCCUUGACAG AGAGCUGUGCGGGCCGAGCUGAAACUUCGGUGCCAUACCCUUGAAUUCUUGGACGCC ACUAGCGAGCGCGAUUUCUGGGCAGAUUCCUGUACUGCAACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGGAAGACCUGAUCCUGUACUGCACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAUAGCACCGGAAGCAGCCUUGAUUGCCCCAGAAAGAACCC GACAGCCUGGAAACUUAUCAGAAAGUAAGGCAAGGCAAGUUGUUCGGCCGUGCGGGUC UGCUAAUGACCCUGAAGGGCCUGCCCCAGCACCUACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUUAGCACCCUUCAGAUCCACAGGAGAAACAUGGGCCAGGCCUUGAGCCCCGACA UGUUAAGCCACAGAUUCUGGCCUACUACCUAGUCAGAAAGGGCAUGCCAUUCGGCAGGC CCACGAGGCCAGUGGCAAGGCCGUGUUCUUGGCCCCGAAACGAGGGAGUUGCGCUGAAC CAGUUGUCUUCACAGGAACUGCAGACCAUUCUCCCCUUGUUCAGCGGCCGACGUGAUCU GCGUGUGGGAACUACGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCCACCGCACG AUCCAGCGUAGACUUGGCAGAUCCGCCAGGUGAGAGCCCCUGCUGCAAGCCCCAGCGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAAGAA GAAAU AUAAG AGCCA CC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCGU GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#19 Cap: Cap1 Poly A tail: 100 nt Chemistry: G5

	1	33	56	108	Chemistry
	MASESGKLWGGRFVGAVDPI ME KENASIA YDRHLWEVDVQGSKA YSRGLEKAGLLTKAEMDQILHG LDKVAEEMAQGTFKLNSNDEDI HTANERRLKE LIGATAGKLHTG RSRNDQVVTDLR LNMRT CSTL SGLLWELIRTMVDRAEAERDVL FPGYTHLQRAQPIRWSHWILSH AVALTRDSERLLEVRKRINVL P LGSGA IAGNPLGVDR ELLRAEL NFGA IITLNSMDATSERDFVAEF LFWASLCMTHLSRMAEDLLIYC TKFE SFVQLSDAYSTGSSLMPQ KKNPDSLELIRSKAGRVFGRCA GLIMTLKGLPSTYNKDLQEDKE AVFEVSDTMSAVLQVATGVIST LQIHQENMGQALS PDM LATDLA YYLVRKGMPFRQAEHESKAVF MAETKGVALNQLSLQELQ TISP LFSGDVICWMDYGHVSVEQYGAL GGTARSSVDWQIRQVRALLQAO QA	AUGCCAGCGAGAGCGGCAAGCUGUGGGGCGGCAGAUUCUGUGGGCGCCGUGGACCCCA UCAUGGAGAAAGUUAACGCCAGCAUCGCCUACGACAGACACACCUUGUGGGAGGUGGACGU GCAGGGCAGCAAGCCUACAGCAGAGGCCUGGAGAAAGCCGCGCUGCUGACCAAGGCC GAGAUGGAC CAGAUUCUGCACGGCCUGGACAAGGUGGCCGAGGAGUGGGCCCCAGGGCA CCUUAAGCUGAACAGCAACGACGAGGACAUCACACCGCCAA C GAGAGAA GACUGAA GGAGCUGAU CGGCGCCACCGCGGGUAAGCUUCACACCGG CAGAA G CAGAAACGACCAG GUGGUGACCGACCCUGAGACUGUGGAUGAGACAGACCCUGCAGCACCCUGAGCGGCCUUC UUUGGGAUUUGAUCAGAAACCAUGGUGGACAGAGCCAGCCCCAUCCGUGGAGCCACUGGAUCCUG CCCCGGCUACACCCACCCUGCAGAGAGCCAGCCCCAUCCGUGGAGCCACUGGAUCCUG AGCCACGCCGUGGGCCUGAC CAGAGAUAGCGAGAGACUGCUGGAGGUGAGAAAGAGAA UCAACGUGCUGCCCCUGGGCAGCGGCCCAUAGCAGGUAACCCUUUAGGCGUUGACAG AGAGCUGCUGCGGGCCGAGCUGAA CUUCGUGCAGAGUUCUGUUCUGGGCCAGCCUGCAUGACCC ACUAGCGAGCGCGAUUUCUGUGGCAGAGUUCUGUACUGACCAAGGAGUUCAGCUUCGU AUCUGAGCAGAAUGGCCGAAGACCUGAUCCUGUACUGCACCAAGGAGUUCAGCUUCGU GCAGCUGAGCGACGCAUAVAGCACCGGAAGCAGCCUGAUUGCCCCAGAAAGAAACCCC GACAGCCUGAA CUUAUCAGAAUAAGG CAGG CAGAGUUCUGGCCGUGCGCGGUC UGCUAAUGACCCUGAAAGGCCUGCCCAGCACCUACAACAAAGGACCUUCAGGAGGACAA GGAGGCCGUGUUCGAGGUGAGCGACACCAUGAGCGCCGUGCUGCAGGUGGCCACCGGC GUGAUAGCACCCUUCAGAUCCAC CAGGAGAA CAUGGGCCAGGCCUGAGCCCCGACA UGUUA GCCACAGAU CUGGCCUACUACUAGUCAGAAAGGCAUGCCAUUCGGCAGGC CCACGAGCCAGUGGCAAGCCGUCUUCAUUGGCCGAAACGAAGGAGUUGCGCUGAAC CAGUUGUCUUC CAGGAA CUGCAGACCAUCUCCCCUCUGUUCAGCGGCGACGUGAUCU GCGUGUGGGA CUA CGGCCACAGCGUGGAGCAGUACGGCGCGCUCGGUGGCACCGCACG AUCCAGCGUAGACUGGCAGAUCCGCCAGGUGAGAGCCUGCUGCAAGCCCCAGCAGGCC	GGGAA AUAAG AGAGA AAAGA AGAGU AAGAA GAAAU AUAAG ACCCC GGCGC CGCCA CC CUGCAC CCGUAC CCCCG GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	UGAUAA UAGGCU GGAGCC UCGGUG GCCUAG CUUCUU GCCCCU UGGGCC UCCCCC CAGCCC CUCCUC CCCUUC CUGCAC CCGUAC CCCCG GGUCUU UGAAUA AAGUCU GAGUGG GCGGC	#20 Cap: Cap1 PolyA tail: 100 nt Chemistry: G5

EXAMPLES

EXAMPLE 1: Materials and Methods

Western Blotting- Measuring Liver ASL levels

Liver (25-30 mg) was added to Precellys homogenising tube (VWR International, Cat # 431-0710) containing 400 µl of cold homogenising buffer (1x RIPA buffer + 1x protease inhibitors). The tube was loaded onto Precellys tissue homogeniser machine (Bertin instruments) and lysed at speed 4500 rpm 30 seconds x 2 cycles. Homogenates were centrifuged at 10,000 x g for 20 min at 4 °C, and supernatants were kept frozen at 80 °C. BCA assay was performed per manufacturer's instructions to determine protein concentration.

40 µg of the liver supernatant were diluted 1:1 with 2x Laemmli sample buffer (containing 10% 2-β-mercaptoethanol (β-ME)) at final volume of 40 µl, vortexed and heated to 95°C for 10 minutes. The samples were stored in -20 °C until use.

SDS-PAGE was used to separate proteins based on their molecular weight (MW). The separated proteins in the acrylamide gel were then transferred to an Immobilon-PVDF membrane. Once the transfer was completed, the membrane was removed from the wet transfer cassette and blocked in 5% non-fat milk powder in PBS-T (PBS+ 0.1% Tween-20 for an hour at room temperature ("rt"). The membrane was then incubated with primary antibody (1:1000 anti-ASL and 1:5000 anti-GAPDH) in 2.5% milk in PBS-T overnight at 4 °C on a shaker. Following incubation, the membrane was washed three times with PBS-T for 5 minutes at rt on to remove any excess unbound primary antibody. The membrane was then incubated with the respective secondary antibodies diluted 1:10,000 in 5% milk in PBS-T in rt for an hour. Again, the membrane was washed three times with PBS-T for 5 mins each to remove excess unbound secondary antibodies at rt on a shaker. Acquiring and analysis was performed in Licor Image Studio software.

ASL Activity Assay (Fumarate Detection Assay)

60 µg protein lysate was added to final concentration of 3.6 mM argininosuccinic acid and made upto 50 µl volume with 50mM phosphate buffer (pH 7.3). The reaction was incubated at 37 °C for 1 h and stopped by heating at 80 °C for 20 min (in a thermocycler).

The mixture was centrifuged at 10,000 x g for 5 minutes and supernatant transferred to a clean tube. 5ul of the supernatant was used to measure fumarate levels using Fumarate Assay Kit (Abcam, Cat # Ab102516) per the manufacturer's specifications. After 60 minutes incubation per kit instruction, plates were read in Absorbance mode at 450 nm in plate reader.

5 Fumarate concentration (ng/ul) was determined per manufacturer's guide.

Ammonia and alanine transaminase (ALT) Measurements

Plasma was collected by adding whole blood collection into a red top EDTA tube and centrifuged at 10,000 x g for 4 minutes at rt instantly. Supernatant is then transferred onto microcentrifuge tube and stored at -80°C until analysis.

10 *Ammonia and ALT reads from plasma:* 10 µl of plasma is spotted on to the plasma cartridges for ammonia and ALT respectively and values read in Fujifilm NX600 machine. Ammonia reads were measured in µmol/l and ALT in Units/L. Plasma samples were diluted in PBS when necessary.

15 *Ammonia reads from whole blood:* 10 µl of whole blood was spotted onto whole blood cartridges for ammonia and read using NX10N Fujifilm machine.

In-Cell Western Assay

Plating cells and treatment:

1. Plated 5000 cells per well of 96-well dish (CellBind 96 well microplate) in Complete DMEM media (DMEM media supplemented with 10%FBS and 50 units of Penicillin and Streptomycin).
2. The next day, replaced media and treat as necessary. In case of LNPs, 0.2 µg LNPs in 500 µl of media per well were added. For positive control, separate wells were transfected with equal amount of ASL mRNA. For negative control, PBS treated wells were also included.
- 25 a. Transfection protocol: 0.3 µl lipofectamine 2000 was added to 20 µl Optimem media and left at rt for 10 minutes.
- b. 0.25 µg unformulated mRNA was added to 20 µl of Optimem media separately.
- c. Post 10 minutes, the mRNA+optimum mixture was added to the lipofectamine mixture, mixed and left at rt for 5 minutes.
- 30

- d. 40 μ l mixture was added per well and incubated at 37 °C for 4 hours.
- e. After 4 hours, the transfection mixture was replaced with Complete DMEM media.
3. Leave for 24h.
- 5 4. Each condition was performed in duplicates.

hASL In-Cell Western:

1. Carefully removed media from plate as not to disturb the cell layer.
2. Washed the plate with 300 μ l/well of PBS.
- 10 3. Fixed the plate with 75 μ l/well of ice cold methanol (stored at -20°C) for 10-15 minutes at room temperature.
4. Inverted plate on bench paper to remove methanol.
5. Washed the plate with 300 μ l/well PBS.
6. Added 300 μ l/well of Licor PBS Blocking Buffer and incubate for 90 minutes at room
- 15 temperature on the Belly Dancer (speed setting 1).
7. Removed blocking buffer.
8. Diluted Anti-ASL antibody 1:1000 in Licor PBS Blocking Buffer and add 100 μ l/well of the dilution to the plate. Incubate for 2 hours at room temperature on the Belly Dancer (speed setting 1).
- 20 9. Removed antibody.
10. Washed plate with 300 μ l/well of PBS containing 0.1% Tween for 5 minutes on the Belly Dancer (speed setting 1). Repeat for a total of 4 washes.
11. Diluted CellTag700 and IRDye 800CW Goat anti-Rabbit IgG 1:1,000 in Licor PBS Blocking Buffer and add 100ul/well of the dilution to the plate. Cover the plate with
- 25 aluminum foil and incubated for 1 hour at room temperature on the Belly Dancer (speed setting 1).
12. Washed plate with 300ul/well of PBS containing 0.1% Tween for 5 minutes on the Belly Dancer (speed setting 1). Repeated for a total of 3 washes.
13. Washed plate with 300 μ l/well of PBS for 5 minutes on the Belly Dancer.
- 30 14. Removed PBS and gently pat the plate dry on bench paper.

Read plates immediately on the Licor Odyssey CLx.

C¹³ Ureagenesis Assay

1. *Animal administration pre-harvest:*

1% acetate (60 mg of C¹³ sodium acetate powder diluted in 6mL of sterile PBSD) was injected intraperitoneally (IP). Injection was performed at 5.4 µL/gram so 135 µL is administered for a 25-gram animal. Plasma was collected after 30 minutes of administration. Plasma was stored at -80°C. Processing was done by Gas Chromatography-Mass Spectrometry (GCMS).

2. *Preparations of reagents from stocks*

a. 6N Hydrochloric acid: To prepare 100ml of 6N HCl from this stock, 49.27 ml of the stock solution was diluted up to a final volume of 100 ml with distilled water.

b. 1M sodium hydroxide: 4g was dissolved in 100 ml water.

c. 5% (v/v) 1,1,3,3-Tetramethoxypropane (MDBMA): 50 µl of MDBMA stock was diluted with 0.95ml of water and shaken well to dissolve. The stock was prepared fresh each day and surplus was discarded.

d. 0.1M Tetrabutylammonium bisulphate (TBA) in 0.1 M Potassium Phosphate buffer, pH 7.4: 3.3395 g of TBA and 1.3609 g of Potassium dihydrogen phosphate were dissolved in 80ml of distilled water. pH was adjusted to 7.4 with KOH and up to a final volume of 100 ml.

e. 0.13M 2,3,4,5,6-Pentafluorobenzyl bromide (PFBBR) in Dichloromethane: 0.1 ml of PFBBR stock was diluted with 5ml of dichloromethane. PFBBR was stored at room temperature in fume cupboard, in a tightly stoppered glass tube.

f. Urea standard stocks: 10mM Urea: 0.0150 g dissolved in 25ml of water. 10mM Urea-13C: 0.01526 g dissolved in 25ml of water. 10mM Urea-13C,15N2: 0.01575 g dissolved in 25ml of water (This is the internal standard stock solution.). All urea standards were stored at -20°C.

3. *Preparation of standards:*

a. Internal standard working solution (1mM final concentration): 0.1ml of the 10mM Stock solutions were mixed with 0.9 ml of water to produce two separate working solutions, which were stored at -20°C.

b. Urea standard curves: Two separate standard curves were prepared, one for unlabelled urea, and a second for urea-13C. Using either the 10mM Urea stock solutions or the 1 mM Urea working solutions, the following standard curve was set up.

<i>Urea (μM)</i>	<i>Vol of 1mM Std (μl)</i>	<i>Vol of Water (μl)</i>
200	200	800
100	100	900
75	75	925
50	50	950
25	25	975
10	10	990
5	100 μ l of 50 μ M	900
2.5	100 μ l of 25 μ M	900
1.0	100 μ l of 10 μ M	900
0.5	100 μ l of 5 μ M	900
0	0	1000

4. ***Procedure on GCMS:***

a. Cyclization: To stoppered glass tubes, the following was added:

- 50 μ l Urea Standard
- 20 μ l of 1mM Urea Internal standard
- 15 μ l Freshly prepared 5% MDBMA
- 200 μ l 6N HCl

The mixture was incubated for 1 hour at room temperature and dried down under Nitrogen gas at room temperature, then the residue was re-dissolved in 100 μ l 1M NaOH.

Derivatization: To the re-dissolved product, 500 μ l of 0.1M TBA was added in phosphate buffer, pH 7.4 and 500 μ l of 0.13M PFBBR in DCM. Tubes were capped, vortexed for 30 seconds then sonicated for 1 hour. 2ml hexane was added, the tubes were shaken then centrifuged 5 mins at 4000rpm.

The upper organic layer was collected to a clean tube, dried down under Nitrogen gas at room temperature and residue was re-dissolved in 100 μ l Hexane. If the samples are cloudy, centrifuge briefly to remove any precipitate before transferring to GC vial.

b. GCMS conditions:

The GC system was programmed with the following temperature gradient for Urea.

<i>Temperature (°C)</i>	<i>Rate (°C/min)</i>
140	-
220	10
300	30

5. ***Other settings:***

- 5 a. Inlet Temperature: 280°C
- b. MS Transfer line: 250°C
- c. Split ratio: 1:12
- d. Carrier gas (Helium) flow rate: 0.8 ml/min
- e. Ion source: 225 °C
- 10 f. Reagent gas: (Methane) flow rate: 2 ml/min. The gas must have been switched on for at least 30 minutes to allow stabilization prior to tuning the instrument and performing chromatography.
- g. Detection Mode: Scan and Selected Ion Monitoring (SIM), collecting data on the following ions:

15

<i>M/Z</i>	<i>Compound detected</i>
94.5 – 95.5	Urea
95.5 – 96.5	Urea+1
97.5 – 98.5	Urea+3 (IS)

2 µl of each prepared sample were injected into the GC system and the data was collected with the XCaliber V 3.0.63 Software.

Immunohistochemistry

- 20 *Liver fixation, sectioning and embedding:* At harvest, livers were fixed in 10% formalin solution for fixing and left at rt for 48 hours before transferring to 70% ethanol solution. The livers were embedded and sectioned, two liver slides per section, 5 µM thickness.

- Staining:* The liver sections were dewaxed in histoclear for 10 mins, shaken at 5 mins, 25 and gradually rehydrated in EtOH. Then, the sections were treated with 3.5mls 30% H₂O₂ in 300 mls MetOH, followed by antigen retrieval using citrate buffer 0.01M. Heat was used to

retrieve antigens. The with slides were blocked with 15% Goat serum in TBST. The slides were incubated overnight at 4°C with primary antibody (ab97370) diluted at 1 in 1000 in 10% Goat serum in TBST. The slides were incubated with polymer-HRP and developed. Finally, the slides were imaged using a Zeiss Axioplan Histology scope.

5 **HILIC-MS/MS for analysis of amino acids - dried blood spot**

This method was adapted from ‘Rapid quantification of underivatized amino acids in plasma by hydrophilic interaction liquid chromatography (HILIC) coupled with tandem mass-spectrometry’ *J. Inherit. Metab. Dis.* 39:651–660 (2016), which is incorporated herein in its entirety.

10 For all the steps outlined below, water (H₂O) refers to fresh Milli-Q® ultrapure water. Guthrie card is spiked with whole blood sample during harvest, dried at room temperature overnight and stored at -20 °C in a foil bag with desiccant.

Sample preparation for DBS:

- a. A 3.2 mm punch from centre of the dried blood spot (equivalent to 3.2 µl blood
15 volume) was placed in a 2 ml glass vial without insert.
- b. 100 µl extraction solution (90 µl methanol + 10 µl mixture of internal standards (IS))
was added to the vial.
Internal standards (IS) refers to the stable isotopes used as internal standard for
quantification.
- 20 c. The 10 µl internal standard (IS) mixture consists of 2 nM L-Arginine-13C₆, L-Citrulline-2,3,3,4,4,5,5-d₇, L-Glutamic-2,3,3,4,4-d₅ Acid and 0.6 nM of L-Glutamine-1,2-13C₂ and L-Ornithine-2,3,3,4,4,5,5-d₇.
- d. The sample was sonicated in a sonication bath for 15 minutes at room temperature.
The supernatant was then transferred on to a 1.5 ml microcentrifuge tube.
- 25 e. The sample is then dried using microcentrifuge speedvac vacuum concentrator
(Settings: room temperature, V-AQ and brakes ‘on’). *IS accounts for any losses in
steps post extraction.*
- f. The dried sample is reconstituted in 40 µl H₂O and vortexed to mix, centrifuged at
13000 rpm for 2 mins and supernatant transferred to a 300 µl insert glass vial.
- 30 g. 40 µl of 0.1M HCl is added to the sample and vortexed to mix.

- h. The sample is topped up with 280 μ l of Solvent A (10 mM ammonium formiate, 85% Acetonitrile (CAN), 15% water (H₂O), 0.15% formic acid) and vortexed to mix before injection into the liquid chromatography-tandem mass spectrometry (LC-MS/MS) system.

5 *Note: Samples post methanol extraction and drying can be stored at -20 °C if mass spectrometry analysis is not performed on the same day as the sample preparation.*

Preparation of calibrators:

- a. Unlabelled amino acids (calibrators) at known concentrations mixed with internal standards were used to produce calibration curve for each run.
- 10 b. All amino acids stocks except glutamine was prepared in 0.1M HCl. Glutamine was IS stocks were prepared in H₂O and were stored at 4 °C.
- c. For each run, 40 μ l of stock mixture of calibrators at various concentrations (0,1,5,10,50 μ M) in 0.1M HCl was mixed with 40 μ l of mixture containing stock glutamine at various concentrations (0,1,5,10,50 μ M) and fixed IS concentration in
- 15 H₂O. The mixture was vortexed and topped with 280 μ l of Solvent A, similar to the sample preparation. The final concentrations of calibrators used are shown in Table 9:

Table 9: Concentration of calibrators used.

Stock concentration μ M	Final concentration μ M	Final pmoles
0	0	0
0.1	0.01	4
0.2	0.02	8
0.5	0.05	10
1	0.1	40
2	0.2	80
5	0.5	200
10	1	400
50	5	2000

Internal standards were used at fixed concentrations at all time for both calibrators and

20 sample preparation. For concentrations refer to section 5.2.

Initial HILIC chromatography and mass spectrometry (MS) setup:

- a. Before samples can be processed through LC-MS/MS system, the column needs to be primed for the HILIC run.

- b. The LC system (solvents and wash tubings) were primed for 3 minutes each at 0.4 ml/min.
- c. Once the column is installed onto the LC system, 10 minutes wash with 95% H₂O is performed at 45 °C column temperature, 0.4 ml/min flow rate.
- 5 d. Following this, 60% Acetonitrile (ACN) and 40% H₂O mixture is flushed through the column at 35 °C column temperature, flow rate of 0.4 ml/min for 50 mins.
- e. The column is then equilibrated with initial conditions (100% Solvent A, 35 °C column temperature at flow rate of 0.4 ml/min) for 30 mins. The column is now primed for the chromatography separation.
- 10 f. For MS/MS system, the cone is cleaned before each run. The cone is initially sonicated in a solution of 50% formic acid:H₂O for 15 mins, followed by further 15 mins sonication in 100% methanol. The cone is then dried before placing back into the MS.
- g. Two gradient runs and one blank run was performed before the calibrators and sample runs were initiated.
- 15

Chromatography separation:

- Amino acid chromatography separation was performed using Acquity UPLC BEH Amide column (2.1 x 100 mm, 1.7 µm particle size) including a Van GuardTM UPLC BEH Amide pre-column (2.1 x 5mm, 1.7 µm particle size) (Waters Limited, UK) with following LC conditions.
- 20 a. Column temperature: 35 °C
 - b. Sample volume injection: 2 µl
 - c. Mobile phase Solvent A: 10 mM ammonium formate in 85 % acetonitrile containing 0.15 % formic acid
 - 25 d. Mobile phase Solvent B: 15 mM ammonium formate in MilliQ-water containing 0.15 % formic acid pH 3.0
 - e. Low Pressure Limit: 0 psi
 - f. High Pressure Limit: 15000 psi
 - g. Seal Wash Period: 3.00 min
 - 30 h. Samples were maintained at 4 °C.
 - i. Gradient conditions for the mobile phase are shown in Table 10.

Table 10: Optimised mobile phase gradient table.

Time (min)	Flow rate (ml/min)	%age A	%age B	Curve
Initial	0.4	100	0	Initial
6	0.4	100	0	6
6.1	0.4	94.1	5.9	6
10	0.4	82.4	17.6	6
13	0.4	70.6	29.4	6
13.1	0.6	70.6	29.4	6
16	0.6	70.6	29.4	6
17	0.6	100	0	6
17.1	0.4	100	0	6
18.5	0.4	100	0	6
18.6	0.4	50	50	6
28	0.4	50	50	6
28.1	0.4	100	0	6
38	0.4	100	0	6

- a. Total run time is 38.00 min. All amino acids are separated in the first 16 minutes, followed by column wash (50% B) for 10 minutes and 10 minutes of re-equilibration with starting (initial) conditions (100% A).
- 5 b. The retention time varies across different day runs as such a fixed retention time cannot be provided. However, the amino acids are separated in the order laid out in Table 2. Samples are spiked with labelled internal standards with different masses but display same retention time acting as a further control for analyte peaks.
- c. Wash Solvent: 90:10 Water:Acetonitrile (ACN)
- 10 d. Pre-Inject Wash Time: 4.0 sec
- e. Post-Inject Wash Time: 6.0 sec
- f. Purge Solvent: 85% ACN
- g. A blank was run after every 10 sample runs to determine any carryovers.

Mass spectrometry parameters:

- 15 a. A XevoTQ-S (Waters) triple quadrupole mass spectrometer with an electrospray ionization (ESI) source and an Acquity UltraPure Liquid Chromatography (UPLC)-system (Waters, Manchester, United Kingdom) were used.

b. MassLynx software (v4.2; Waters, Manchester, United Kingdom) was used for instruments' control and data acquisition.

c. Mass spectrometer operation modes:

- ESI-positive
- Capillary voltage 1.00 kV
- Desolvation temperature 550 °C
- Source temperature 150 °C
- Cone gas flow 150 L/hour
- Desolvation gas flow 500 L/h

d. Multiple reaction monitoring mode, cone voltage and collision energies were used for maximal ion detection of both precursor and product ions as shown in Table 11.

Table 11: Optimised mass spectrometry parameters used for detection of underivatised urea cycle amino acids and their corresponding labelled internal standards in positive ion mode.

Analyte	Precursor ion (m/z)	Product ion (m/z)	Cone Voltage (Volts)	Collision energy (Volts)
Glutamine	147.1	84.1	15	15
L-Glutamine-13C2	149.1	85.1	15	15
Citrulline	176.1	159	15	10
L-Citrulline-d7	183.15	166.05	15	10
Glutamic Acid	148	84	18	16
L-Glutamic Acid-d5	153.06	89.06	18	16
Arginine	175.1	70.1	25	20
L-Arginine-13C6	181.1	76.1	25	20
Ornithine	133.1	70.1	15	15
L-Ornithine-d7	139.98	77.08	15	15
ASA anhydrides	273.2	70.2	25	25
Argininosuccinic Acid (ASA)	291.2	70.2	25	25

Analysis:

- a. Data was acquired using MassLynx v4.2 software.
- b. TargetLynx™ application manager was used for subsequent batch data processing and reporting of results. Application was automated to calculate the standard curve calibration and concentration of the unknown analytes, thereby avoiding biases. Manual correction was performed where necessary provided the peaks are not detected by the software.
- c. All target amino acids (ASA and their anhydrides, arginine, ornithine, citrulline, glutamine and glutamate) were acquired separately. Analysis were combined where necessary.
 - Glutamine can be converted into glutamate spontaneously during storage or via glutaminase activity. Hence, both glutamine and glutamate were measured separately and presented together.
 - Argininosuccinic acid (ASA) can exist as a five-membered ring when acyl group binds an amine group located within the aliphatic chain of molecule. This results in loss of water creating anhydride derivatives. Two forms of anhydrides have been identified which have same molecular mass and are readily interconvertible. Along with ASA, both forms of anhydrides were measured and combined for accurate quantification of ASA levels.
- d. Each analyte have corresponding stable isotope isoforms which is used as their internal standard. However, ASA and ASA anhydrides do not have commercially available stable isotope. So, L-Citrulline-2,3,3,4,4,5,5,-d7 was used as internal standard control for ASA and their anhydrides.
- e. Using the standard curve, pmoles of amino acids per DBS was determined which is then converted to μM .

In vitro ASL activity- fibroblasts***1. Plating cells and treatment.***

- a. Plated ~800,000 cells in a 6 cm TC dish in 4 ml of Complete DMEM media (DMEM supplemented with 10% FBS and 50 units of Penicillin and Streptomycin).
- b. The next day, replaced media and treat as necessary. In case of LNPs, 2.5 μg LNPs in 4 mls of media per well were added. For positive control, a separate dish was treated with equal amount of ASL mRNA.

c. Transfection protocol:

- i. 3 μ l lipofectamine 2000 was added to 250 μ l Opti-MEM media and left at rt for 10 minutes.
- ii. 2.5 μ g unformulated mRNA was added to 250 μ l of Opti-MEM media separately.
- iii. Post 10 minutes, the mRNA+opti-MEM mixture was added to the lipofectamine mixture, mixed and left at rt for 5 minutes.
- iv. 500 μ l mixture was added per well and incubated in TC incubator for 4 hours.
- v. After 4 hours, the transfection mixture was replaced with Complete DMEM media.

d. Left for 48h.

- e. After 48h, cells were harvested using the assay buffer provided in the Abcam fumarate kit.

2. Harvest:

- a. Placed dishes containing cells on ice. Aspirated the media and wash with 4ml of PBS.
- b. Aspirated the PBS and add 250 μ l of assay buffer (fumarate kit).
- c. Placed on a rotating plate shaker for 5 mins on low speed.
- d. After 5 minutes, scraped the dishes using plate scraper and collected the lysed samples in an Eppendorf tube.
- e. Vortexed the tube and leave on ice for 10 minutes.
- f. Centrifuged on maximum speed at 4 °C for 15 minutes.
- g. Transferred the supernatant into a new tube.
- h. Performed BCA assay on the samples to determine protein concentration.

3. ASA incubation and fumarate assay

- a. On ice, took 60 μ g protein per sample and incubate with ASA (final concentration of 300 μ M) in a final volume of 100 μ l.
- b. Negative control: incubated 300 μ M ASA in 200 μ l volume without any samples, to measure the background read.
- c. Incubated the mixture for 15 minutes at 37°C and terminate the reaction at 95C for another 15 minutes.
- d. Centrifuged samples at maximum speed at rt for 5 minutes and transfer the supernatant.

- e. For fumarate reaction, added 50 μ l of supernatant per well (perform in duplicates) and followed fumarate detection protocol described herein.
- f. After 60 minutes incubation per kit instruction, plates were read in Absorbance mode at 450nm in plate reader.
- 5 g. Fumarate concentration (ng/ μ l) was determined per manufacturer's guide.

Preparation of LNPs Containing mRNA Constructs

LNP formulations were prepared using a modified procedure of a method previously described for siRNA. In brief, lipids were dissolved in ethanol at molar ratios of A:B:C:D (ionizable lipid: DSPC:cholesterol:PEG-lipid) corresponding to the molar ratios used in the particular LNP formulation (e.g. LNP-1A, LNP, 1B, LNP-2A, etc.). The lipid mixture was combined with a 25 mM sodium acetate buffer (pH 5) containing mRNA at a ratio of 3:1 (aqueous:ethanol) using a multi-inlet vortex mixer. Formulations were dialyzed against PBS (pH 7.4) in dialysis cassettes for at least 18 hr. Formulations were concentrated using Amicon ultra centrifugal filters (EMD Millipore, Billerica, MA), passed through a 0.22- μ m filter, and stored at 4°C until use. All formulations were tested for particle size, RNA encapsulation, and endotoxin.

[¹⁸F]FSPG PET Production and Procedure

2-3 week-old WT or *Asl^{neo/neo}* mice were anaesthetised and injected with [¹⁸F]FSPG, and dynamically scanned on a preclinical NanoPET/CT plus system (Mediso; Budapest, Hungary). This was done as follows: up to four mice were placed in a warm box (37°C) for circa 10 min; syringes (31G) with 40-60 μ L of a [¹⁸F]FSPG saline solution were prepared and activity measured on a Capintech dose calibrator (Mirion medical, NJ, United States); mice were anaesthetised (O₂ flow rate of 1.00 l/min and isoflurane levels of 1-2.5%) one by one in an induction box and then moved onto a mobile rig for injections; intravenous tail injections were done by illuminating the tail of the mice using a swan neck light and with the help of magnifying glasses. After injection, mice were transferred onto a four-bed animal hotel (Mediso) on the PET/CT. As injections were done in series scanning time was set up to cover the 40-90 minutes post injection for all scanned mice.

In general, following [¹⁸F]FSPG PET, mice were culled by cervical dislocation and liver, skin, and other tissue was collected, snap frozen and moved to a -80°C freezer for later ex vivo analysis.

PET Reconstruction and Data Quantification

PET data was reconstructed in Nucline (v2.01, Bartec Technologies, Farnborough, UK) by dividing each dynamic PET scan into time frames of 10 mins of PET acquisition creating a series of dynamic images spanning circa 30-100 mins post injection (individual differences for the scanned mice). A dynamic iterative reconstruction algorithm, Tera-Tomo 3D ($0.4 \times 0.4 \times 0.4 \text{ mm}^3$ voxel size), was used with attenuation, scatter, and random coincidences correction.

PET and CT images for individual mice were set up by cropping the time frames and the CT image for each PET scan so that each new volume only contained one mouse. Thereafter the data was saved, and the procedure repeated for the remaining mice of that data set using the initial data set. For each mouse, a region of interest (ROI) were drawn over the liver and one over part of the skin of the lower right flank. For each ROI, data for the 10 min frames in the 40-90 mins post injection interval, were extracted.

***Ex vivo* Total Glutathione Quantification**

Frozen liver and skin tissue was thawed on ice and 25-50 mg of tissue was added to prechilled Lysing Matrix D tube (MP Biomedicals) containing ice cold 400 μL 1X passive lysis buffer (Promega; E1941). The tissue was then lysed at 4°C on a Precellys Evolution (Bertin Technologies, Montigny-le-Bretonneux, France); samples were run for five 30s cycles at 6700 RPM. Lysates were centrifuged at $15,000 \times g$ for 10 min at 4°C , and the supernatant was taken for analysis. Total intracellular GSH was determined using the luminescent-based GSH/GSSG-Glo Assay Kit (Promega; V6611) according to manufacturer's instructions in a white 96-well plate prepared with 5 μL of sample supernatant (neat or 1:10 diluted) along with 5 μL of GSH standards (1-100 μM). Results were normalized to protein concentration, determined using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) as per the manufacturer's instructions.

Transcriptomics

RNA was extracted from liver samples using Qiagen RNeasy kits (74004) following kit instructions. 4 liver samples each from the WT, Luc mRNA and hASL mRNA neonatal treated group of *Asl^{Neo/Neo}* mutants were analysed. The samples were sent to UCL genomics for RNA-sequencing using NextSeq 1000/2000 P2 Reagents (100 cycles) v3~16M 2x50bp read pairs/sample.

EXAMPLE 2: *In vitro* expression mRNA sequence selection for hASL

Codon Optimization: Human ASL (hASL) sequence was codon-optimized (SEQ ID NOs:30–33) as different constructs. All constructs (SEQ ID NOs:330–334) were Alpha G5 with 3xmiR142 and screened in SNU423 and HEK293T cells. The expression of hASL was determined by western blot over a range of timepoints (24 to 96 hours). SEQ ID NO:333 (with ORF having SEQ ID NO:33) showed the best expression among all codon optimized variants and was used to build protein variants.

hASL Variant Selection: Through protein sequence analysis, ten hASL variants (SEQ ID NOs:2–11) were selected for their predicted effect on hASL protein stability and ubiquitination: T233V, C307L, A104G (conservation), K266P, G233D, G438N (backbone entropy), K51R, K57R, K80R, and K97R (de-ubiquitination). (All substitutions are with respect to the wild-type human ASL sequence (SEQ ID NO:1) and are disclosed above in Table 2. Table 8 shows the constructs incorporating the ORFs encoding the hASL polypeptide variants.) mRNA sequences encoding the hASL variants were screened in SNU423 and HEK293T cells, and SEQ ID NO:20 showed the highest expression at 96 hours (FIG. 1A) and was used to build UTR variations.

UTR Variant Selection: Different 5'UTR and 3'UTR variants of the K51R variant were tested for expression. Namely (5'UTR SEQ ID NO:50, 3'UTR SEQ ID NO:128; 5'UTR SEQ ID NO:78, 3'UTR SEQ ID NO:137; 5'UTR SEQ ID NO:50, 3'UTR SEQ ID NO:138; and 5'UTR SEQ ID NO:78, 3'UTR SEQ ID NO:139). UTRS were cloned as the UTR of the CO7 K51R variant. Western blot analysis showed that the construct according to SEQ ID NO:300 (ORF SEQ ID NO:20; 5'UTR SEQ ID NO:50; 3'UTR SEQ ID NO:108) trended to have higher expression (FIG. 1B).

EXAMPLE 3: ASL protein produces fumarate with Argininosuccinic Acid incubation in HEK293 cells transfected with ASL mRNA

Using recombinant ASL protein, Fumarate detection assay was optimized. In the presence of ASL protein Fumarate production increases with Argininosuccinic acid incubation.

HEK293 cells transfected with ASL mRNA produced fumarate with Argininosuccinic Acid incubation, whereas cells transfected with eGFP mRNA did not produce detectable levels of fumarate (FIG.2).

EXAMPLE 4: ASL Hypomorph Mice (ASL^{tm1Brle}) Experiments

5 The ASL Hypomorph Mice (ASL^{tm1Brle}) has a targeted mutation comprising an FRT-flanked neomycin cassette insterted into intron 9 of the ASL gene that decreases expression and activity of the gene. The hypomorph mice express about 15-20% of the normal levels of ASL protein and die at 3-4 weeks of age. The hypomorph mice show severe postnatal growth retardation, abnormal hair patterning, multiple organ system dysfunction,
10 hyperammonemia, and nitric oxide deficiency. Histological defects are found in multiple organs, involving the immune, hematopoietic, and cardiovascular systems. Heterozygous animals have no phenotype.

In order to test the effect of expressing hASL in the hypomorph mice, 1 miligrams per kilogram (mpk) LNP-1A with SEQ ID NO:334 mRNA construct was administered to the
15 hypomorph mice once every two weeks. The hypomorphs that received the hASL (ASL HOM) showed substantial weight gain, but the weight gain was still about 50% less than that of their wild type littermates. ASL HOM also showed restoration of coat, normal locomotion, and behavior. The survival rate of ASL HOM have increased to between 50 and 180 days (FIG. 3) (as opposed to ~3/4 weeks).

20 EXAMPLE 5: *In vitro* study of human fibroblasts

Healthy and ASL patient fibroblasts were compared *in vitro*. The genotype of Patient 1 fibroblasts was c.437G>A / c.437G>A; R146Q / R146Q, and the genotype of Patient 2 fibroblasts was c.719-2A>G / c.857A>G; ? / GIN286Arg.

At the outset, the ASL levels were measured with in-cell western assays, as described
25 in Example 1. Basal ASL levels in patient fibroblasts were lower than that of the wild type control fibroblasts (FIG. 4A). The ASL levels in patient fibroblasts were about 50% of their wild type counterparts (FIG. 4A).

Healthy (WT) and Patient fibroblasts were transfected with luciferase mRNA packaged in lipid nanoparticles (Luc-LNP), ASL mRNA in packaged in lipid nanoparticles

(ASL-LNP), naked ASL mRNA or vehicle (PBS). ASL levels increased *in vitro* post ASL mRNA treatment (FIG. 4B). ASL activity was measured using a fumarate detection assay 48 hours after ASL mRNA treatment. ASL activity increased *in vitro* post ASL mRNA treatment (FIG. 4C).

EXAMPLE 6: *In vivo* Studies with Mild Phenotype (late death) ASA Mice

This *in vivo* study was conducted over 4 months (120 days). Mice with mild phenotype argininosuccinic aciduria (mild phenotype ASA mice) were administered v1 (LNP-1A with SEQ ID NO:334) or v2 (LNP-3A with SEQ ID NO:300 (without idT)) ASL mRNA (ASL-LNP) constructs. or a luciferase-LNP (Luc-LNP, in LNP-1A) control. On Day 1/Day 2, ASL-LNP v1 or Luc-LNP were administered by intravenous (IV) injection. On Days 7, 14, 21 and 28, ASL-LNP v1 or Luc-LNP were administered by intraperitoneal (IP) injection. From Day 35 onward, ASL-LNP v2 or Luc-LNP were administered by IV injection weekly (every 7 days). Mice were sacrificed on Day 120, 7-10 days after last injection.

Administration of ASL-LNP partially rescued the growth of the mice, as compared to Luc-LNP administered mice (FIG. 5A). Survival of the mice ASL-LNP administered mice was also significantly improved (FIG. 5B). Overall, significant improvement in growth and survival was observed in mild phenotype ASA mice. Increased ASL levels were also detected.

EXAMPLE 7: Pharmacokinetics/Pharmacodynamics Studies

A third generation hASL mRNA construct (SEQ ID NO:300 (+ idT)) was used for the pharmacokinetics/pharmacodynamics studies. Mild phenotype ASA mice were IV injected a single dose of a LNP (LNP-3A) comprising SEQ ID NO:300 (+ idT) or Luc-LNP (in LNP-3A) at 3-4 weeks of age. Mice were sacrificed at timepoints 2 hours, 24 hours, 72 hours or 7 days after the injection. Liver ASL levels were measured using western blot as described herein in Example 1. ASL activity was measured using fumarate assay as described herein in Example 1. Plasma ammonia and ALT levels were measured as described herein in Example 1. Orotate levels were measured using mass spectroscopy as described herein in Example 1.

Results: Liver ASL protein levels were restored to near-normal levels 24 hours after ASL-LNP v3 (SEQ ID NO: 300 (+ idT) in LNP-3A) injection, however, liver ASL protein levels dropped at 72-hour and 7-day timepoints (FIG. 6A).

Histological analysis of liver slides also showed a similar result: percentage of ASL-positive regions in liver sections was significantly higher at 24 hours, but 72-hour and 7-day time points were not different from control in a statistically significant manner (FIG. 6B).

Similarly, ASL activity levels were restored to near-normal levels 24 hours after ASL-LNP v3, and liver ASL activity levels dropped at 72-hour and 7-day timepoints (FIG. 6C).

Next, plasma ammonia, Argininosuccinic acid (ASA) and citrulline levels were measured. Beginning from the 24 hour timepoint and continuing through the 72-hour and 7-day timepoints, plasma ammonia (FIG. 6D), Argininosuccinic acid (FIG. 6E) and citrulline levels (FIG. 6F) were significantly lower in ASL-LNP treated mice, as compared to Luc-LNP treated mice.

Finally, plasma orotate (orotic acid) levels were measured 24 hours post-administration. ASL-LNP-treated mice showed significantly low amounts of orotate (similar to WT levels) as compared to Luc-LNP-treated mice (FIG. 6G).

15 **EXAMPLE 8: *In vivo* Studies with Severe (early death) Phenotype ASA Mice**

This *in vivo* study was conducted over 48 days. Mice with severe phenotype argininosuccinic aciduria (severe phenotype ASA mice) were administered ASL mRNA (SEQ ID NO: 300 (+ idT) in LNP-3A) (ASL-LNP) constructs, or a luciferase-LNP (Luc-LNP) control (in LNP-3A). On Day 1/Day 2, ASL-LNP v3 or Luc-LNP were administered by IV injection. From Days 7 onward, ASL-LNP v3 or Luc-LNP were administered by IV injection weekly (every 7 days). Mice were sacrificed on Day 48, 48 hours after last injection.

Administration of ASL-LNP v3 completely rescued the growth of the severe phenotype mice, as compared to Luc-LNP control (FIG. 7A). In addition, survival of the ASL-LNP administered mice was also significantly improved, all ASL-treated mice surviving up to 50 days (FIG. 7B). Overall, significant improvement in growth and survival was observed in severe phenotype ASA mice. Liver ASL protein levels (FIG. 7C) and liver ASL activity (FIG. 7D) were increased significantly higher to normal levels in ASL-LNP-treated mice. Plasma ammonia and amino acid (ASA + ASA anhydrite and citrulline) were decreased to normal (WT) levels after ASL-LNP treatment (FIGS. 7E-7G). ALT levels were

increased to normal/above normal levels (FIG. 7H). Plasma amino acid and ALT levels were measured as described herein in Example 1.

Next, a longitudinal analysis of plasma amino acid levels over 50 days was performed. In ASL-LNP-treated mice, average ASA levels (FIG. 7I) and average citrulline levels (FIG. 7J) were consistently low (as low as wild type levels, when there was a wild type measurement) throughout the observation period.

In ASL-LNP-treated mice, average orotate levels were decreased to non-detectable levels, similar to wild-type mice (FIG. 7K, 7L).

Next, a C¹³ ureagenesis assay was performed on mice treated with ASL-LNP v3 as described herein in Example 1. C¹³ ureagenesis was restored completely to normal levels upon ASL-LNP treatment (FIG. 7M).

In conclusion, longer survival and improved growth were observed in the survival study with ASL-LNP v2 construct (Example 6). On the other hand, single intravenous administration of ASL-LNP v3 construct showed restoration of ASL levels and improvement in biomarkers up to 7 days (Example 7). Rescue of survival and growth in severe ASL mice was observed following weekly injection of ASL-LNP v3 construct. Rescue of disease biomarkers was also observed with the ASL-LNP v3 administration.

EXAMPLE 9: *In vivo* Studies with ASA Mice

ASA mice were administered ASL mRNA (SEQ ID NO: 300 (+ idT) in LNP-3A) (ASL-LNP) constructs, or a luciferase-LNP-3A (Luc-LNP) control. Beginning on day 21, ASL-LNP v3 or Luc-LNP were administered by IV injection (1 mg/kg bodyweight) once per week.

Administration of ASL-LNP v3 greatly improved the growth of the ASA mice, as compared to Luc-LNP control (FIG. 8A). In addition, survival of the ASL-LNP administered mice was also significantly improved, all but one of the ASL-treated mice surviving up to the study endpoint (60 days) (FIG. 8B). Overall, significant improvement in growth and survival was observed in ASA mice. FIGs. 9A-C show the improvement in growth for ASL-LNP administered mice when compared to the Luc-LNP cohort. The growth was nearly the same as for the wildtype cohort.

FIGs. 10A-D show growth velocity curves for the ASL-LNP-administered mice versus the Luc-LNP-administered mice. The growth velocity data corroborates the improvement in growth observed for ASL-LNP-administered ASA mice versus the Luc-LNP-administered cohort.

5 **EXAMPLE 10: *In vivo* Studies - ASL Levels and Activity After Neonatal ICV Administration in WT CD-1 Mice**

10 Liver ASL protein levels and ASL activity were assessed after single intracerebroventricular (ICV) administration of ASL-LNP v3 (SEQ ID NO:300 (+ idT) in LNP-3A) versus Luc-LNP (in LNP-3A) administration in neonatal WT CD-1 mice. At 24 hours post-administration, liver ASL protein levels (FIG. 11A) and liver ASL activity (FIG. 11B) were increased significantly higher compared to normal levels in ASL-LNP-treated mice, with the ASL kinetics approaching (but still significantly higher than) baseline levels after 7 days, 2 weeks, and 3 weeks.

15 **EXAMPLE 11: *In vivo* evidence of impaired glutathione metabolism in *Asl*^{Neo/Neo} mice**

The biosynthesis of glutathione is dependent on cellular import of cystine in exchange of the efflux of glutamate via cystine/glutamate antiporter (xCT), a transmembrane transport protein. The protein expression of xCT in 2 week-old *Asl*^{Neo/Neo} mice was observed to be significantly higher than that of their WT littermates (FIG. 12).

20 [¹⁸F]FSPG is a glutamate analog, positron emission tomography (PET) tracer and non-invasive marker of cellular redox balance and xCT activity, which has been used in the clinic to assess chemoresistance to cancer drugs. [¹⁸F]FSPG was administered intravenously to 2-3 week-old *Asl*^{Neo/Neo} mice and WT littermates, and uptake/retention was then dynamically imaged using PET coupled to computerised tomography. Liver [¹⁸F]FSPG retention for WT mice (5.2 ± 1.5 %ID/g) was significantly lower (*p*=0.002) than that observed in *Asl*^{Neo/Neo} mice (14 ± 4 %ID/g). Consequently, the liver was barely visualized, images being dominated by pancreatic and kidney activity, whereas it was difficult to distinguish between pancreatic and liver uptake in *Asl*^{Neo/Neo} mice (FIG. 13A-B). High [¹⁸F]FSPG retention was also observed in the skin of *Asl*^{Neo/Neo} mice (13 ± 1.8 %ID/g), which was not the case in WT littermates (5.3 ± 2.3 %ID/g) (FIG. 13B). These results were

30

consistent with results showing impaired glutathione metabolism via increased intracellular cystine flux through upregulated expression of xCT antiporter.

EXAMPLE 12: *hASL* mRNA therapy corrects the metabolic dysfunction and

5 liver pathophysiology in *Asl*^{Neo/Neo} mice

To determine the extent of the correction of liver metabolic dysfunction following *hASL* mRNA treatment in *Asl*^{Neo/Neo} mice, transcriptomic analyses using RNA-sequencing (RNA-seq) were performed. Differences in gene expression between WT and *Asl*^{Neo/Neo} mice treated at birth with either *hASL* mRNA or *Luc* mRNA therapy were studied. Principal component analysis comparing the PC1 and PC2 variances showed analogous clustering of the *hASL* mRNA treated and WT liver samples, suggesting a similar profile of gene expression in both groups. In contrast, *Luc* mRNA treated *Asl*^{Neo/Neo} mice grouped further apart (FIG. 14). Volcano plots were used to illustrate genes that were significantly upregulated (red) or downregulated (blue) between WT, *Luc* mRNA and *hASL* mRNA *Asl*^{Neo/Neo} groups. Comparing WT vs *Luc* mRNA *Asl*^{Neo/Neo} groups, 2705 genes were highlighted to be significantly upregulated (1257 genes) or downregulated (1448 genes) (FIG. 15A). Remarkably and unexpectedly, only 7 genes (1 upregulated and 6 downregulated) were significantly different between WT vs *hASL* mRNA *Asl*^{Neo/Neo} mice livers, demonstrating the efficacy of mRNA therapy in correcting liver dysfunction (FIG. 15B). Reiterating this, *Luc* mRNA vs *hASL* mRNA *Asl*^{Neo/Neo} comparison showed plot profile similar to the WT vs *Luc* mRNA *Asl*^{Neo/Neo} comparison (FIG. 15C) with expression of 4297 genes being significantly altered (1962 genes upregulated and 2335 downregulated). Of note, the murine *Asl* gene was significantly downregulated in both groups of *Asl*^{Neo/Neo} livers, likely due to the codon optimised mRNA sequence not being recognised by the database which utilised WT *hASL* cDNA sequence alignment and/or translation of a majority of the administered therapeutic ASL mRNA at 48 hours post injection.

To further study the dysregulation of glutathione function, an analysis of the pathways affecting glutathione metabolism was performed on the RNA-seq data. The analysis highlighted downregulation of multiple genes involved in glutathione biosynthesis and metabolism alongside alterations of genes of the methionine cycle, transsulfuration and antioxidant pathways between *Luc* mRNA (control) *Asl*^{Neo/Neo} mice and WT livers (FIG. 16).

These findings additionally support disruption of glutathione metabolism in *Asl^{Neo/Neo}* mice. These pathways were corrected post- *hASL* mRNA treatment, as shown in the post pathways analysis comparison between *hASL* mRNA and *Luc* mRNA treated *Asl^{Neo/Neo}* mice (FIG. 16).

EXAMPLE 13: *hASL* mRNA therapy corrects the dysfunction of glutathione metabolism in *Asl^{Neo/Neo}* mice

As described in Example 12 above, downregulation of multiple genes involved in glutathione biosynthesis and metabolism was observed in the *Asl^{Neo/Neo}* mice. Indeed, total glutathione (GSH) levels in *Luc* mRNA treated *Asl^{Neo/Neo}* mice were observed to be lower than those of WT mice (FIG. 17). Following *hASL* mRNA administration in both neonatal and adult treated *Asl^{Neo/Neo}* mice, liver glutathione levels were restored to levels that were not significantly different to WT levels (FIG. 17).

To investigate the potential of [¹⁸F]FSPG PET as a non-invasive tool in assessing therapeutic efficacy, [¹⁸F]FSPG was administered intravenously to 2 weeks-old untreated and *hASL* mRNA treated *Asl^{Neo/Neo}* livers. *Asl^{Neo/Neo}* mice were given IV administration of 1mg/kg *hASL* mRNA at birth followed by weekly IP administration of 2mg/kg mRNA before imaging at 2 weeks of age. Remarkably, there was a significant reduction in the liver [¹⁸F]FSPG retention in *hASL* mRNA treated (11 ± 2.0 %ID/g) versus untreated *Asl^{Neo/Neo}* mice (22 ± 2.3 %ID/g; $p = 0.026$) (FIG. 18A-B), in line with reduced glutathione levels observed. This did not normalise to [¹⁸F]FSPG retention baseline levels from WT livers (5.0 ± 2.8 %ID/g). However, skin [¹⁸F]FSPG retention was not affected by mRNA therapy. [¹⁸F]FSPG skin retention was 4.2 ± 3.4 %ID/g in WT mice, and 15 ± 3.7 %ID/g and 15 ± 4.2 %ID/g in untreated and *hASL* mRNA treated *Asl^{Neo/Neo}* mice, respectively (FIG. 19A-B).

Additionally, and in line with improvement of glutathione metabolism and [¹⁸F]FSPG retention, the expression of cystine/glutamate antiporter system xCT was massively reduced in livers from *hASL* mRNA treated versus *Asl^{Neo/Neo}* mice (FIG. 20).

Glutathione metabolism was corrected with corresponding with a restoration of hepatic liver glutathione for both neonatal and adult treated *Asl^{Neo/Neo}* mice similar to that of WT levels (See FIG. 17). This restoration of glutathione levels was associated with a significant reduction of total homocysteine ratio compared to WT in liver from *hASL* mRNA treated versus *Asl^{Neo/Neo}* mice (FIG. 21).

It is to be appreciated that the Detailed Description section, and not the Summary and Abstract sections, is intended to be used to interpret the claims. The Summary and Abstract sections can set forth one or more but not all exemplary embodiments of the present disclosure as contemplated by the inventor(s), and thus, are not intended to limit the present disclosure and the appended claims in any way.

The present disclosure has been described above with the aid of functional building blocks illustrating the implementation of specified functions and relationships thereof. The boundaries of these functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternate boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed.

The foregoing description of the specific embodiments will so fully reveal the general nature of the present disclosure that others can, by applying knowledge within the skill of the art, readily modify and/or adapt for various applications such specific embodiments, without undue experimentation, without departing from the general concept of the present disclosure. Therefore, such adaptations and modifications are intended to be within the meaning and range of equivalents of the disclosed embodiments, based on the teaching and guidance presented herein. It is to be understood that the phraseology or terminology herein is for the purpose of description and not of limitation, such that the terminology or phraseology of the present specification is to be interpreted by the skilled artisan in light of the teachings and guidance.

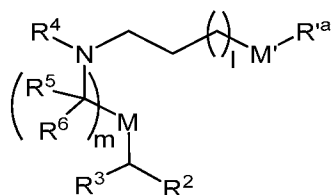
The breadth and scope of the present disclosure should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with the following claims and their equivalents.

All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

WHAT IS CLAIMED IS:

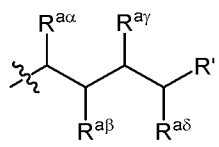
1. A lipid nanoparticle comprising:

(a) an ionizable lipid of Formula (I):



(I) or its N-oxide, or a salt or isomer thereof,

wherein R'^a is R['] branched,



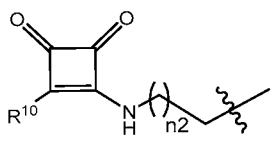
wherein R['] branched is:

wherein  denotes a point of attachment;

wherein R^{aα}, R^{aβ}, R^{aγ}, and R^{aδ} are each independently selected from the group consisting of H, C₂₋₁₂ alkyl, and C₂₋₁₂ alkenyl;

R² and R³ are each independently selected from the group consisting of C₁₋₁₄ alkyl and C₂₋₁₄ alkenyl;

R⁴ is selected from the group consisting of -(CH₂)_nOH, wherein n is selected from

the group consisting of 1, 2, 3, 4, and 5, and ,

wherein  denotes a point of attachment;

wherein R¹⁰ is N(R)₂; each R is independently selected from the group consisting of C₁₋₆ alkyl, C₂₋₃ alkenyl, and H;

wherein n₂ is selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10;

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

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

M and M' are each independently selected from the group consisting of -C(O)O- and -OC(O)-;

R' is a C₁₋₁₂ alkyl or C₂₋₁₂ alkenyl;

1 is selected from the group consisting of 1, 2, 3, 4, and 5; and

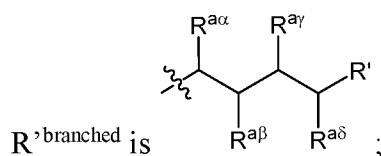
m is selected from the group consisting of 5, 6, 7, 8, 9, 10, 11, 12, and 13; and

(b) an mRNA molecule comprising an open reading frame (ORF) encoding an ASL

5 polypeptide having an amino acid sequence identical to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5.

2. The lipid nanoparticle of claim 1, wherein in the compound of Formula (I):

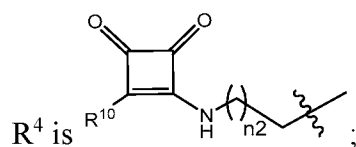
R^a is R^b -branched.



$\frac{\zeta}{\zeta}$ denotes a point of attachment; and

R^{aα} is C₂₋₁₂ alkyl; R^{aβ}, R^{aγ}, and R^{aδ} are each H;

R² and R³ are each C₁₋₁₄ alkyl;



R¹⁰ is NH(C₁₋₆ alkyl);

n2 is 2;

each R^5 is H ;

each R^6 is H ;

M and M' are each -C(O)O-;

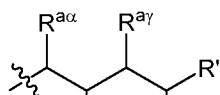
20 R' is a C₁₋₁₂ alkyl;

1 is 5; and

m is 7.

3. The lipid nanoparticle of claim 1, wherein in the compound of Formula (I):

R^a is R[']branched;



R[']branched is $\text{---}\overline{\text{S}}\text{---}$;

$\text{---}\overline{\text{S}}\text{---}$ denotes a point of attachment;

5 R^{a α} , R^{a β} , R^{a γ} , and R^{a δ} are each H;

R² and R³ are each C₁₋₁₄ alkyl;

R⁴ is -(CH₂)_nOH; n is 2;

each R⁵ is H;

each R⁶ is H;

10 M and M' are each -C(O)O-;

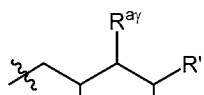
R' is a C₁₋₁₂ alkyl;

l is 5; and

m is 7.

- 15 4. The lipid nanoparticle of claim 1, wherein in the compound of Formula (I):

R^a is R[']branched;



R[']branched is $\text{---}\overline{\text{S}}\text{---}$;

$\text{---}\overline{\text{S}}\text{---}$ denotes a point of attachment;

R^{a β} and R^{a δ} are each H;

20 R^{a γ} is C₂₋₁₂ alkyl;

R² and R³ are each C₁₋₁₄ alkyl;

R⁴ is -(CH₂)_nOH; n is 2;

each R⁵ is H;

each R⁶ is H;

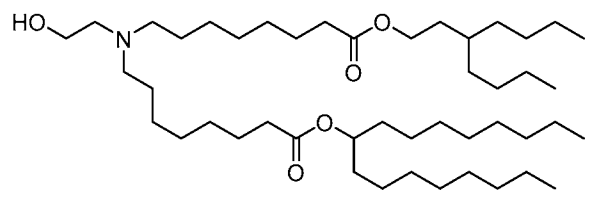
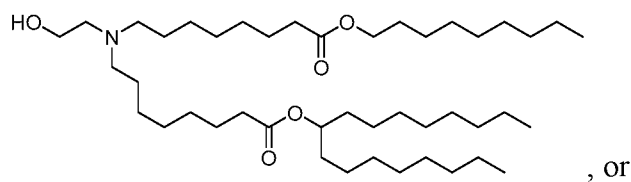
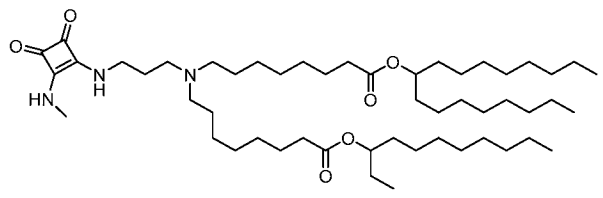
25 M and M' are each -C(O)O-;

R' is a C₁₋₁₂ alkyl;

l is 5; and

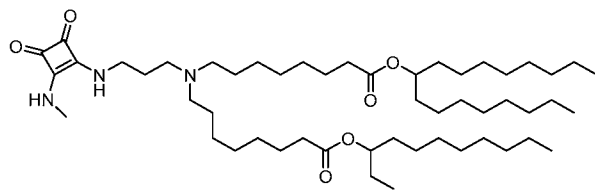
m is 7.

5. The lipid nanoparticle of claim 1, wherein the ionizable lipid is:



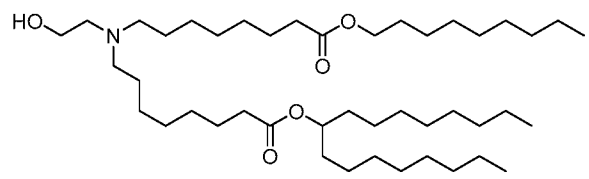
or N-oxides, salts, or isomers thereof.

6. The lipid nanoparticle of claim 1, wherein the ionizable lipid is



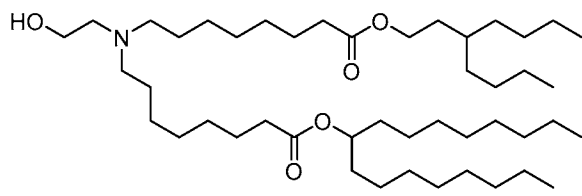
or an N-oxide, salt, or isomer thereof.

7. The lipid nanoparticle of claim 1, wherein the ionizable lipid is



or an N-oxide, salt, or isomer thereof.

8. The lipid nanoparticle of claim 1, wherein the ionizable lipid is



or an N-oxide, salt, or isomer thereof.

9. The lipid nanoparticle of any one of claims 1-8, further comprising:
a phospholipid;
a structural lipid; and
a PEG-lipid.
10. The lipid nanoparticle of claim 9, wherein the lipid nanoparticle comprises:
40-50 mol% of the ionizable lipid, 30-45 mol% of the structural lipid, 5-15 mol% of the phospholipid, and 1-5 mol% of the PEG-lipid.
11. The lipid nanoparticle of claim 9 or claim 10, wherein the lipid nanoparticle comprises:
45-50 mol.% of the ionizable amino lipid, 8-12 mol.% of the phospholipid, 35-40 mol.% of the structural lipid, and 1.5-3.5 mol.% of the PEG-modified lipid.
12. The lipid nanoparticle of any one of claims 9-11, wherein the phospholipid comprises
comprises 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-3-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-3-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinenoyl-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-dilinenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-

sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and mixtures thereof.

5 13. The lipid nanoparticle of any one of claims 9-12, wherein the phospholipid comprises DSPC.

14. The lipid nanoparticle of any one of claims 9-13, wherein the structural lipid is selected from the group consisting of: cholesterol, fecosterol, sitosterol, ergosterol, campesterol,
10 stigmasterol, brassicasterol, tomatidine, tomatine, ursolic acid, alpha-tocopherol, hopanoids, phytosterols, steroids, derivatives thereof, and mixtures thereof.

15. The lipid nanoparticle of any one of claims 9-14, wherein the structural lipid comprises cholesterol or a derivative thereof.

15 16. The lipid nanoparticle of any one of claims 9-15, wherein the PEG-lipid is selected from: 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol (PEG-DMG), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)] (PEG-DSPE), PEG-disteryl glycerol (PEG-DSG), PEG-dipalmetoleyl, PEG-dioleoyl, PEG-distearoyl, PEG-diacylglycamide
20 (PEG-DAG), PEG-dipalmitoyl phosphatidylethanolamine (PEG-DPPE), PEG-1,2-dimyristyloxylpropyl-3-amine (PEG-c-DMA), or combinations thereof.

17. The lipid nanoparticle of any one of claims 9-16, wherein the PEG-lipid comprises PEG-DMG.

25 18. The lipid nanoparticle of any one of claims 1-17, wherein the mRNA molecule, when administered as a single dose to a human subject, is sufficient to:

- (i) reduce plasma ammonia levels in the human subject to a level of less than or equal to about 500 μ M, less than or equal to about 400 μ M, less than or equal to about 300 μ M,
30 less than or equal to about 200 μ M, less than or equal to about 100 μ M, or less than or equal to about 50 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or
- (ii) reduce plasma citrulline levels in the human subject to a less than or equal to about 100 μ M, less than or equal to about 50 μ M, less than or equal to about 25 μ M, or less than

or equal to about 10 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(iii) reduce plasma argininosuccinic acid (ASA) levels to less than or equal to about 150 μ M, less than or equal to about 100 μ M, less than or equal to about 50 μ M, or less than or equal to about 25 μ M, or less than or equal to about 10 μ M, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration.

19. The lipid nanoparticle of any one of claims 1-18, wherein the mRNA molecule, when administered as a single dose to a human subject, is sufficient to:

(iv) increase cellular ASL levels in the human subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.5-fold, or at least about 2.0-fold, compared to the human subject's baseline cellular ASL levels for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120 hours, or at least one week post-administration; or

(v) increase ASL activity in the human subject by at least about 1.5-fold, at least about 2.0-fold, or at least about 2.5-fold, compared to the human subject's baseline ASL activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(vi) reduce plasma orotate levels to less than or equal to about 150 μ mol/mmol creatinine, less than or equal to about 100 μ mol/mmol creatinine, less than or equal to about 50 μ mol/mmol creatinine, less than or equal to about 25 μ mol/mmol creatinine, or less than or equal to about 10 μ mol/mmol creatinine, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration.

20. The lipid nanoparticle of any one of claims 1-19, wherein the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:1.

21. The lipid nanoparticle of any one of claims 1-19, wherein the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:2.

22. The lipid nanoparticle of any one of claims 1-19, wherein the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:3.

23. The lipid nanoparticle of any one of claims 1-19, wherein the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:4.

24. The lipid nanoparticle of any one of claims 1-19, wherein the ASL polypeptide encoded by the mRNA molecule has an amino acid sequence identical to SEQ ID NO:5.

25. The lipid nanoparticle of any one of claims 1-24, wherein the ORF is at least 65% identical to SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, or SEQ ID NO:33.

26. The lipid nanoparticle of any one of claims 1-25, wherein the ORF is at least 65% identical to SEQ ID NO:20.

27. The lipid nanoparticle of any one of claims 1-26, wherein the mRNA molecule further comprises a 5'UTR having a nucleic acid sequence according to SEQ ID NO:50, SEQ ID NO:55, SEQ ID NO:56, or SEQ ID NO:78.

28. The lipid nanoparticle of any one of claims 1-27, wherein the mRNA molecule further comprises a 3'UTR having a nucleic acid sequence according to SEQ ID NO:108, SEQ ID NO:111, SEQ ID NO:128, SEQ ID NO:137, SEQ ID NO:138, or SEQ ID NO:139.

29. The lipid nanoparticle of any one of claims 1-28, wherein the mRNA molecule comprises a poly-A region, wherein the poly-A region about 100 nucleotides in length.

30. The lipid nanoparticle of any one of claims 1-29, wherein the mRNA molecule comprises at least one chemically modified nucleobase, sugar, backbone, or any combination thereof.

31. The lipid nanoparticle of claim 30, wherein the at least one chemically modified nucleobase is selected from the group consisting of pseudouracil (ψ), N1-methylpseudouracil ($m1\psi$), 1-ethylpseudouracil, 2-thiouracil ($s2U$), 4'-thiouracil, 5-methylcytosine, 5-methyluracil, 5-methoxyuracil, and any combination thereof.

32. The lipid nanoparticle of any one of claims 1-31, wherein the mRNA molecule further comprises a 5' terminal cap.

33. The lipid nanoparticle of claim 32, wherein the 5' terminal cap comprises a m7G-ppp-Gm-AG, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azidoguanosine, Cap2, Cap4, 5' methylG cap, or an analog thereof.

5

34. The lipid nanoparticle of claim 33, wherein the 5' terminal cap comprises m7G-ppp-Gm-AG or Cap1.

35. A pharmaceutical composition comprising:

10 the lipid nanoparticle of any one of claims 1-34; and
a pharmaceutically acceptable carrier.

36. A pharmaceutical composition according to claim 35 for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing
15 plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, reducing plasma orotate levels, or reducing ureagenesis in a subject.

37. A lipid nanoparticle according to any one of claims 1-34 for use in the manufacture of a
20 medicament or pharmaceutical composition for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, reducing plasma orotate levels, or reducing ureagenesis in a subject.

25 38. A method of treating or delaying the onset and/or progression of argininosuccinic aciduria in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

30 39. A method of reducing plasma ammonia levels in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

40. A method of reducing plasma argininosuccinic acid (ASA) levels in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

41. A method of reducing plasma citrulline levels in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

42. A method of increasing ASL levels in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

43. A method of increasing ASL activity in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

44. A method of reducing orotate levels in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

45. A method of reducing ureagenesis in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

46. The method of any one of claims 39-45, wherein the administration is a single administration.

47. The method of any one of claims 39-45, wherein the administration is a repeated administration.

48. The method of claim 47, wherein the administration is about once per day, about once per week, about once per two weeks, or about once per month.

49. The method of any one of claims 39-48, wherein the pharmaceutical composition is administered intravenously, intramuscularly, or subcutaneously.

50. An ASL polypeptide having an amino acid sequence identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11.

51. The ASL polypeptide of claim 50, wherein the amino acid sequence is identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, or SEQ ID NO:5.

52. The ASL polypeptide of claim 50, wherein the amino acid sequence is identical to SEQ ID NO:2.

53. A pharmaceutical composition comprising:
the ASL polypeptide of any one of claims 50-52; and
a pharmaceutically acceptable carrier.

54. A cell comprising the ASL polypeptide according to any one of claims 50-52.

55. A codon-optimized mRNA molecule encoding the ASL polypeptide according to any one of claims 50-52.

56. The codon-optimized mRNA molecule of claim 55, having an open reading frame (ORF) having at least about 60% identity to the nucleic acid sequence of SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, or SEQ ID NO:29.

57. The codon-optimized mRNA molecule of claim 55 or 56, wherein the ORF has at least about 60% identity to the nucleic acid sequence of SEQ ID NO:20.

58. A cell comprising the nucleic acid construct of claim 56 or claim 57.

59. The lipid nanoparticle of any one of claims 1-19, wherein the mRNA molecule, when administered as a single dose to a human subject, is sufficient to:

(i) increase liver glutathione levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(ii) reduce liver xCT antiporter levels in the subject by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 92%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater, for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(iii) reduce liver xCT antiporter activity in the subject in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline ASL activity for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration;

(iv) reduce total plasma homocysteine (HCyS) levels in the subject by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-fold, at least about 5.0-fold, or greater, compared to the subject's baseline total plasma HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration; or

(v) reduce total liver homocysteine (HCyS) levels in the subject at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 2.0-fold, at least about 2.5-fold, at least about 3.0-fold, at least about 4.0-

fold, at least about 5.0-fold, or greater, compared to the subject's baseline total liver HCyS level for at least 12 hours, at least 24 hours, at least 48 hours, at least 72 hours, at least 96 hours, at least 120, or at least one week post-administration.

60. A pharmaceutical composition according to claim 35 for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, increasing glutathione levels, increasing liver glutathione levels, reducing plasma orotate levels, reducing ureagenesis, reducing liver xCT antiporter levels, reducing liver xCT antiporter activity, reducing total plasma HCyS levels, or reducing total liver HCyS levels in a subject.

61. A lipid nanoparticle according to any one of claims 1-34 for use in the manufacture of a medicament or pharmaceutical composition for use in treating or delaying the onset and/or progression of argininosuccinic aciduria, reducing plasma ammonia levels, reducing plasma argininosuccinic acid (ASA) levels, reducing plasma citrulline levels, increasing ASL levels, increasing ASL activity, increasing glutathione levels, increasing liver glutathione levels, reducing plasma orotate levels, reducing ureagenesis, reducing liver xCT antiporter levels, reducing liver xCT antiporter activity, reducing total plasma HCyS levels, or reducing total liver HCyS levels in a subject.

62. A method of increasing glutathione levels in a human subject, comprising:
administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

63. A method of increasing liver glutathione levels in a human subject, comprising:
administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

64. A method of reducing liver xCT antiporter levels in a human subject, comprising:
administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

65. A method of reducing liver xCT antiporter activity in a human subject, comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

66. A method of reducing total plasma homocysteine (HCyS) levels in a human subject,
5 comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

67. A method of reducing total liver homocysteine (HCyS) levels in a human subject,
10 comprising:

administering to the human subject an effective amount of a pharmaceutical composition according to claim 35.

15

FIG. 1A

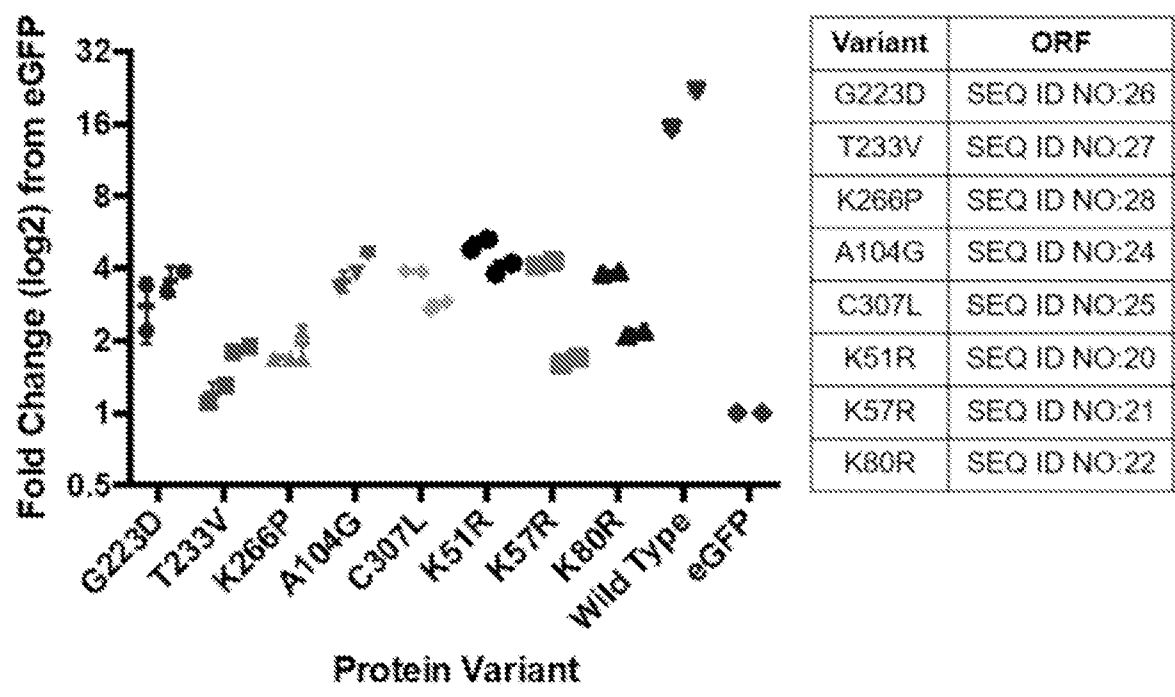


FIG. 1B

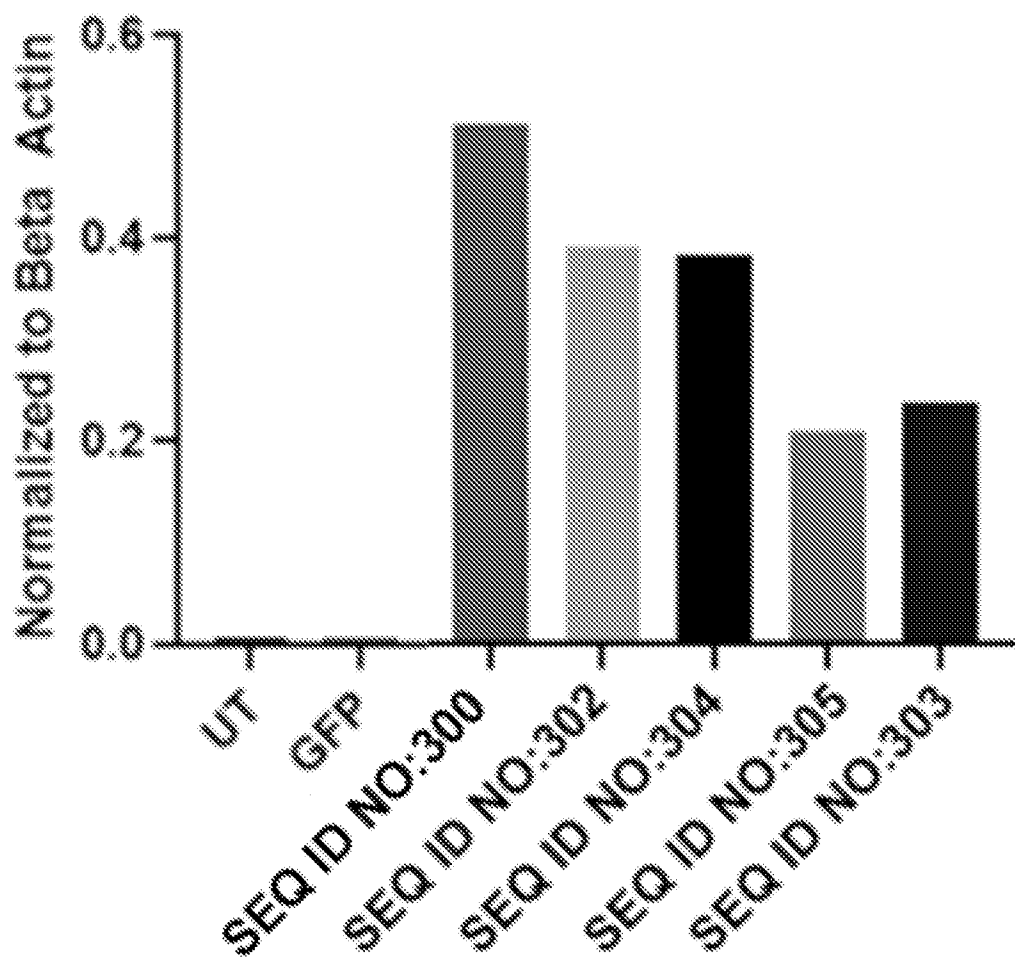


FIG. 2

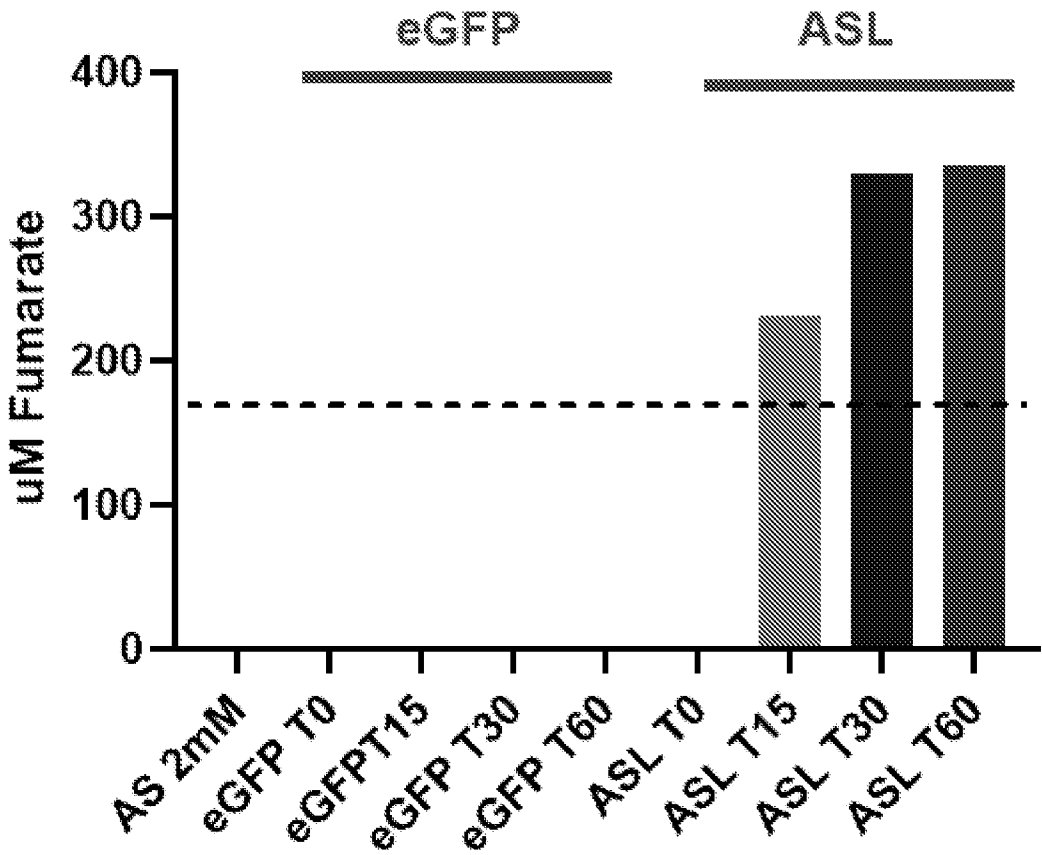


FIG. 3

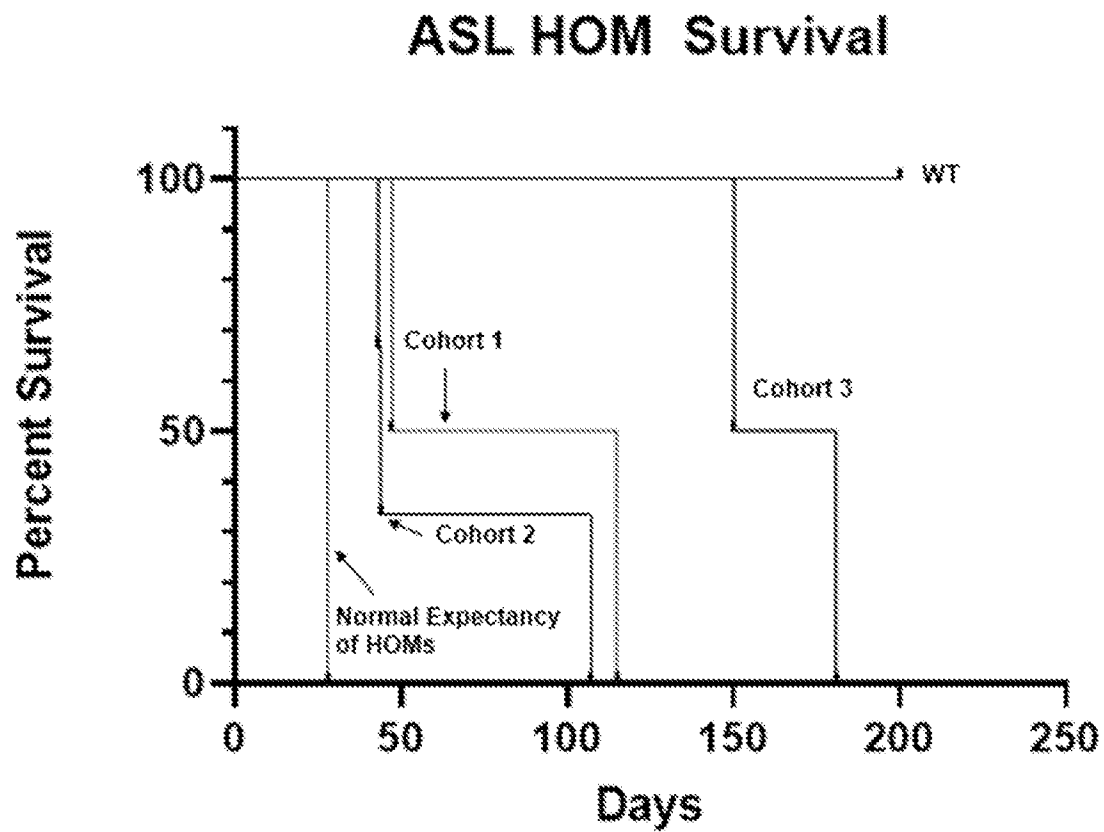


FIG. 4A

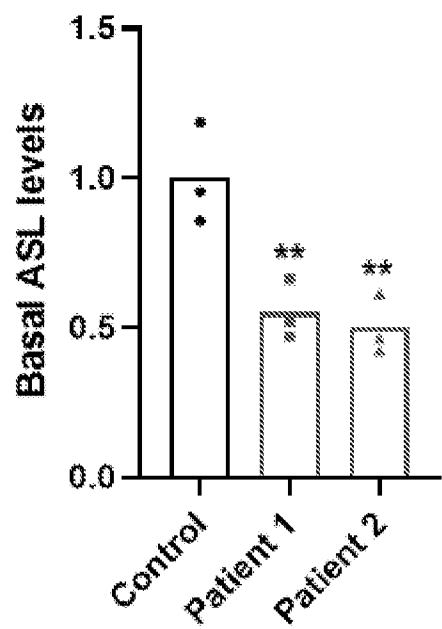


FIG. 4B

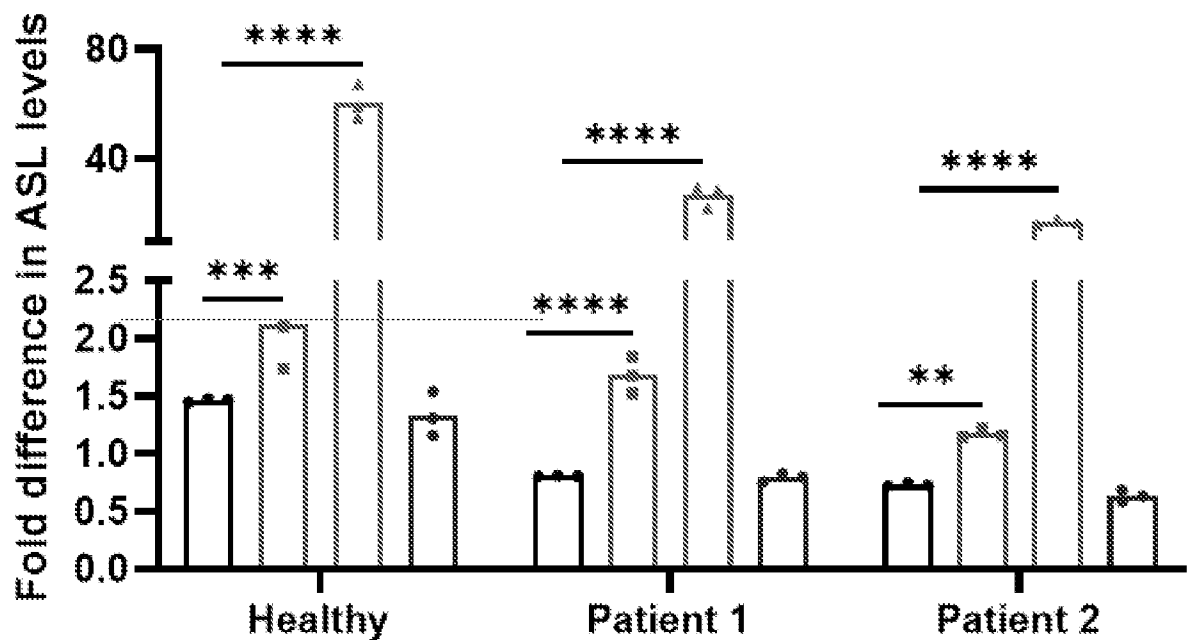


FIG. 4C

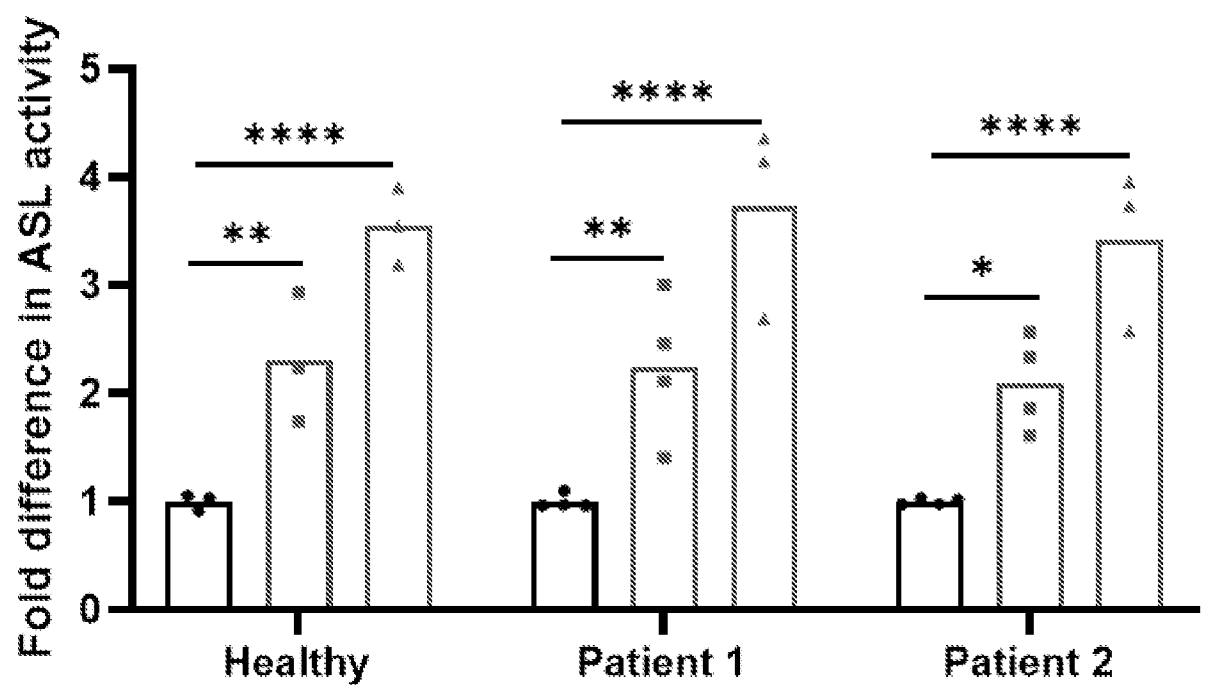


FIG. 5A

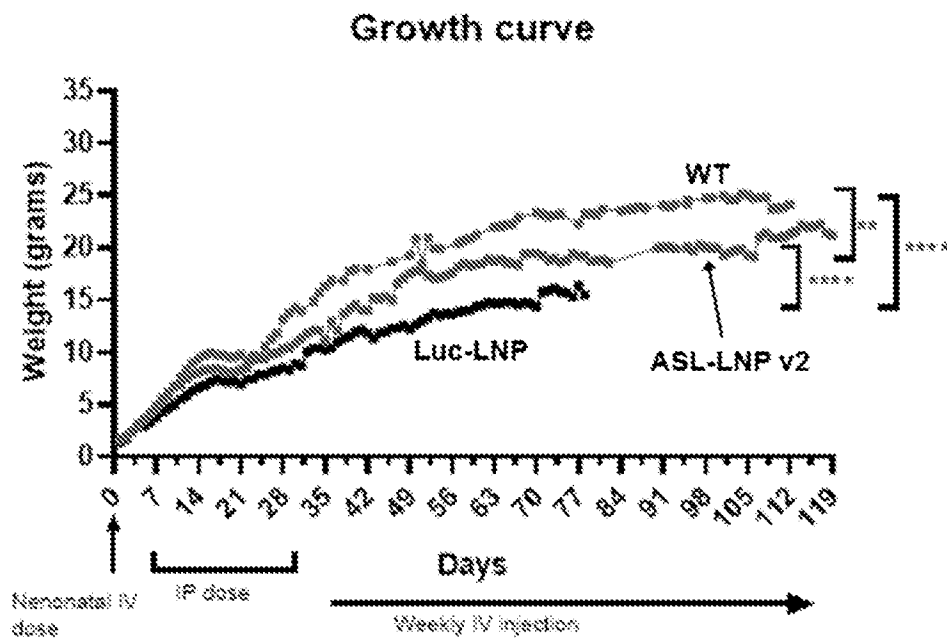


FIG. 5B

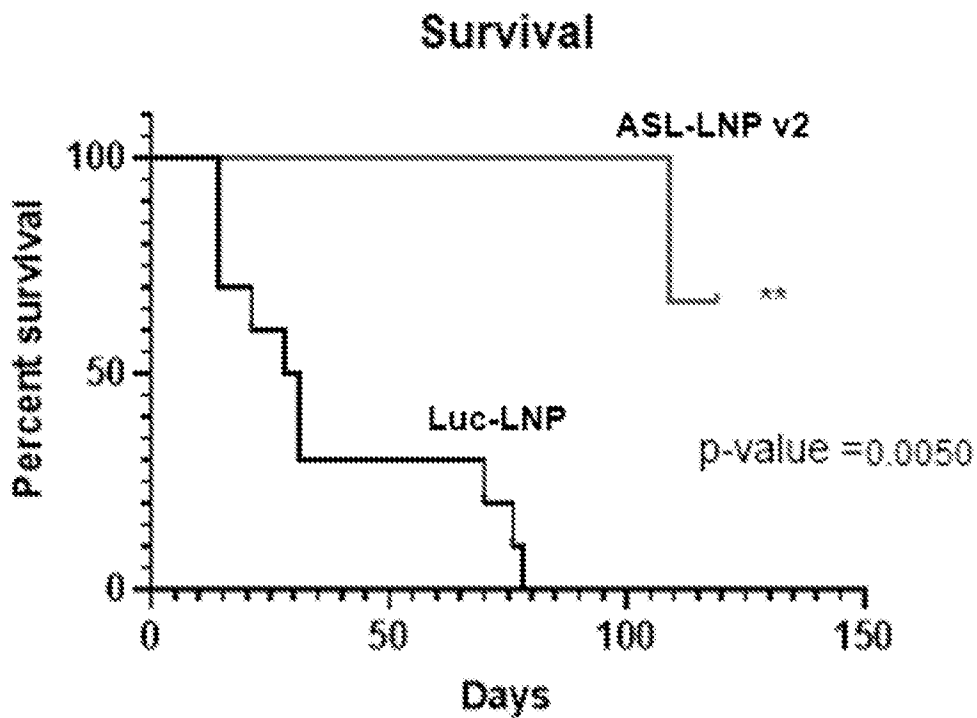


FIG. 6A

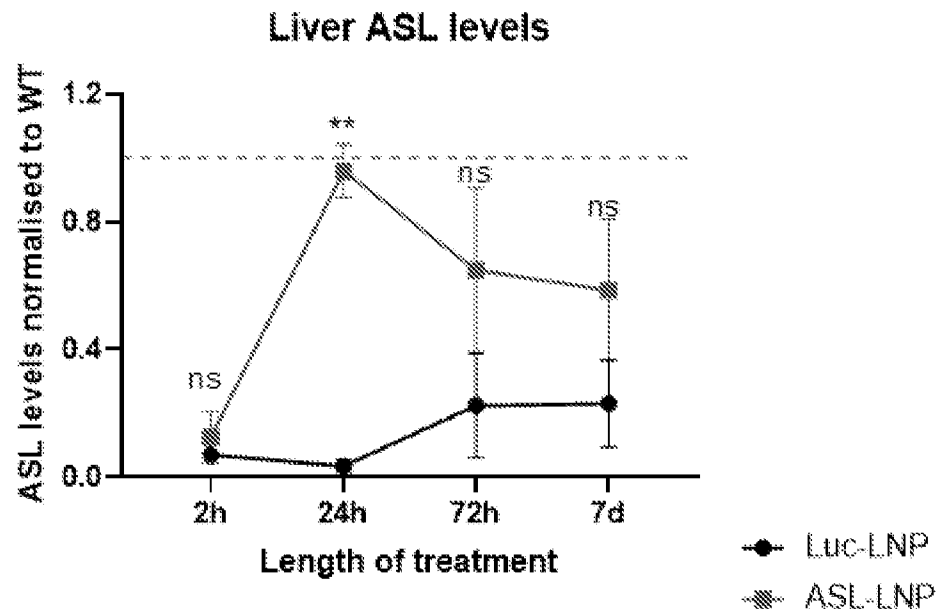


FIG. 6B

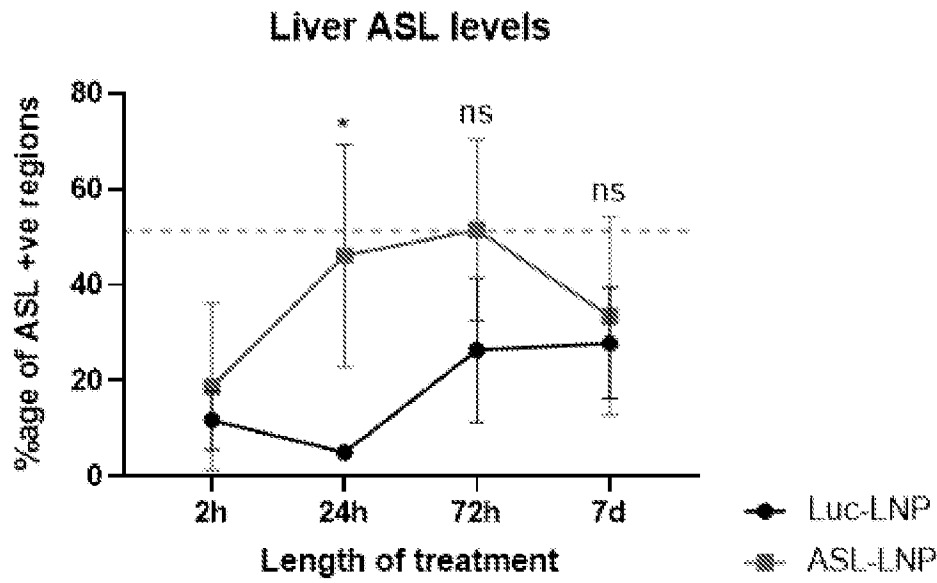


FIG. 6C

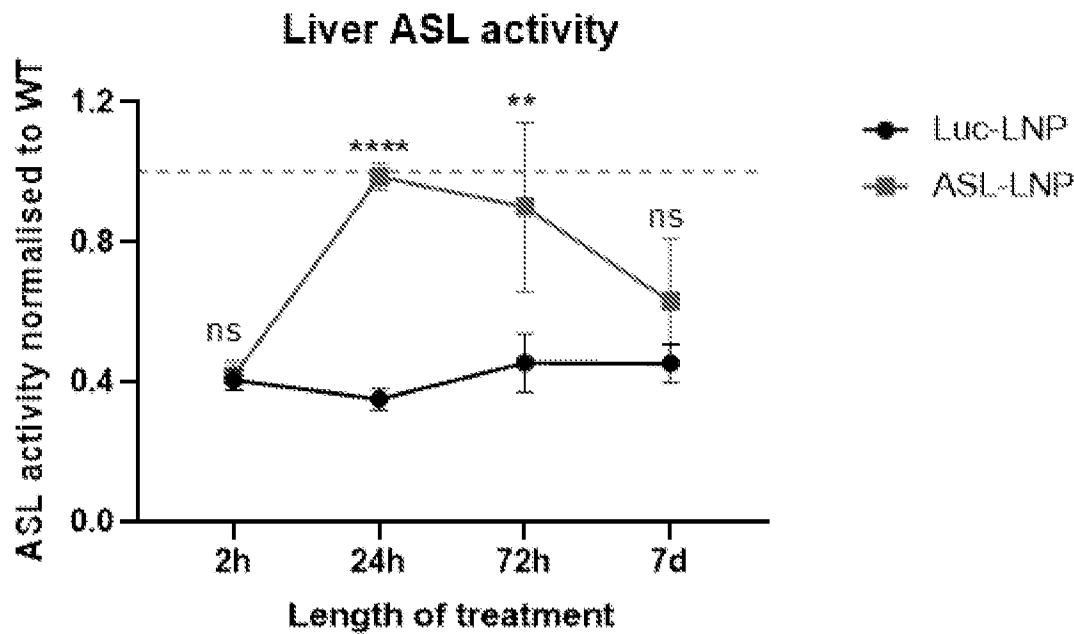


FIG. 6D

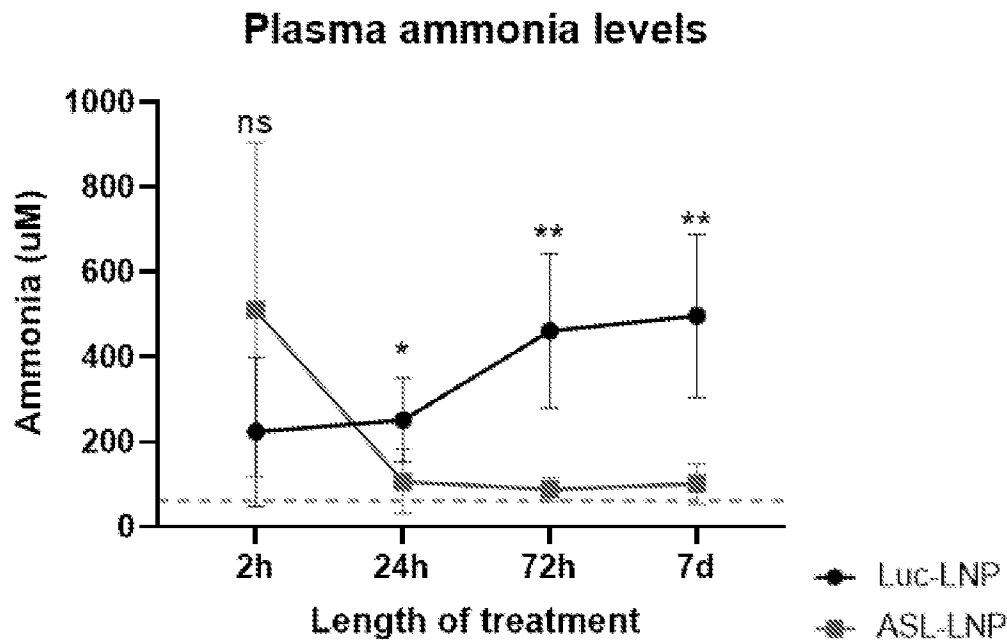


FIG. 6E

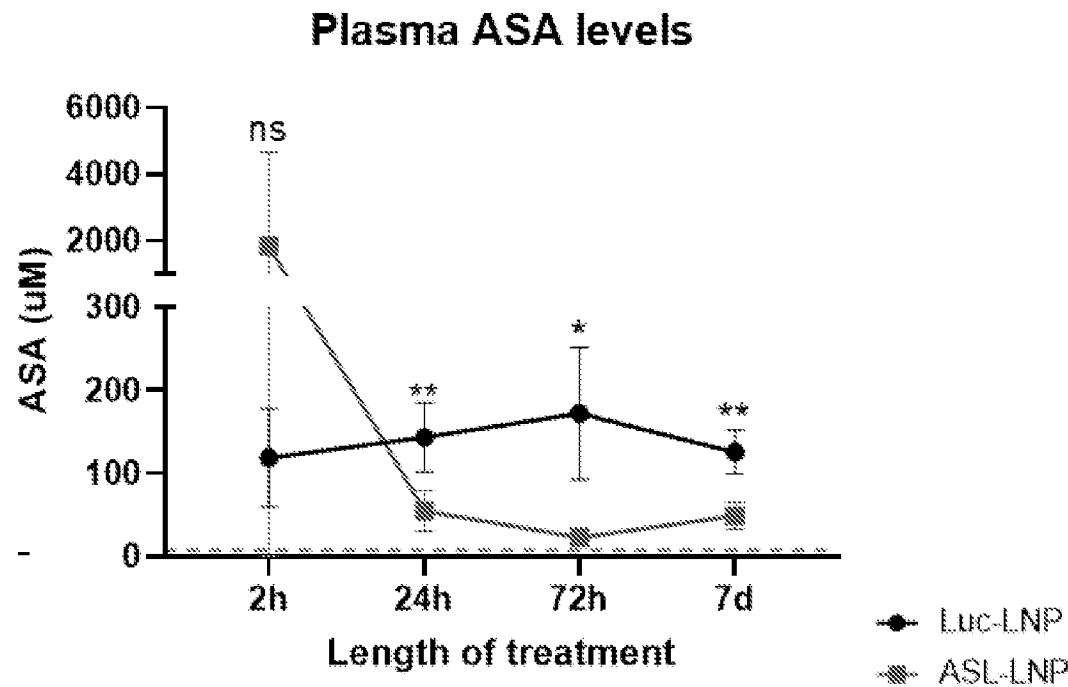


FIG. 6F

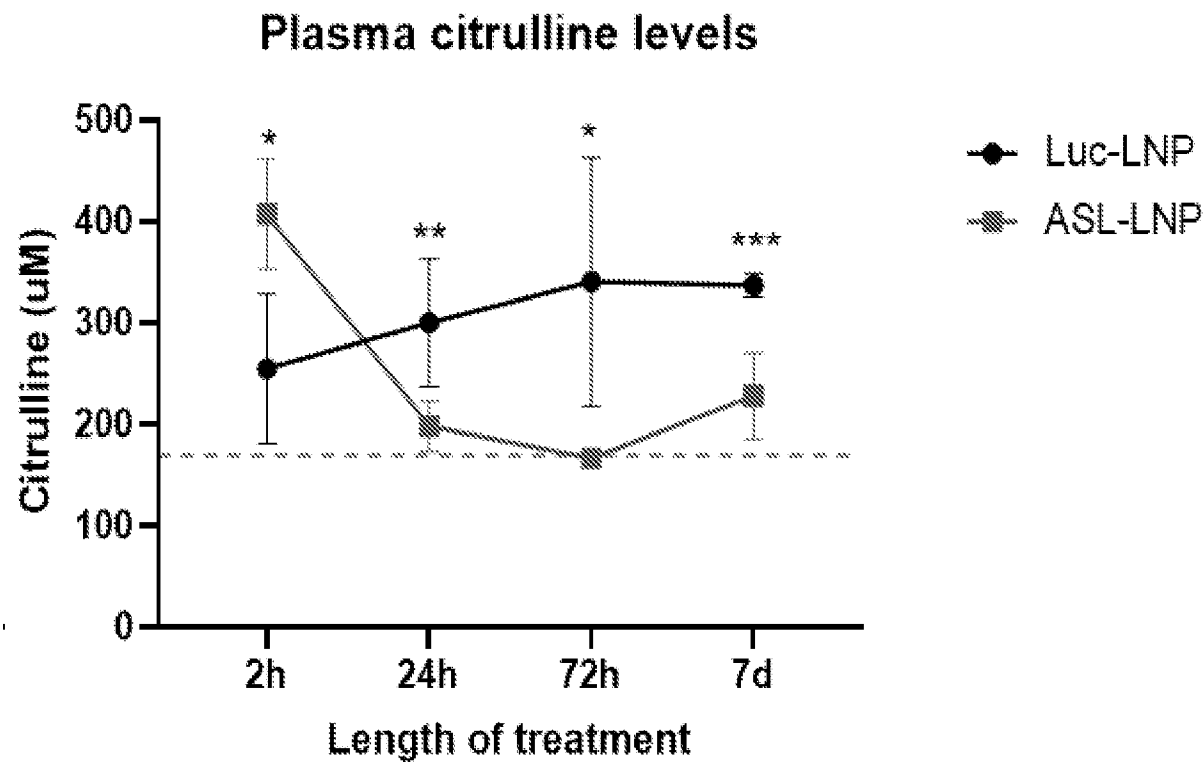


FIG. 6G

orotate levels post 24h administration

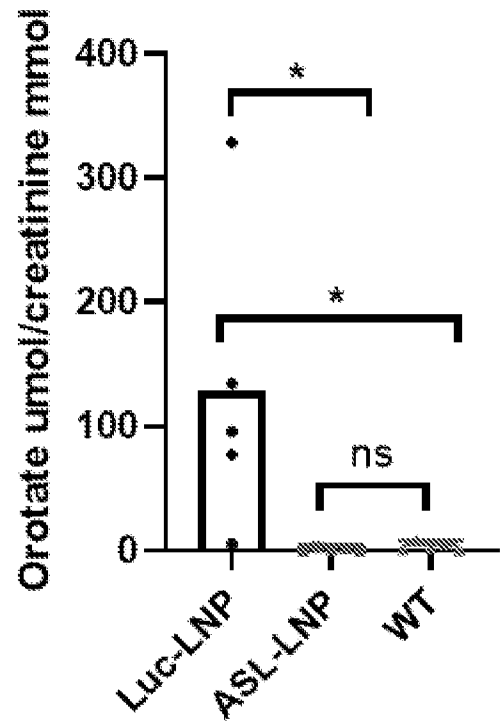


FIG. 7A

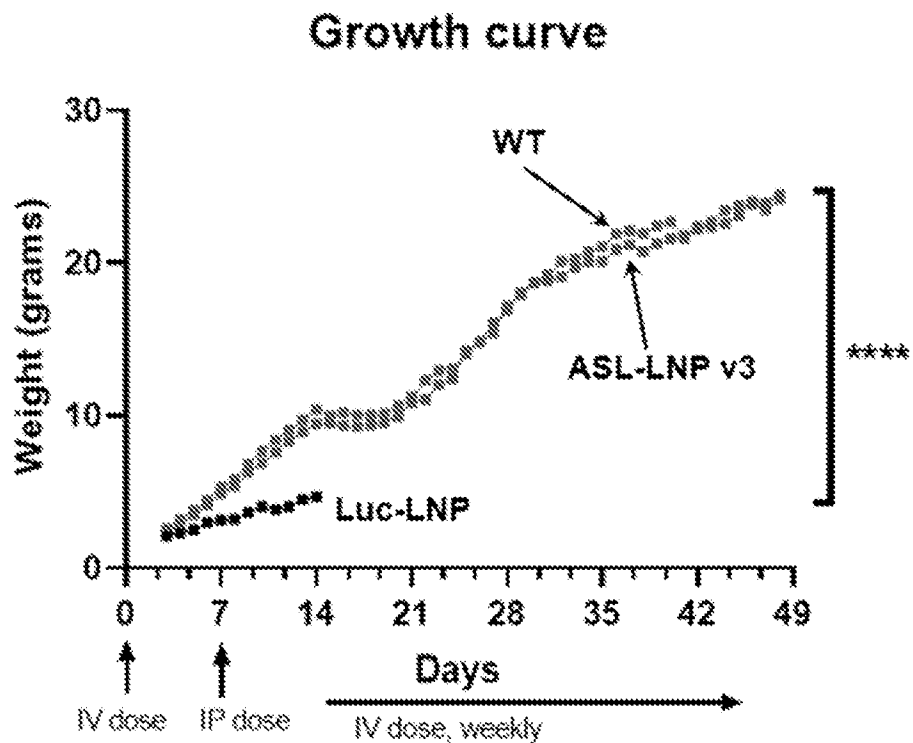


FIG. 7B

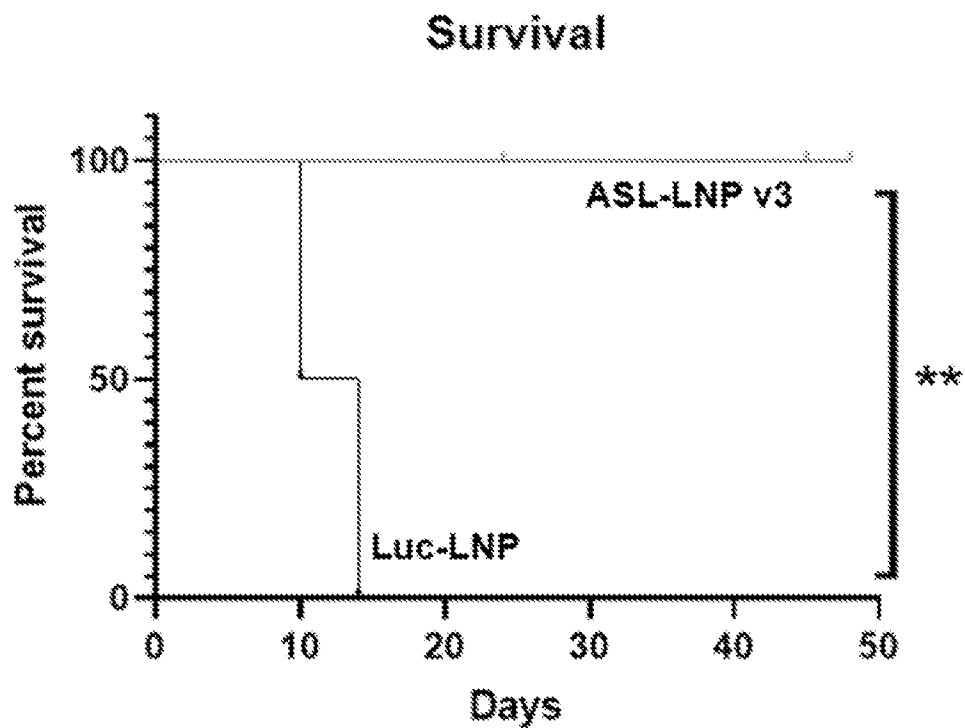


FIG. 7C

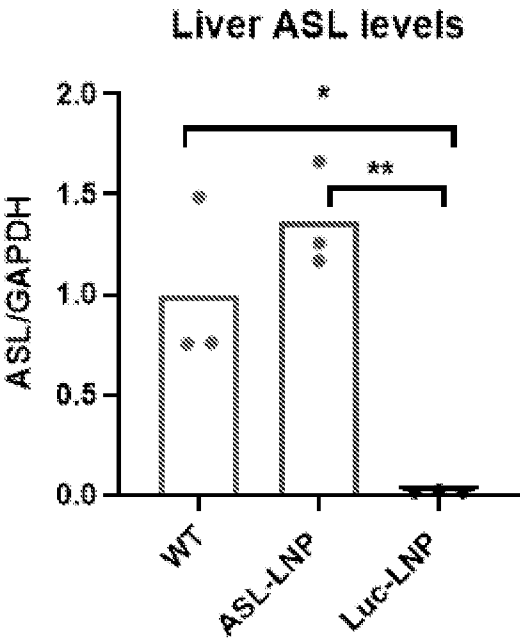


FIG. 7D

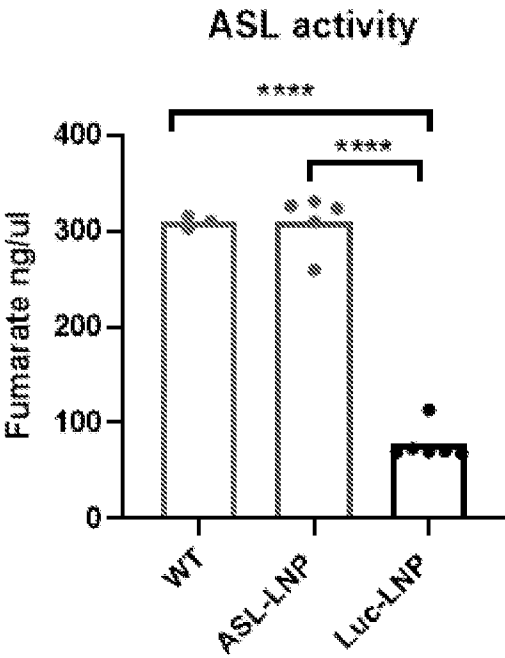


FIG. 7E

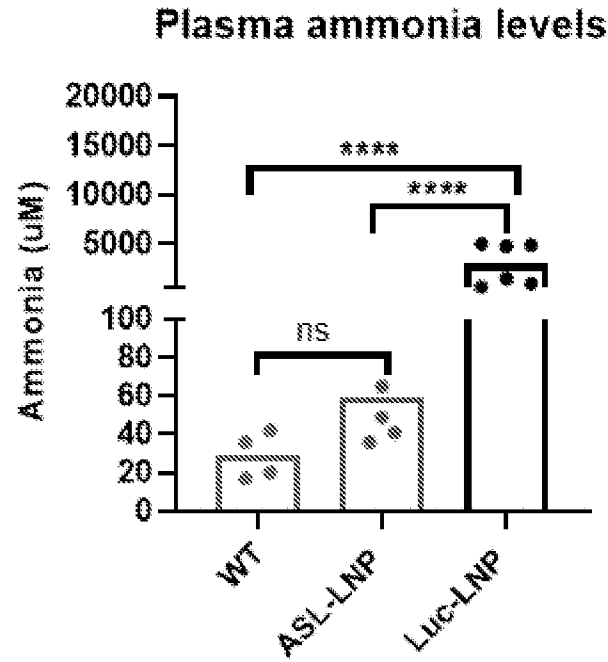


FIG. 7F

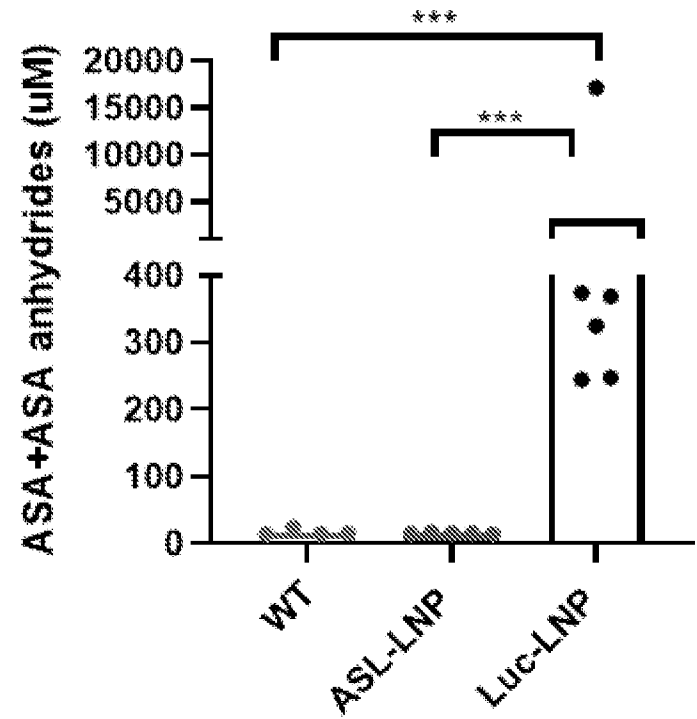


FIG. 7G

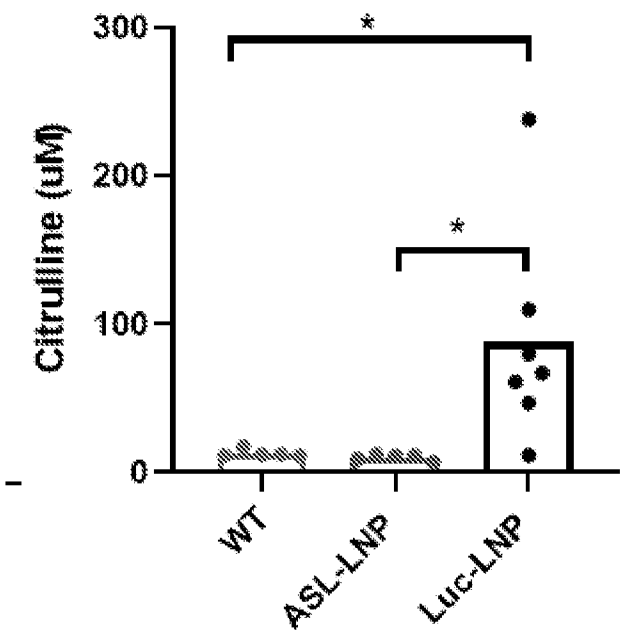


FIG. 7H

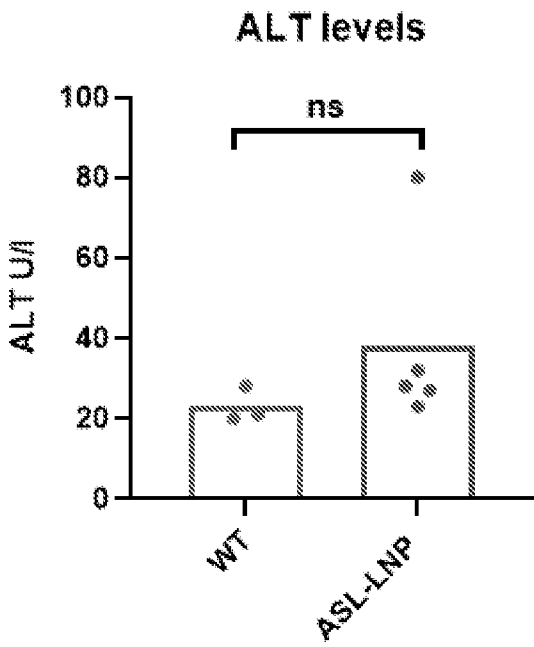


FIG. 7I

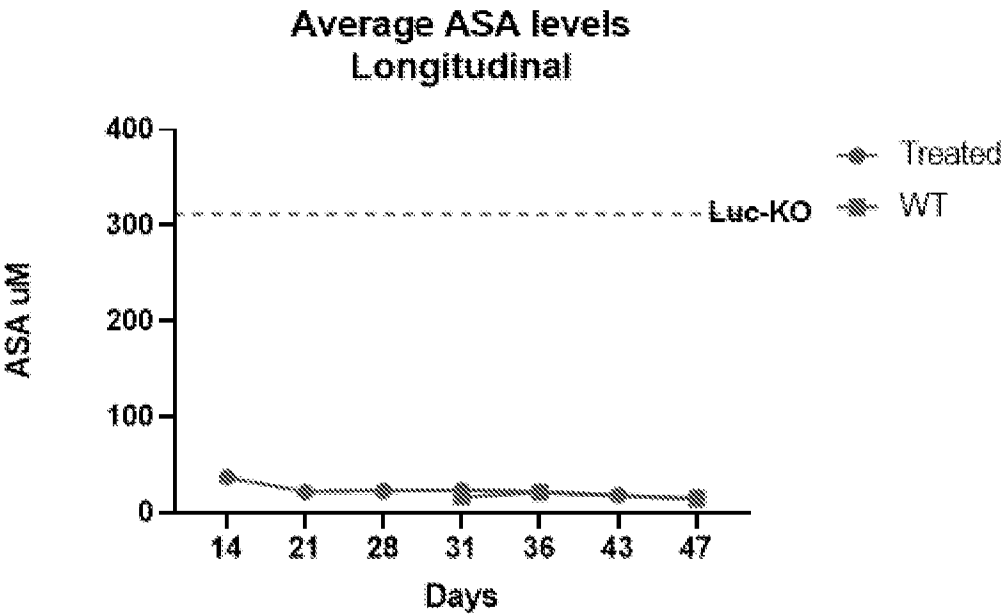


FIG. 7J

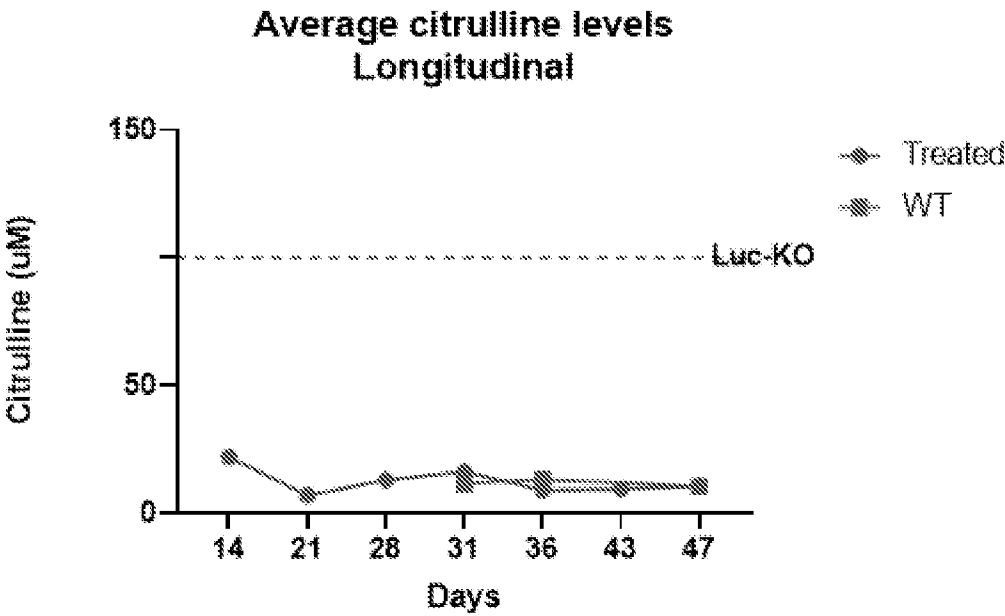


FIG. 7K

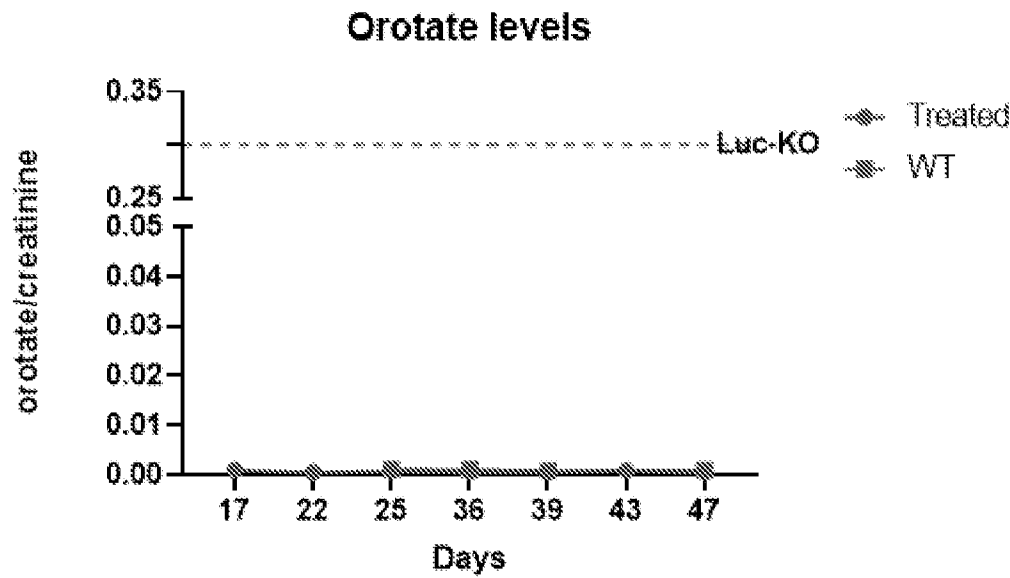


FIG. 7L

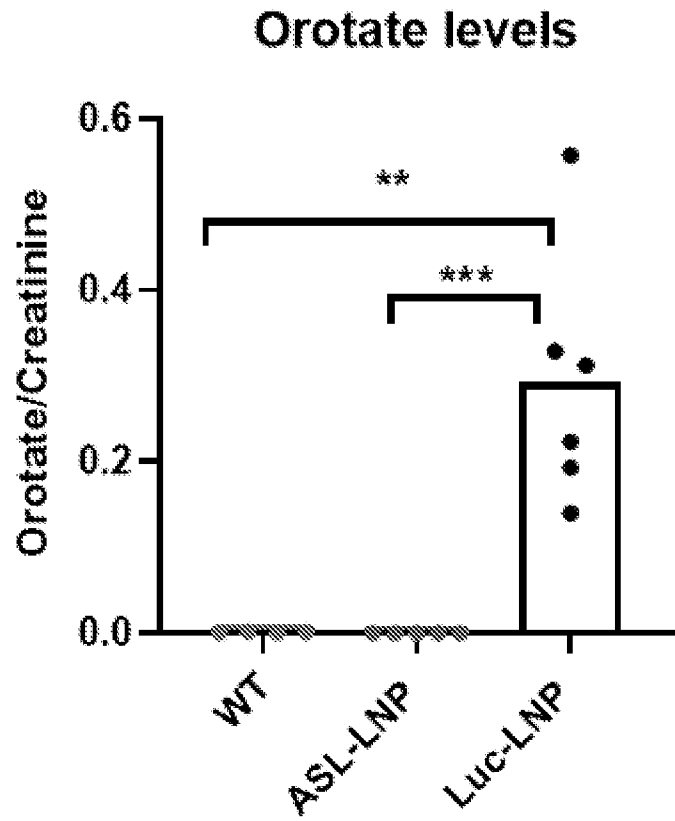


FIG. 7M

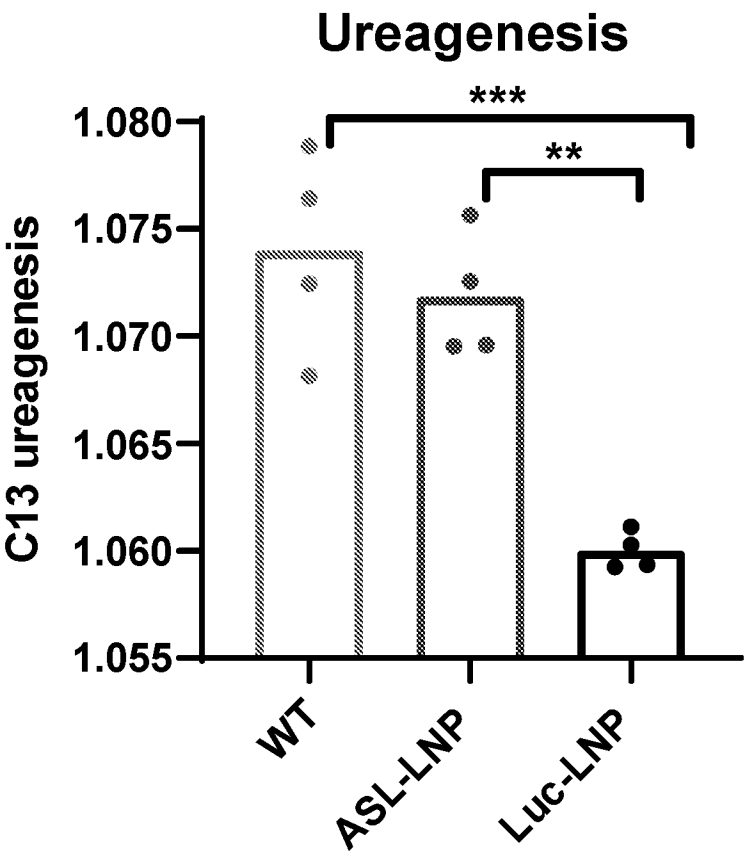


FIG. 8A

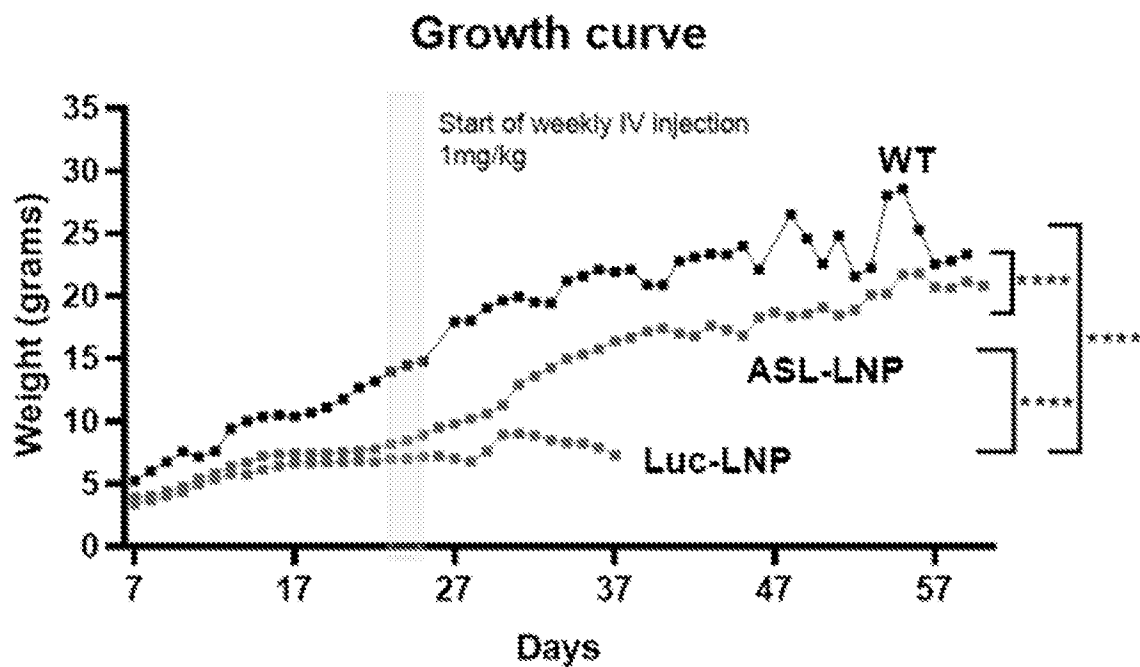


FIG. 8B

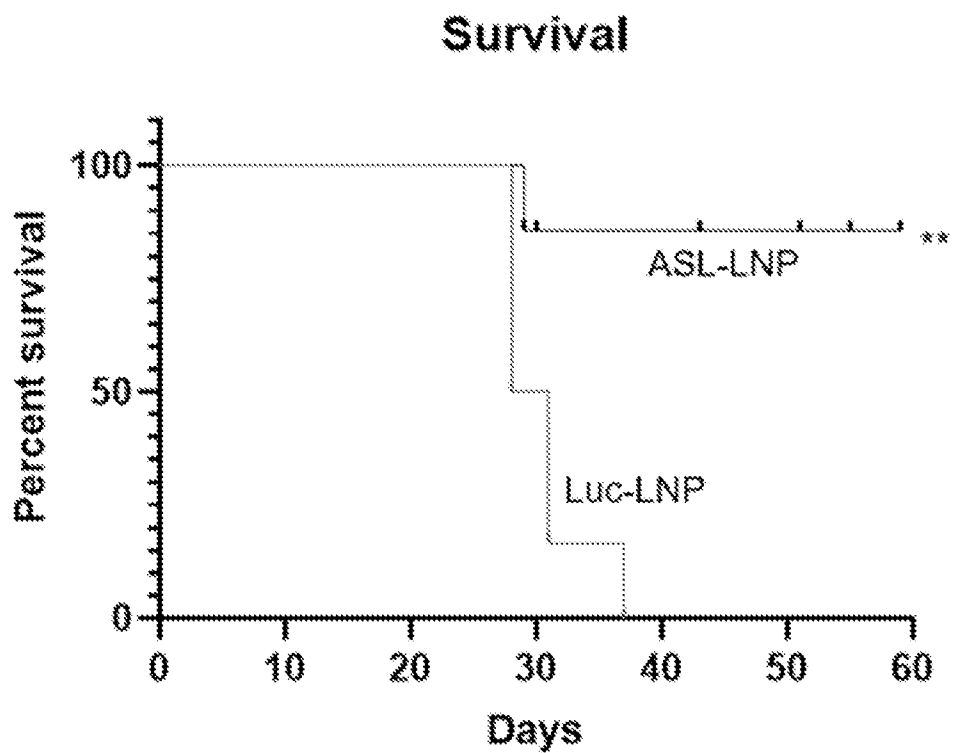
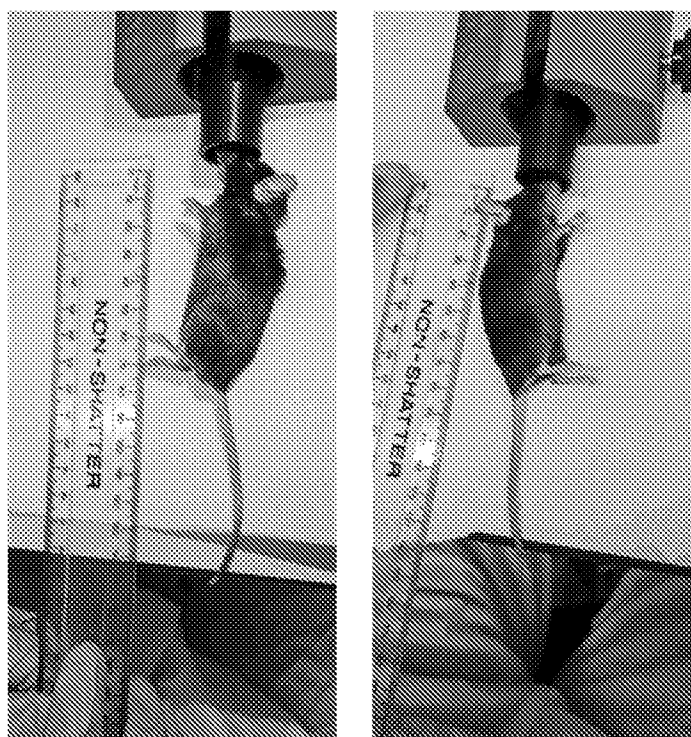
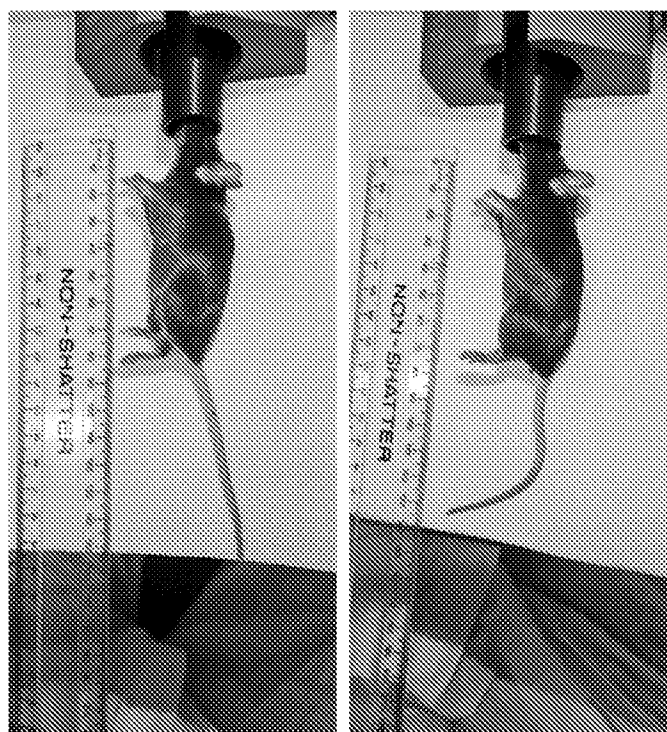


FIG. 9A



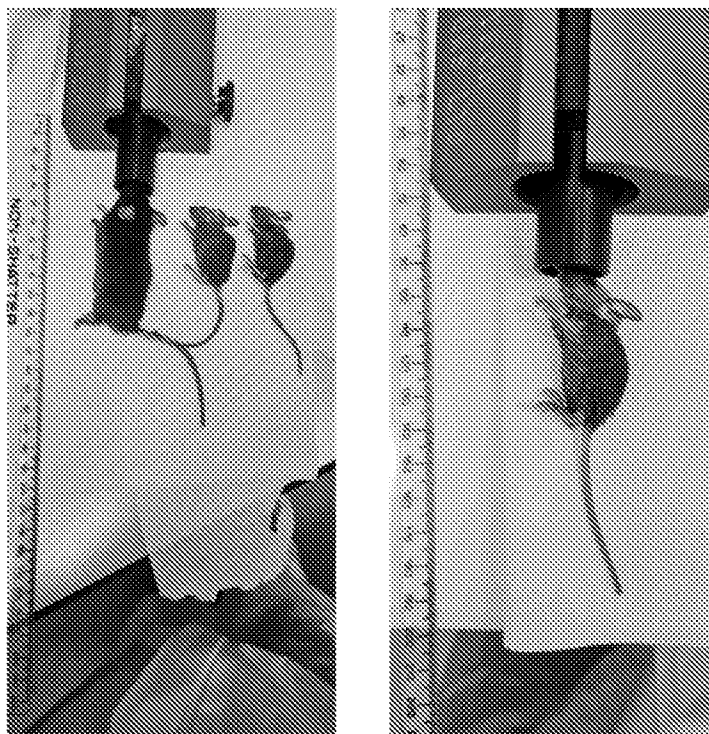
ASL treated d57

FIG. 9B



WTs d57

FIG. 9C



WT (left) and Luc
d27 (two right)

Luc d31

FIG. 10A

Growth velocity - 1st injection

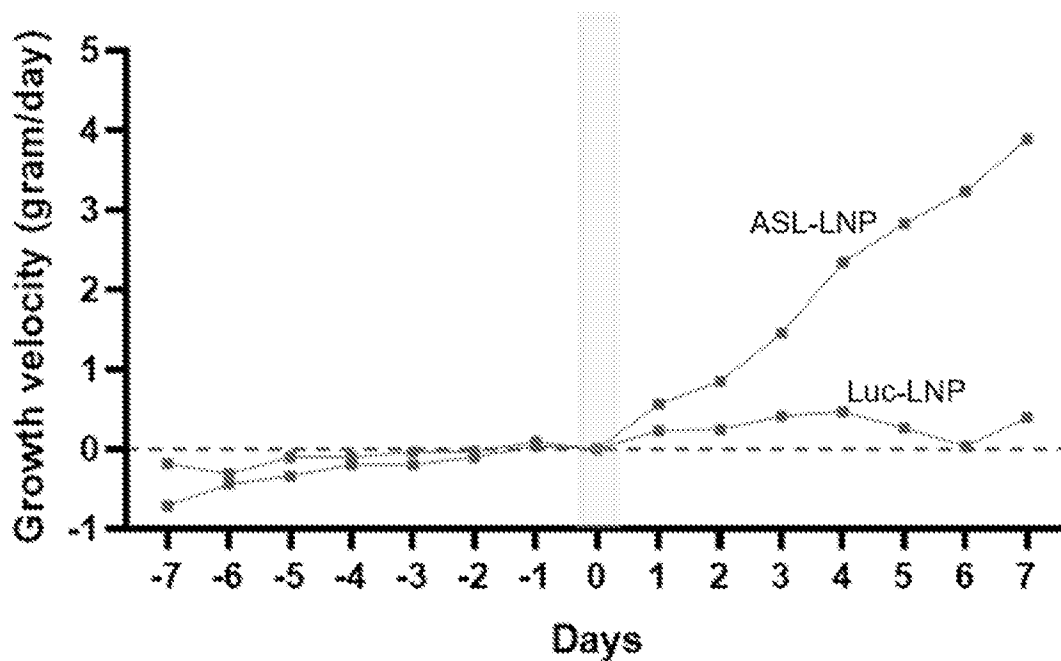


FIG. 10B

Growth velocity - 2nd injection

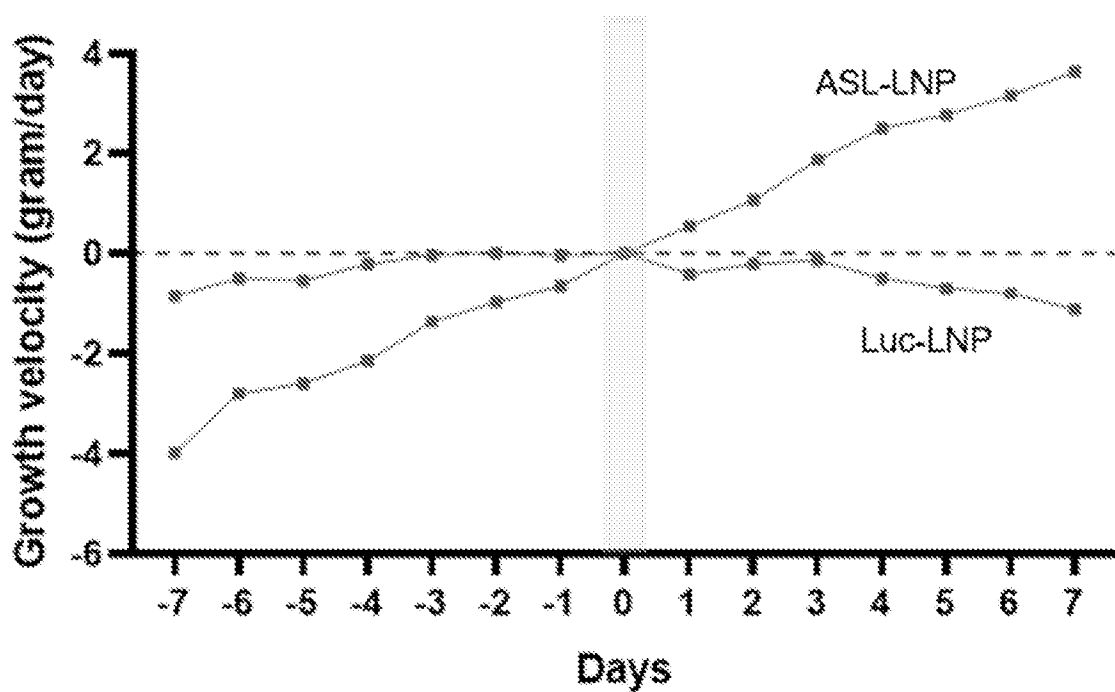


FIG. 10C

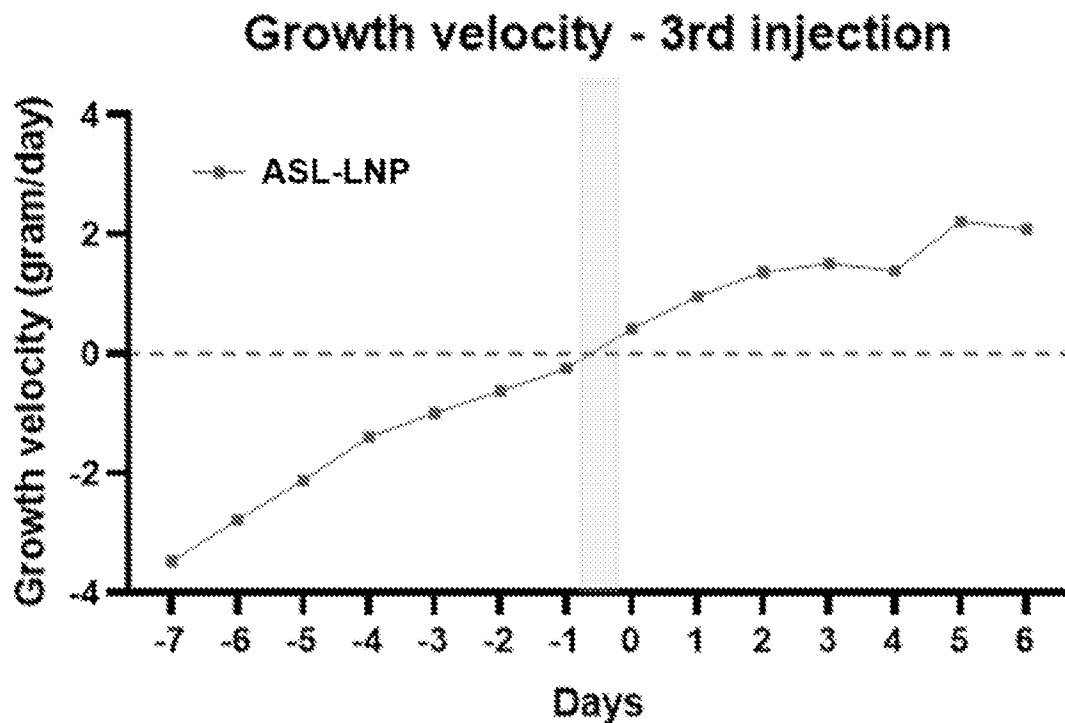


FIG. 10D

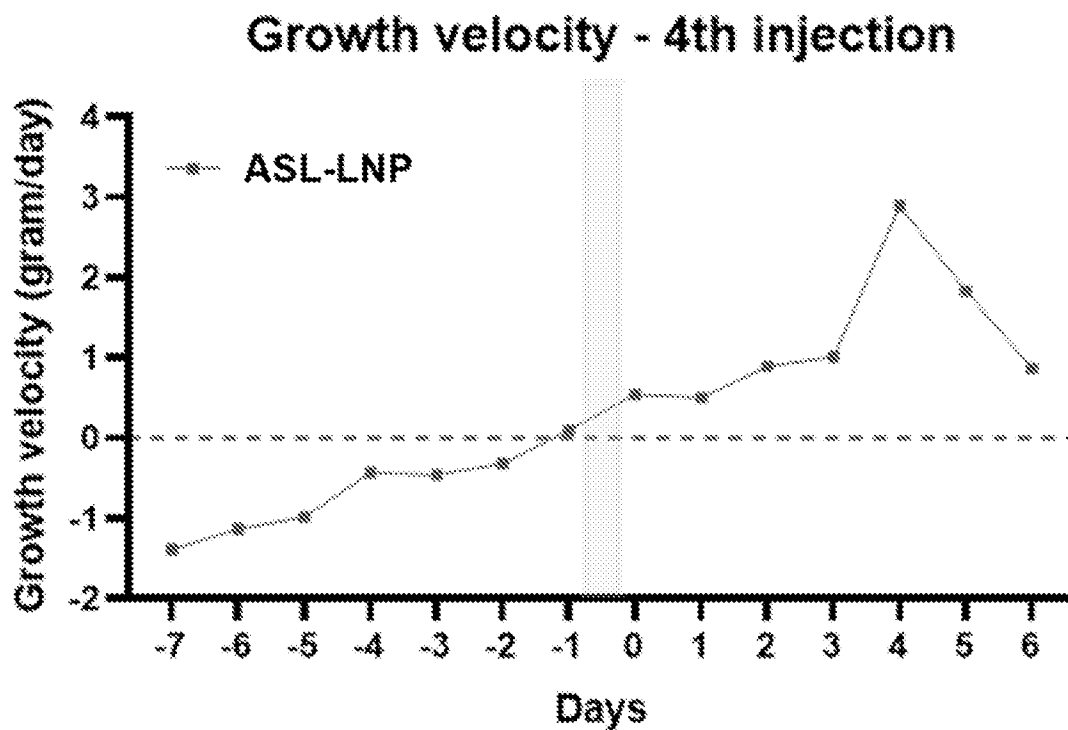


FIG. 11A

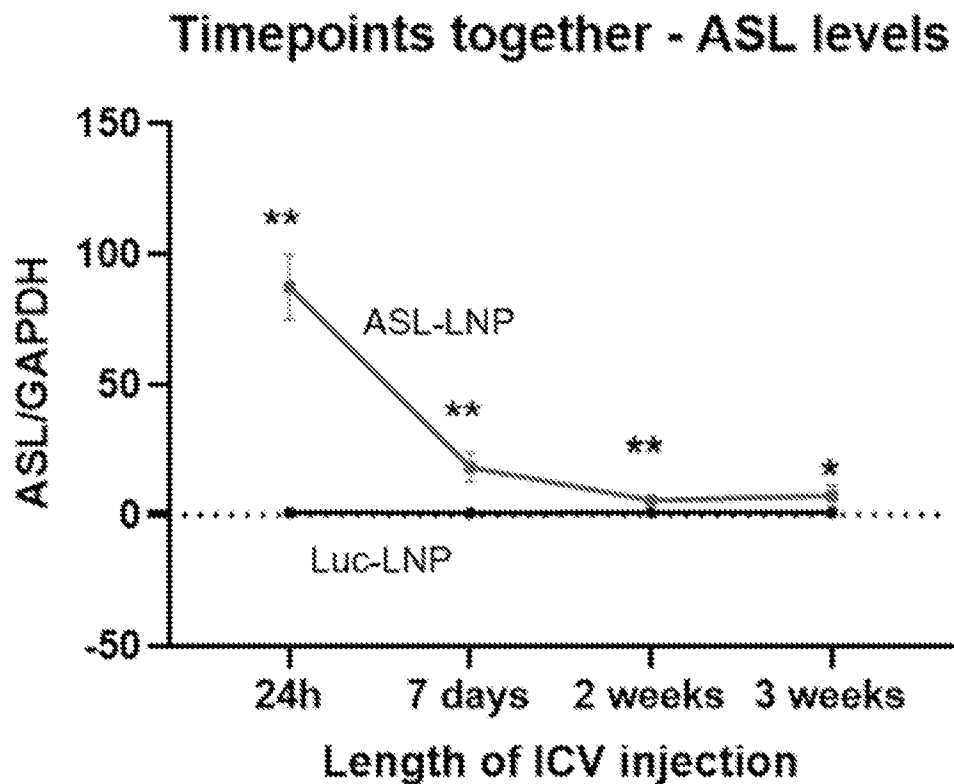


FIG. 11B

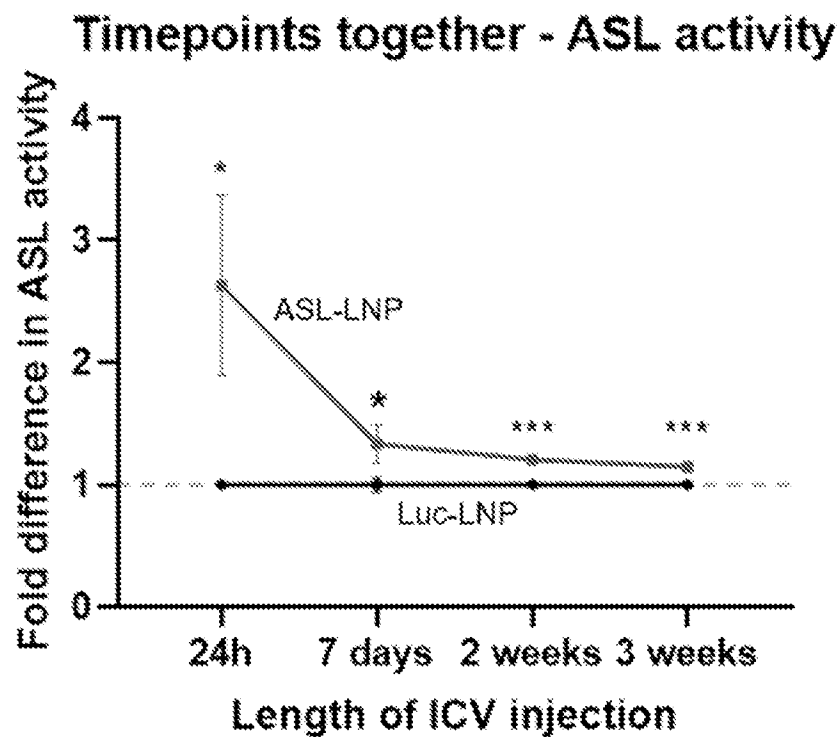


FIG. 12

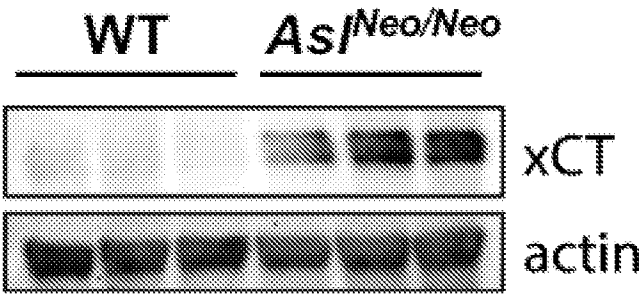


FIG. 13A

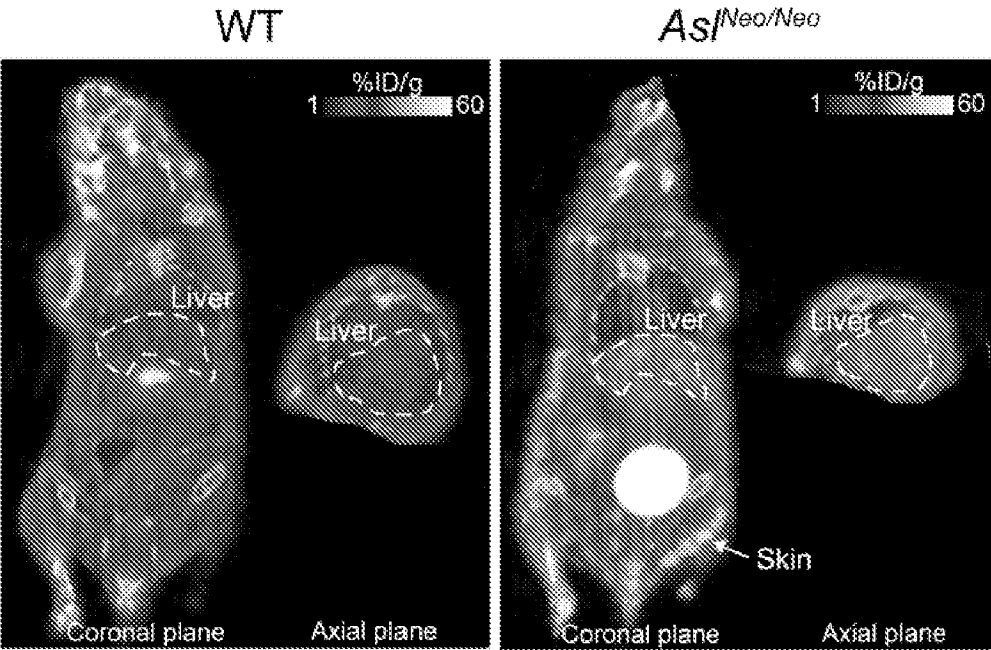


FIG. 13B

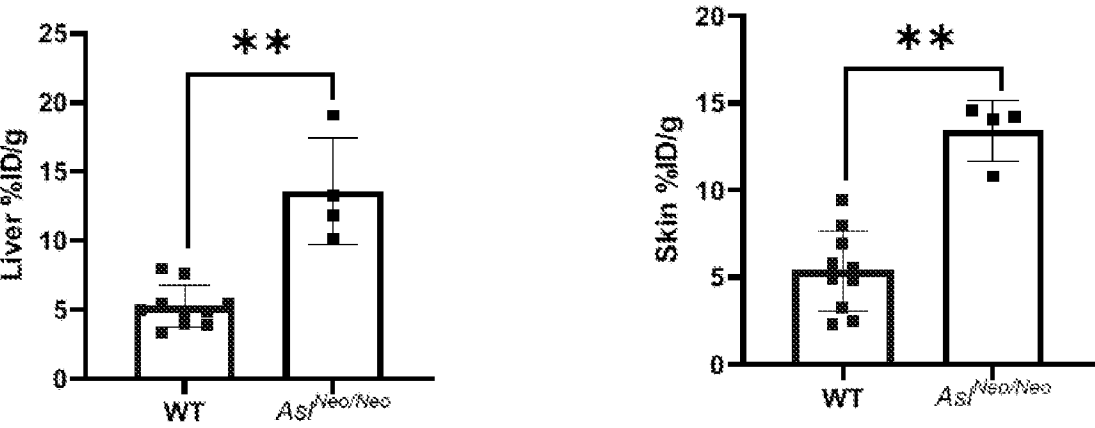


FIG. 14

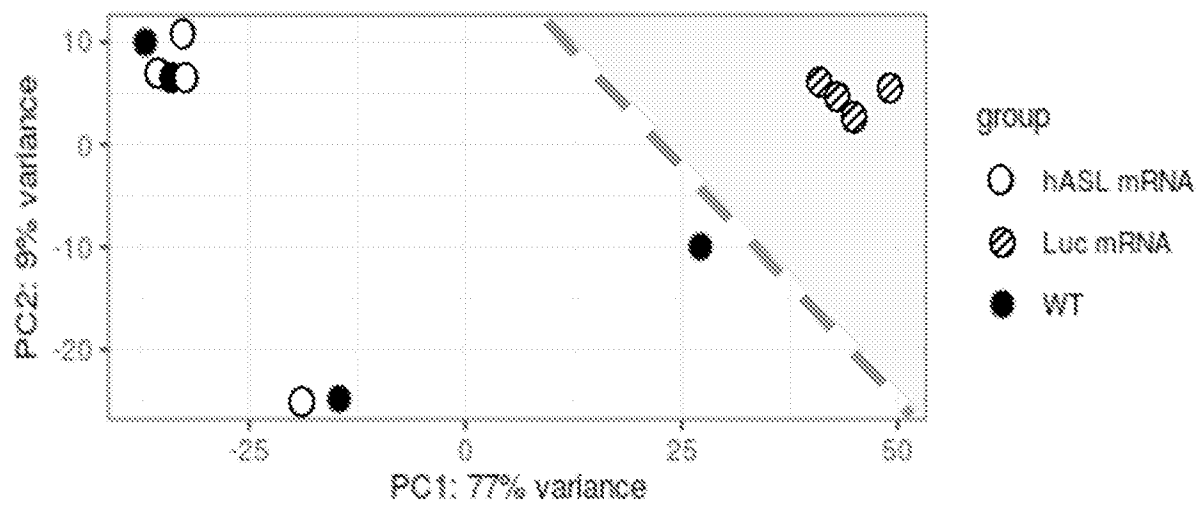


FIG. 15A

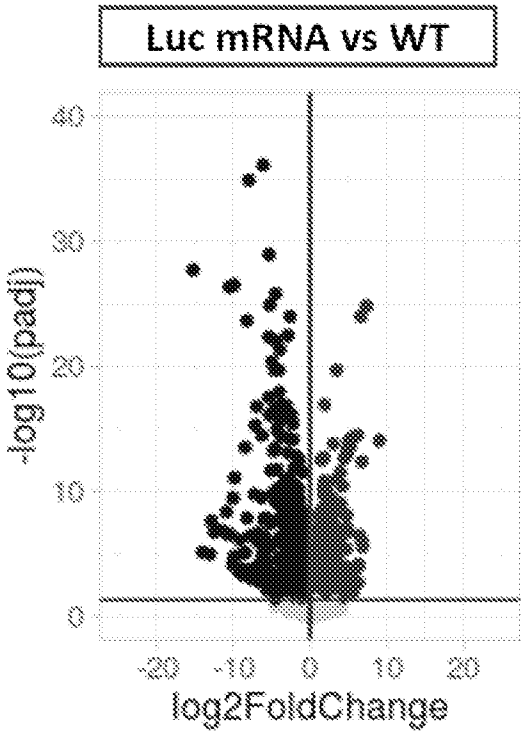


FIG. 15B

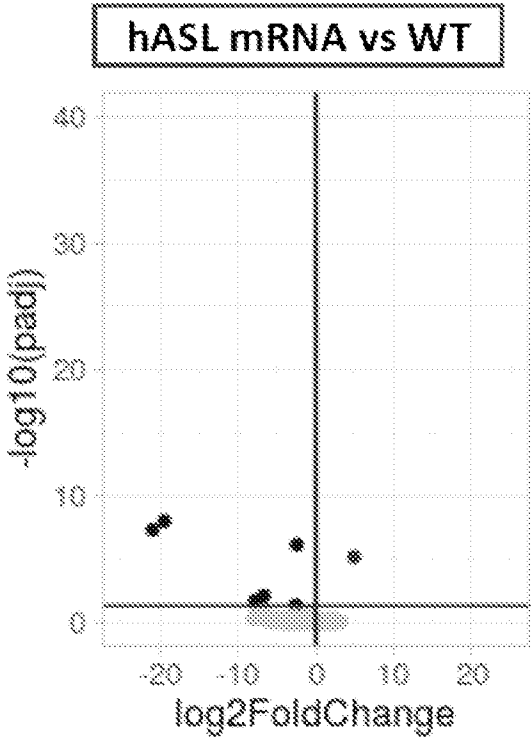


FIG. 15C

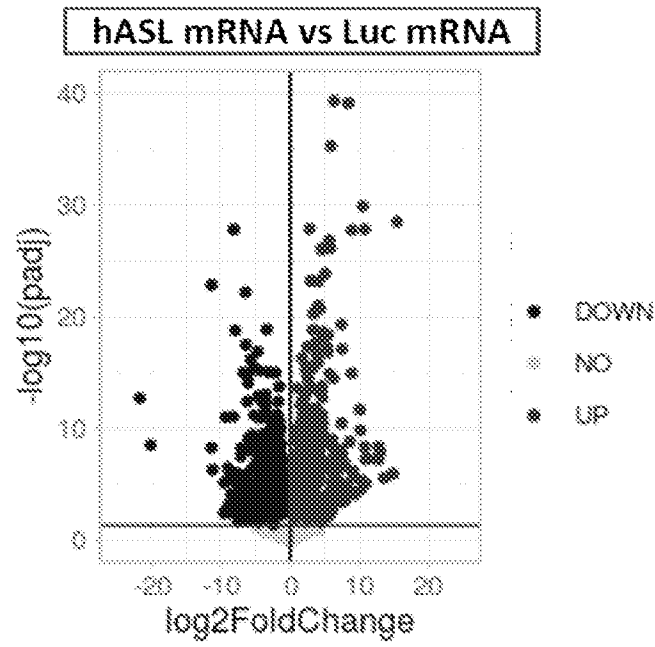


FIG. 16

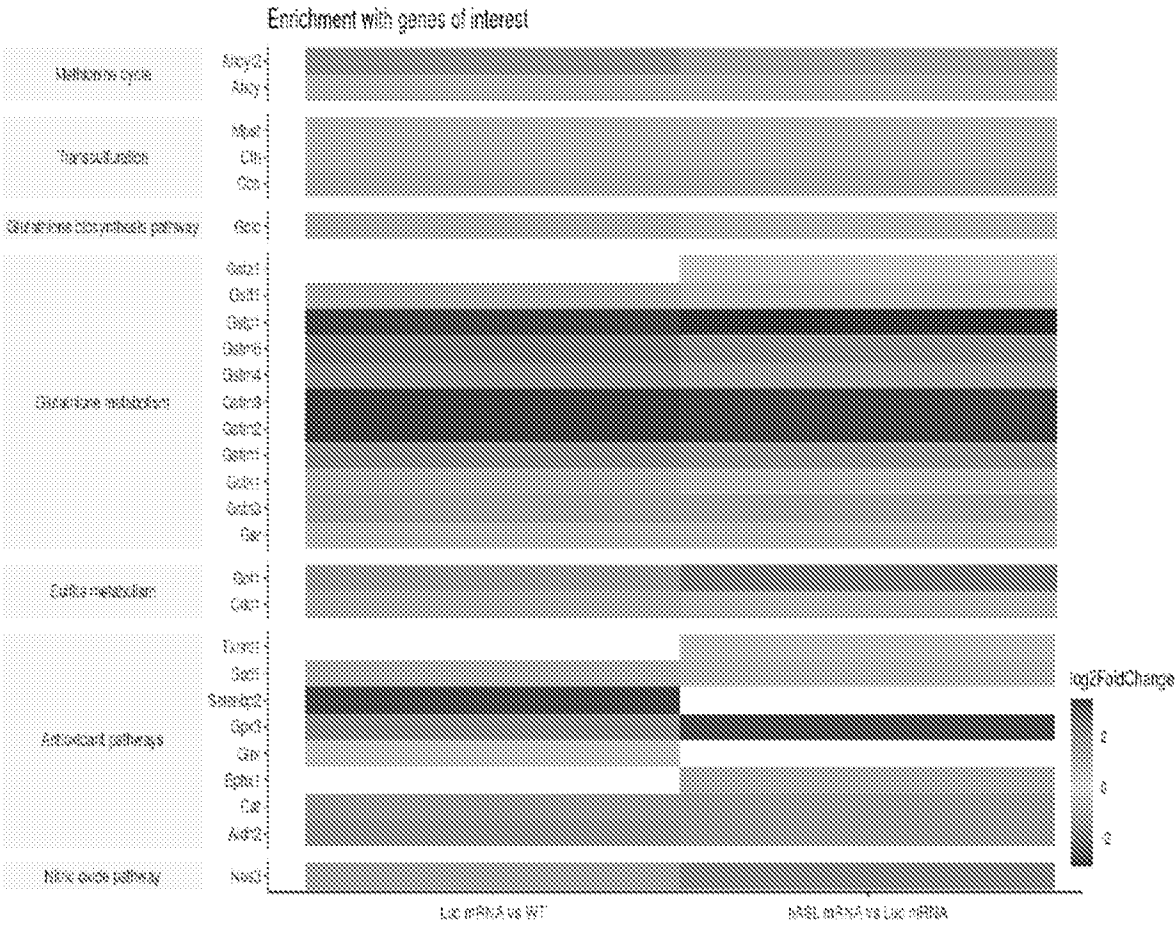


FIG. 17

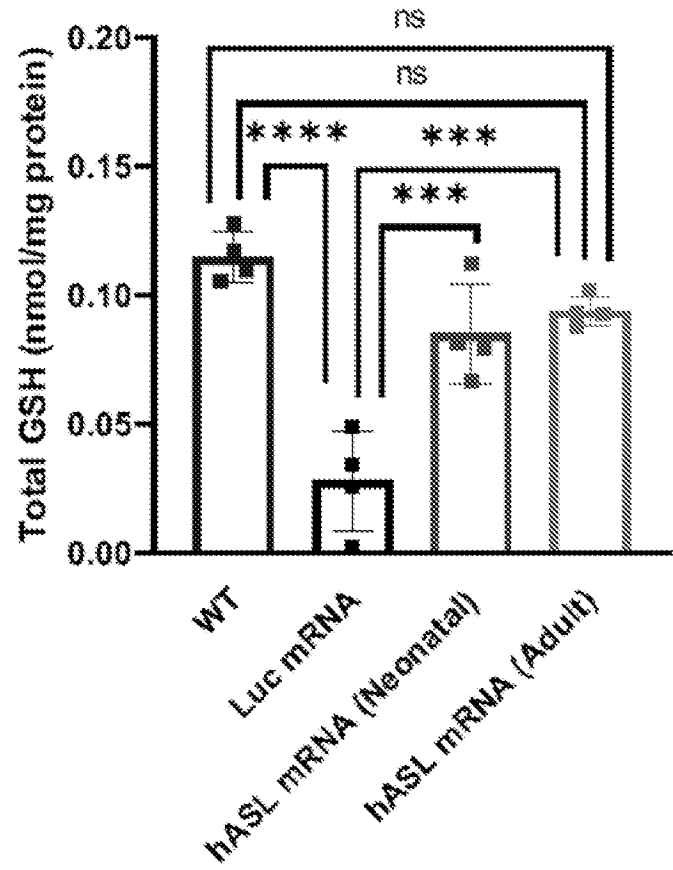


FIG. 18A

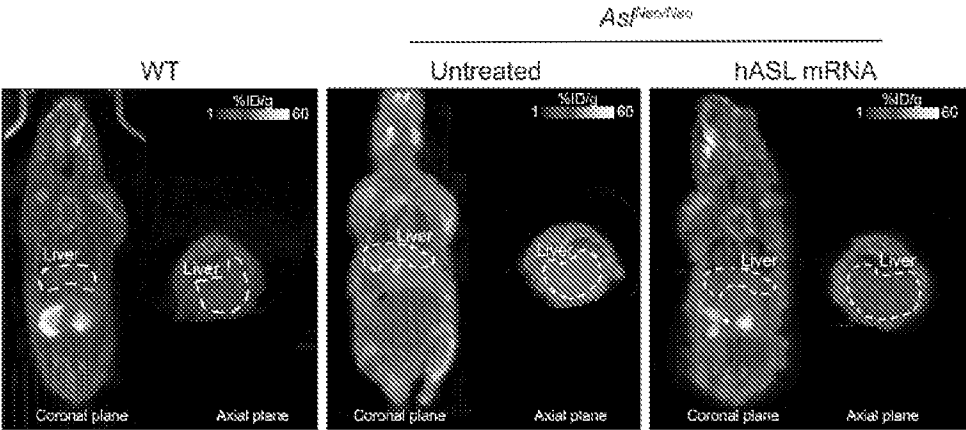


FIG. 18B

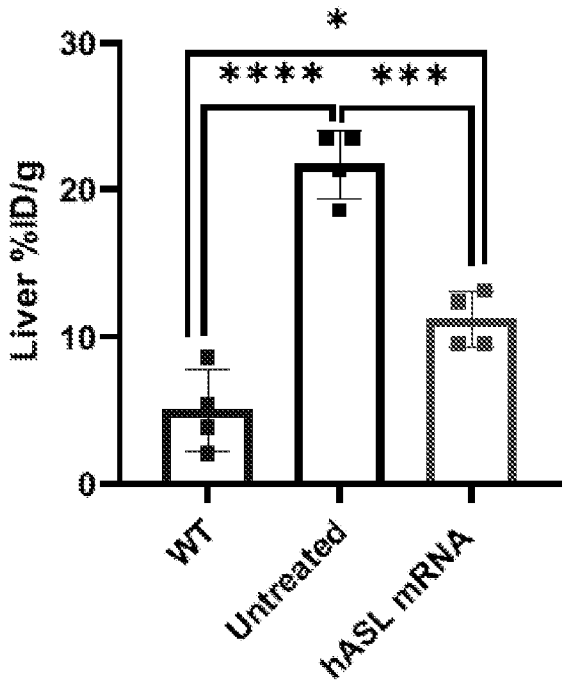


FIG. 19A

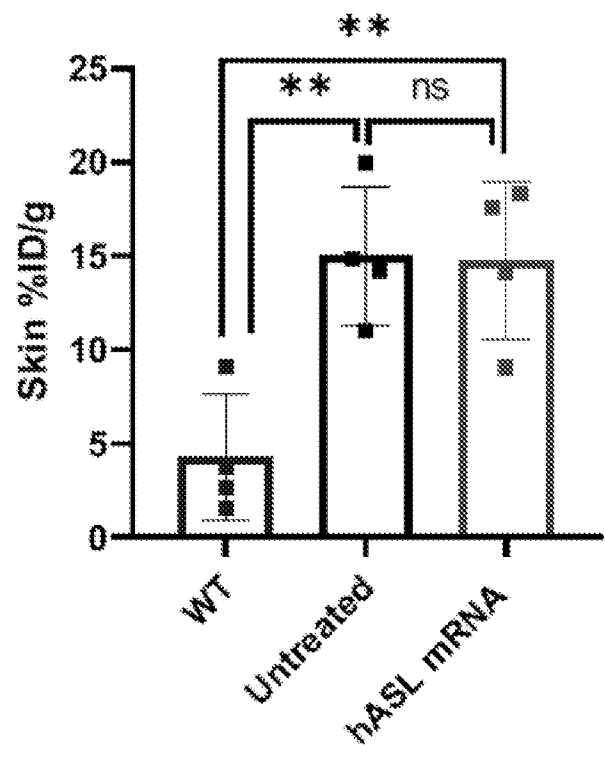


FIG. 19B

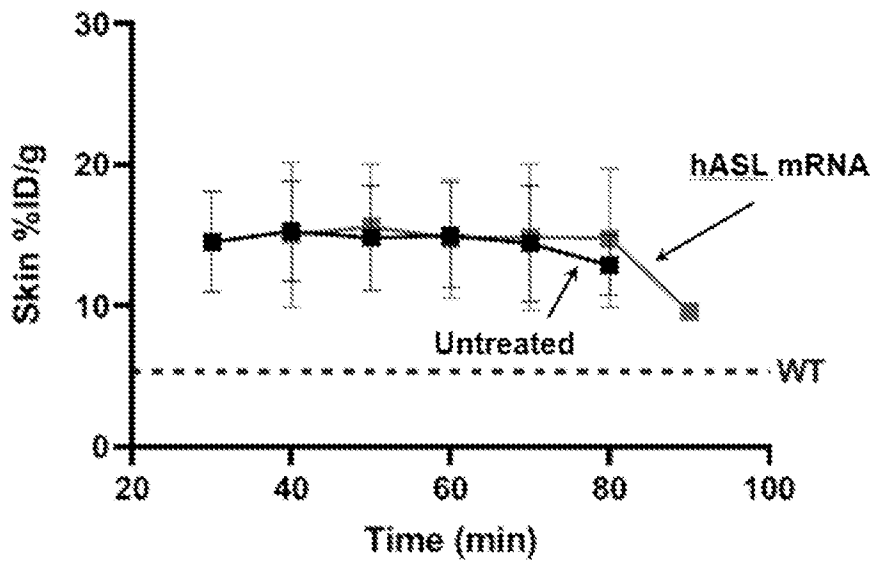


FIG. 20

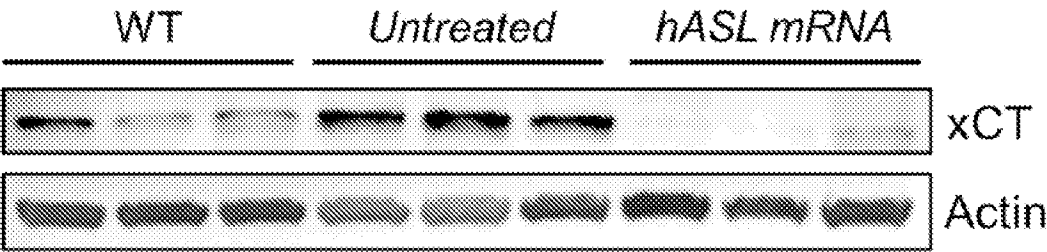
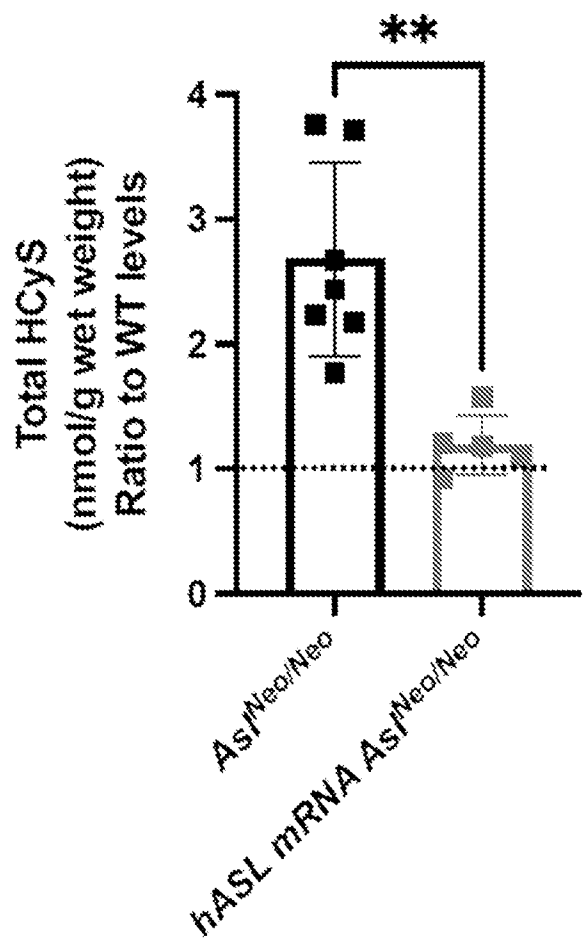


FIG. 21



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/017573

A. CLASSIFICATION OF SUBJECT MATTER		
INV.	A61K9/127	A61K9/51
	A61P3/00	C12N9/88
		C12N15/88
ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
A61K C12N A61P		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
EPO-Internal, BIOSIS, CHEM ABS Data, EMBASE, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2018/126084 A1 (PHASERX INC [US]) 5 July 2018 (2018-07-05) paragraph [0006] - paragraph [0056] paragraph [0165] paragraph [0168] paragraph [0249] paragraph [0261] claims	1-67
X	WO 2016/118697 A1 (PHASERX INC [US]) 28 July 2016 (2016-07-28)	50, 51, 53-58
Y	the whole document paragraph [0007] - paragraph [0052] example 24 claims	1-20, 22-49, 59-67
	----- -/--	
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
17 July 2023		27/07/2023
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Marchand, Petra

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2023/017573

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2021/262909 A2 (MODERNATX INC [US]) 30 December 2021 (2021-12-30) the whole document summary of invention; page 1, line 16 – page 103, line 10 page 224, line 27 – page 264, line 12 examples claims -----	1-20, 22-49, 59-67
Y	STACI SABNIS ET AL: "A Novel Amino Lipid Series for mRNA Delivery: Improved Endosomal Escape and Sustained Pharmacology and Safety in Non-human Primates", MOLECULAR THERAPY, vol. 26, no. 6, 1 June 2018 (2018-06-01), pages 1509-1519, XP055644778, US ISSN: 1525-0016, DOI: 10.1016/j.ymthe.2018.03.010 the whole document abstract figures 1-5 -----	1-5, 7, 9-20, 22-49, 59-67
Y	CORNEBISE MARK ET AL: "Discovery of a Novel Amino Lipid That Improves Lipid Nanoparticle Performance through Specific Interactions with mRNA", ADVANCED FUNCTIONAL MATERIALS, vol. 32, no. 8, 12 February 2022 (2022-02-12), page 2106727, XP093028016, DE ISSN: 1616-301X, DOI: 10.1002/adfm.202106727 Retrieved from the Internet: URL:https://onlinelibrary.wiley.com/doi/full-xml/10.1002/adfm.202106727> the whole document abstract figures "Results" "Experimental Section" -----	1-7, 9-20, 22-49, 59-67

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/017573

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed.
 - b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13~~ter~~.1(a)).
☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2023/017573

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018126084 A1	05-07-2018	EP 3562510 A1	06-11-2019
		US 2020078313 A1	12-03-2020
		WO 2018126084 A1	05-07-2018
WO 2016118697 A1	28-07-2016	AU 2016209295 A1	31-08-2017
		CA 2974503 A1	28-07-2016
		CN 107530436 A	02-01-2018
		CN 114642735 A	21-06-2022
		EP 3247328 A1	29-11-2017
		HK 1248548 A1	19-10-2018
		JP 7189921 B2	14-12-2022
		JP 2018509387 A	05-04-2018
		JP 2021050212 A	01-04-2021
		US 2018221402 A1	09-08-2018
		US 2023038687 A1	09-02-2023
		WO 2016118697 A1	28-07-2016
WO 2021262909 A2	30-12-2021	AU 2021297248 A1	02-02-2023
		BR 112022025991 A2	14-03-2023
		CA 3187261 A1	30-12-2021
		CN 116194151 A	30-05-2023
		EP 4168556 A2	26-04-2023
		IL 299196 A	01-02-2023
		KR 20230042005 A	27-03-2023
		WO 2021262909 A2	30-12-2021