

LIPID NANOPARTICLES FOR DELIVERY OF THERAPEUTIC PAYLOADS TO CELLS

5 CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of, and priority to, U.S. Provisional Application 63/582,457, filed on September 13, 2023; U.S. Provisional Application 63/592,195, filed on October 22, 2023; and U.S. Provisional Application 63/677,892, filed on July 31, 2024; the contents of each of which are hereby incorporated herein by reference in
10 their entirety.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] The content of the electronic sequence listing (252052001440seqlist.xml; Size: 167,341 bytes; and Date of Creation: September 13, 2024) is herein incorporated by reference in its entirety.

15 BACKGROUND

[0003] The development of lipid nanoparticles (LNPs) has recently made significant advances towards intracellular delivery of payloads such as nucleic acids (e.g., mRNA or siRNA). LNPs are generally comprised of multiple components including an ionizable lipid, a PEGylated lipid, a helper lipid and cholesterol, all of which play important roles in effectively
20 delivering the payload to diseased tissue.

[0004] Despite advances in the development of LNPs, substantial challenges still remain. Specifically, LNPs that provide a sufficient quantity of a payload to target cells has been challenging. In particular, development of LNPs that can deliver larger payloads and/or multiple payloads, such as those used in gene editing, is an important goal.

25 [0005] Additionally, in vivo delivery of therapeutic payloads to cells, such as to hematopoietic stem cells (HSCs) and immune cells such as T cells, is currently being investigated but can be challenging. For example, specific delivery of a genetic medicine payload, such as a gene therapy or gene editor agent, to the rare long-term HSC (LT-HSC) population or to T cells be difficult as any therapeutic agent delivered systemically must evade
30 or overwhelm myriad mechanisms that will prevent its delivery to the relevant microenvironment (for example, liver uptake, phagocytic cell uptake, complement activation,

opsonization, phagocytic cell uptake, etc.). Furthermore, after delivery to the target cell, the payload(s) must be able to disassociate from the endosome so it can be expressed in sufficient quantities in the targeted cell.

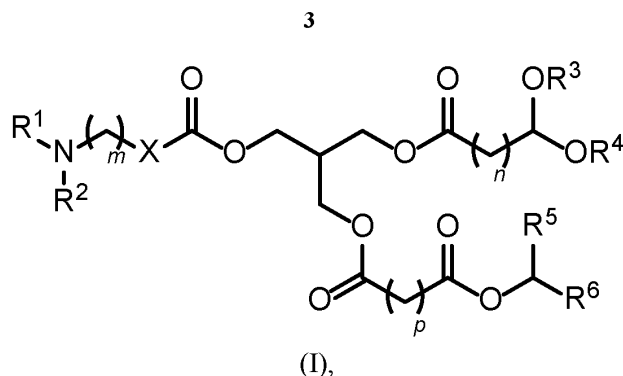
5 [0006] Moreover, LNPs often exhibit unfavorable toxicological profiles, that precludes their use as therapeutic delivery vehicles. LNPs may accumulate and deliver payloads to cells other than the intended target, which results in potential toxicity. The unfavorable safety profile often stem from the ionizable lipid used in the LNP.

10 [0007] Accordingly, there exists a need to develop LNPs with ionizable lipids that have favorable safety profiles and that can be used to promote cell-type specificity, high transfection efficiency, and a high degree of endosomal escape.

SUMMARY OF THE INVENTION

15 [0008] The disclosure provides lipid nanoparticles (LNPs) or conjugates comprising a lipid nanoparticle (LNP) encapsulating a payload, such as a therapeutic agent (e.g., a gene modifying system) or reporter construct, for delivery to cells. In some embodiments, the cells are liver cells. In some embodiments, the cells are T cells. In some embodiments, the cells are hematopoietic stem cells (HSCs) or hematopoietic progenitor cells (HPCs). In some
20 embodiments, the cells are long-term hematopoietic stem cells (LT-HSCs). In some cases, the LNPs and conjugates described herein can deliver the payload to cells (e.g., T cells, HSCs, or liver cells) more effectively than a baseline conjugate or baseline LNP. In some embodiments, the disclosure also provides conjugates comprising a lipid nanoparticle (LNP) encapsulating a
25 payload, such as a therapeutic agent (e.g., a gene modifying system) or reporter construct, for delivery to cells (e.g., T cells, HSCs, or liver cells) that have a targeting moiety on its surface. The conjugates having at least one targeting moiety described herein will also be referred to as targeted LNPs or tLNPs.

 [0009] In one aspect, the disclosure provides an ionizable lipid that is a compound of formula (I):



or a pharmaceutically acceptable salt thereof, wherein:

X is -O- or -CH₂-;

m is 0, 1, 2, or 3;

R¹ is C₁₋₄alkyl;

R² is C₁₋₄alkyl;

n is 1, 2, 3, or 4;

R³ is C₄₋₁₀alkyl;

R⁴ is C₄₋₁₀alkyl;

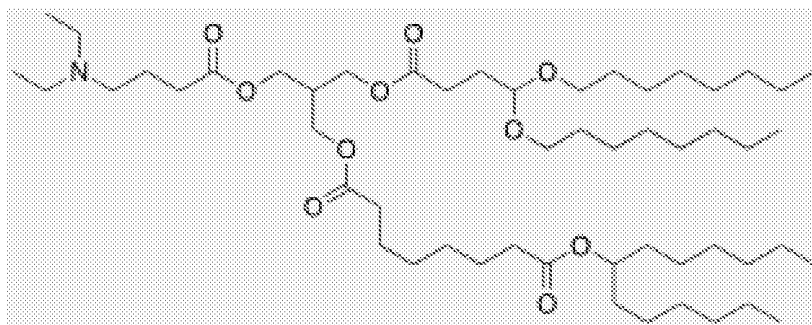
p is 2, 3, 4, 5, or 6;

R⁵ is C₄₋₁₀alkyl; and

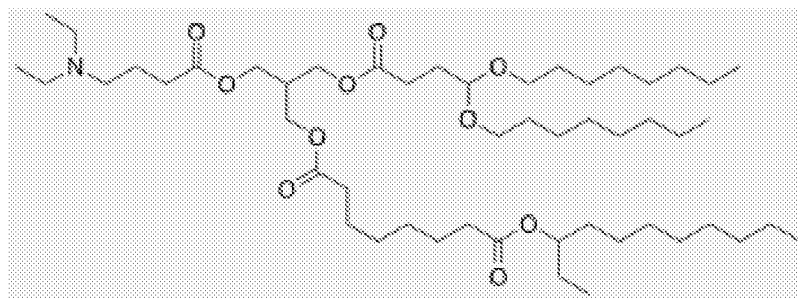
R⁶ is C₄₋₁₀alkyl.

[0010] In another aspect, the disclosure provides an ionizable lipid has a structure as set forth in Table 1, as disclosed herein.

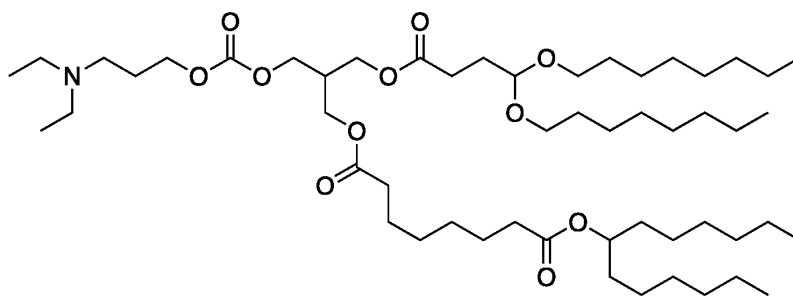
[0011] In some embodiments, the ionizable lipid has one of the structures depicted below:



Lipid 154

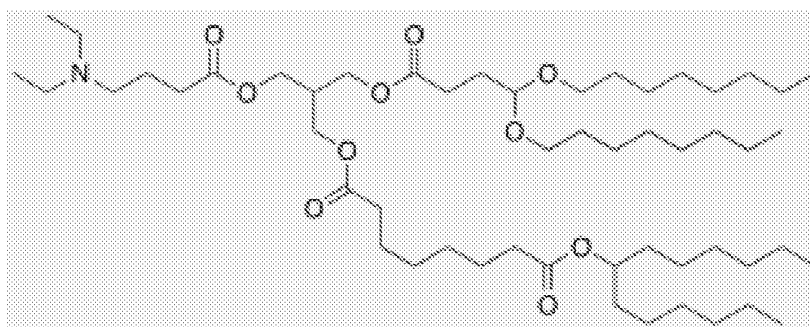


Lipid 155

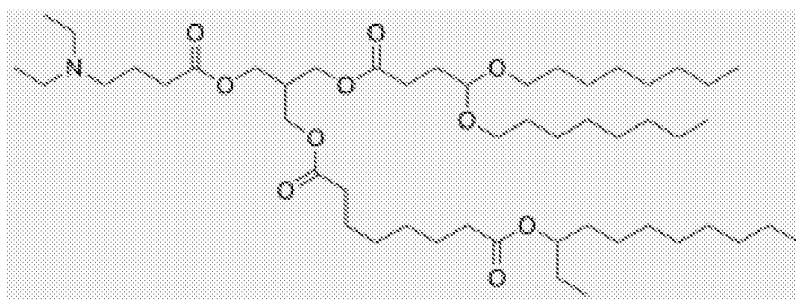


Lipid 232

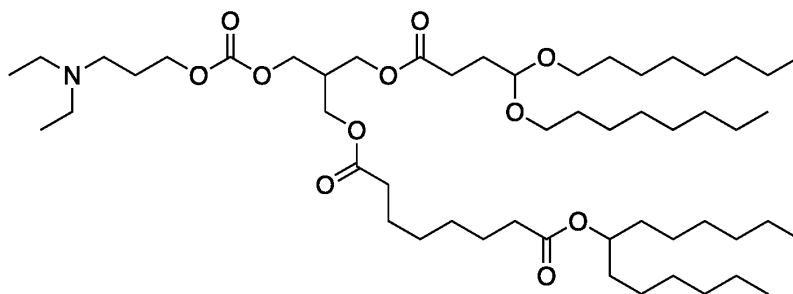
[0012] In some embodiments, the ionizable lipid has one of the structures depicted below:



Lipid 154

**Lipid 155**

[0013] In some embodiments, the ionizable lipid has the structure depicted below:

**Lipid 232**

5

[0014] In another aspect, the disclosure provides a lipid nanoparticle (LNP) for delivery of a therapeutic agent encapsulated in the LNP (to, e.g., T cells, HSCs, or liver cells), wherein the LNP comprises an ionizable lipid and a helper lipid, wherein the ionizable lipid is a compound of formula (I). In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, and Lipid 263. In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, Lipid 232, Lipid 262, and Lipid 263. In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, and Lipid 232. In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155 and Lipid 163. In some embodiments, the ionizable lipid is Lipid 154. In some embodiments, the ionizable lipid is Lipid 155. In some embodiments, the ionizable lipid is Lipid 163. In some embodiments, the ionizable lipid is Lipid 209. In some embodiments, the ionizable lipid is Lipid 232. In some embodiments, the ionizable lipid is Lipid 262. In some embodiments, the ionizable lipid is Lipid 263.

[0015] In some embodiments, the LNP is delivered in vivo to a patient in need thereof. In some embodiments, the LNPs are delivered ex vivo to the isolated cells (e.g., T cells, HSCs, or liver cells) of a subject (e.g., human patient) in need thereof. In some such embodiments, HSCs are first collected from the bone marrow of a subject. Following isolation of the desired cell subpopulation, the targeted LNPs of the disclosure can be mixed with the isolated cells, thereby resulting in effective transduction of the payload (e.g., a gene modifying system) into the cells. The modified cells can be administered to the patient for engraftment, as described in the art.

[0016] In another aspect, the disclosure provides conjugates (targeted LNPs) comprising an LNP and a targeting moiety. In some cases, inclusion of a targeting moiety of a targeted LNP enhances the ability of the LNP to deliver the payload to cells. In some embodiments, the cells are T cells. In other embodiments, the cells are HSCs and/or HPCs (collectively HSPCs). The conjugates (targeted LNPs) can be formulated in a pharmaceutical composition and can be directly administered to a subject (e.g., patient) in need thereof (e.g., by in vivo administration). Once administered, the compositions can deliver significant quantities of a therapeutic agent to the targeted cells (e.g. T cells, liver cells, or LT- HSCs of the subject). Alternatively, the conjugates can be administered to isolated cells such as T cells, HSCs (e.g., LT-HSCs), or liver cells ex vivo.

[0017] In some embodiments, the targeting moiety of the conjugate is an antibody, Fab fragment or single chain variable fragment (scFv), a DARPIN, a VHH domain antibody, a FN3 domain, a nanobody, a single domain antibody or a Centyrin. In other embodiments, the targeting moiety is a folate moiety, an antibiotic mimetic, a polynucleotide (such as a DNA or RNA aptamer), a carbohydrate, a vitamin or a N-Acetylgalactosamine (GalNac). In some embodiments, the targeting moiety is a peptide or protein, such as a ligand or part of a ligand, that binds to a receptor (e.g., a receptor on the surface of a T cell, HSC, or liver cells). In some embodiments, the targeting moiety is a peptide or protein that binds to a receptor or ligand on the surface of a T cell, HSC, HPC or LT-HSC, wherein the affinity of the peptide or protein targeting moiety for the receptor or ligand on the surface of the T cell, HSC, HPC or LT-HSC is modulated via phage display.

[0018] In some embodiments, the conjugate comprises an ionizable lipid and a helper lipid, wherein the ionizable lipid is a compound of formula (I). In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, and

Lipid 263. In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, Lipid 232, Lipid 262, and Lipid 263. In some embodiments, the ionizable lipid is selected from Lipid 154, Lipid 155, and Lipid 232. In some embodiments, the lipid is selected from Lipid 154, Lipid 155 and Lipid 163. In some embodiments, the lipid is Lipid 154. In some
5 embodiments, the lipid is Lipid 155. In some embodiments, the lipid is Lipid 163. In some embodiments, the ionizable lipid is Lipid 209. In some embodiments, the ionizable lipid is Lipid 232. In some embodiments, the ionizable lipid is Lipid 262. In some embodiments, the ionizable lipid is Lipid 263.

[0019] In some embodiments where the cells are T cells, the cells can be activated prior
10 to ex vivo delivery of the LNPs or conjugates. In other embodiments, the T cells are not activated prior to the delivery of the LNPs or conjugates ex vivo to T cells.

[0020] The disclosure further provides delivery of particular payloads to cells, wherein the payloads are encapsulated within LNPs or conjugates of the disclosure. In some embodiments, the payload can be a nucleic acid encoding a reporter gene, such as green
15 fluorescent protein (GFP). In some embodiments, the payload can be a therapeutic agent, for example, a therapeutic peptide or protein, a nucleic acid comprising a therapeutic agent, or a nucleic acid encoding a therapeutic agent. In some embodiments, the therapeutic agent can be a genetic medicine, wherein the therapeutic agent is capable of modifying, altering or effecting a change in the genomic DNA of a cell (e.g., a cell in the subject). In some embodiments, the
20 therapeutic agent is a gene therapy agent or gene editing agent.

[0021] In certain embodiments, the payload (e.g., therapeutic agent) delivered by the LNP or conjugates of the disclosure can be a small molecule, peptide or protein, non-coding RNA (ncRNA), gRNA, siRNA or miRNA, a nucleic acid (e.g., mRNA) encoding a peptide or protein (e.g., a protein for replacement gene therapy, or a protein for modifying or altering the
25 genome or epigenome, e.g., a protein for gene editing), a nucleic acid encoding or comprising one or more components of a system for altering a genome (e.g., one or more components of a ribonucleoprotein (RNP) complex for editing or altering the genome or epigenome (e.g., for introducing insertion-deletion mutations (indels), base editing, epigenetic editing, or target-primed reverse transcription (TPRT), e.g., by a mechanism that requires a recombinase,
30 transposase, retrotransposase, helicase, reverse transcriptase, polymerase, deaminase, methylase, demethylase, or ligase function), or combinations thereof. In certain embodiments, the therapeutic agent (i.e., payload) delivered by the LNP or targeted LNP can be a gene

modifying protein, a nucleic acid encoding a gene modifying protein, or a gene modifying system, as described herein. In some embodiments, the therapeutic agent (i.e., payload) comprises a template RNA, as described herein. In some embodiments, the template RNA, as described herein, e.g., a template RNA of a gene modifying system. In some embodiments, the template RNA is used for insertion of a nucleic acid sequence at a specific location of a genomic DNA. In some embodiments, the template RNA comprises a spacer of Table 5, Table 6, or Table 11. In some embodiments, template RNA comprises a heterologous sequence of Table 7, Table 8, or Table 12. In some embodiments, the template RNA comprises a PBS sequence of Table 13. In some embodiments, the template RNA comprises a heterologous object sequence and a PBS sequence of Table 14. In some embodiments, the template RNA comprises a sequence of Table 9, Table 10, or Table 16.

[0022] In some embodiments, an LNP or conjugate comprising an ionizable lipid as described herein is capable of delivering a payload (e.g., a therapeutic payload) to 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 80%, 85%, 90%, 95%, or 99% more T cells, HSCs, HPSCs, LT-HSCs, or liver cells than an LNP comprising a baseline lipid (such as V003). In some embodiments, an LNP comprising an ionizable lipid as described herein is capable of delivery a payload (e.g., a therapeutic payload) to 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 80%, 85%, 90%, 95%, or 99% more T cells, HSCs, HSPCS, LT-HSCs, or liver cells than an LNP comprising a baseline lipid (such as V003) in an animal model, as described herein. In some embodiments, the LNP is a conjugate.

[0023] In some embodiments, a payload (e.g., a therapeutic payload) delivered to T cells, HSCs, HPCs, LT-HSCs, or liver cells by an LNP comprising an ionizable lipid as described herein is expressed at levels that are at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% greater than a payload comprising a baseline lipid (such as V003). In some embodiments, a payload (e.g., a therapeutic payload) delivered to T cells, HSCs, HPCS, LT-HSCs, or liver cells by an LNP comprising an ionizable lipid as described herein is expressed at levels that are at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% greater than a payload comprising a baseline lipid (such as V003). In some embodiments, the LNP is a conjugate.

[0024] In some embodiments, at least 2%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, or at least 40% of T cells in a subject comprise a genome that have been modified or altered following administration of an LNP described herein, such as a targeted LNP (conjugate) encapsulating a therapeutic agent (e.g., a system for modifying, altering or editing a genome). In some embodiments, from about 5% to about 50%, from about 5% to about 20%, from about 10% to about 30%, from about 20% to about 50%, from about 20% to about 40%, or from about 20% to about 30% of T cells in a subject comprise a genome that has been modified or altered following administration of an LNP described herein, such as a conjugate.

10 [0025] In some embodiments, at least 2%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, or at least 40% of HSCs (e.g., LT-HSCs) in a subject comprise a genome that have been modified or altered following administration of an LNP described herein, such as a targeted LNP (conjugate) encapsulating a therapeutic agent (e.g., a system for modifying, altering or editing a genome). In some embodiments, from about 5% to about 50%, from about 5% to about 20%, from about 10% to about 30%, from about 20% to about 50%, from about 20% to about 40%, or from about 20% to about 30% of HSCs in a subject comprise a genome that has been modified or altered following administration of an LNP described herein, such as a conjugate.

20 [0026] In some embodiments, at least 2%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, or at least 40% of liver cells in a subject comprise a genome that have been modified or altered following administration of an LNP described herein, such as a targeted LNP (conjugate) encapsulating a therapeutic agent (e.g., a system for modifying, altering or editing a genome). In some embodiments, from about 5% to about 50%, from about 5% to about 20%, from about 10% to about 30%, from about 20% to about 50%, from about 20% to about 40%, or from about 20% to about 30% of liver cells in a subject comprise a genome that has been modified or altered following administration of an LNP described herein, such as a conjugate.

30 [0027] In some embodiments, the LNPs of the present disclosure, e.g., conjugates, comprise one or more pegylated lipid molecules. In some embodiments, the targeting moiety (e.g., antibody, Fab fragment or scFv) of a conjugate is associated with or chemically bonded to at least one of the pegylated lipid molecules. In some embodiments, the conjugate comprises from about 0.05 mol % to about 2 mol % of the pegylated lipid bonded to the targeting moiety.

In some embodiments, the PEG spacer between the lipid and the targeting moiety comprises at least about 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 200, or 110 ethylene glycol units. In some embodiments, the PEG spacer comprises about 10-120 ethylene glycol units. In some embodiments, the molecular weight of the pegylated lipid bonded to the targeting moiety is from about 500 (i.e., PEG500) to about 5,000 (i.e., PEG5000). In some embodiments, the molecular weight of the pegylated lipid bonded to the targeting moiety is from about 1,000 (i.e., PEG1000) to about 3,000 (i.e., PEG5300). In some embodiments, the lipid component of the pegylated lipid bonded to the targeting moiety is selected from DMG, DPG, DSG, DTA, DOPE, DPPE, DMPE, DSPE, sphingosine, sphingomyelin, and stearic acid.

[0028] As disclosed herein, different pegylated lipids in the targeted LNPs of the disclosure can imbue different properties to the LNPs, including pharmacokinetics, distribution and endosomal release. Surprisingly, it has been discovered that LNPs comprising different pegylated lipids not covalently bound to a targeting moiety could show different toxicological properties when all other components of the LNPs (e.g., ionizable lipid) and relative proportions are the same. In particular, it has been discovered that LNPs comprising pegylated lipids with PEG lipid anchors having 16 carbon atoms (i.e., C16 PEG lipid anchors) or 18 carbon atoms (i.e., C18 PEG lipid anchors) could reduce or eliminate the toxicity associated with identical LNPs other than the pegylated lipid component. In some embodiments, administration (in vivo or ex vivo) of the LNPs comprising C16 PEG lipid anchors or C18 PEG lipid anchors does not result in substantial liver toxicity. In some embodiments, LNPs comprising pegylated lipids with PEG lipid anchors having 16 carbon atoms (i.e., C16 PEG lipid anchors) or 18 carbon atoms (i.e., C18 PEG lipid anchors) exhibit longer circulation times *in vivo* relative to LNPs comprising pegylated lipids with shorter PEG lipid anchors (e.g., C14 PEG lipid anchors). In some embodiments, administration of the targeted LNPs results in no or minimal increases in serum levels of the secreted liver enzyme alanine aminotransferase (ALT). Reduction of liver exposure is particularly useful when targeted delivery of LNPs to particular cell types is desired, such as delivery to immune cells or HSCs, as disclosed herein.

[0029] In other embodiments, transduction of the LNPs into liver cells (e.g., hepatocytes) is desirable. For instance, transduction of the LNPs into liver cells is desirable to treat liver diseases and conditions. In some embodiments, the LNPs useful for treating liver diseases and conditions do not include a targeting moiety. In some embodiments, the LNPs useful for treating liver diseases and conditions comprise N-acetylgalactosamine (GalNac).

[0030] In some embodiments, the pegylated lipid has at least one C16 (palmitoyl) PEG lipid anchor. In some embodiments, the pegylated lipid has two C16 PEG lipid anchors (i.e., dialkyl chains of 16 carbons long). In some embodiments, the pegylated lipid is 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine (DPPE-PEG2000.). In some embodiments, the pegylated lipid is 1,2-Dipalmitoyl-rac-glycero-3-methylpolyoxyethylene (DPG-PEG2000). In some embodiments, the pegylated lipid is C16 PEG ceramide. In some embodiments, the targeted LNPs comprising the C16 pegylated lipids show significantly reduced liver uptake than otherwise identical LNPs comprising C14 pegylated lipids.

[0031] In some embodiments, the pegylated lipid has at least one C18 PEG lipid anchor. In some embodiments, the pegylated lipid has two C18 PEG lipid anchors (i.e., dialkyl chains of 18 carbons long). In some embodiments, the C18 pegylated lipid is 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (DSPE-PEG2000). In some embodiments, the C18 pegylated lipid is distearoyl-rac-glycerol-PEG2000 (DSG-PEG2000).

[0032] In some embodiments, the pegylated lipid has at least one C14 PEG lipid anchor. In some embodiments, the pegylated lipid has two C14 PEG lipid anchors (i.e., dialkyl chains of 14 carbons long).

[0033] In some embodiments, an LNP of the present disclosure, e.g., conjugate, comprises one or more non-pegylated lipids (e.g. non-pegylated phospholipids). Such lipids are often referred to as helper lipids. In some embodiments, the targeting moiety (e.g., antibody, Fab fragment or ScFv) is associated with or chemically bonded to at least one of the non-pegylated lipids. In some embodiments, the non-pegylated lipid (helper lipid) is selected from POPC, DOPC, DOPE, and DSPC. In some embodiments, the non-pegylated lipid (helper lipid) is a sphingolipid. In some such embodiments, the non-pegylated lipid is a sphingomyelin. In some embodiments, the sphingomyelin has a head group selected from, phosphocholine, phosphoethanolamine or ceramide. In some embodiments, the sphingomyelin is egg sphingomyelin.

[0034] In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP, or conjugate, is from about 18% to about 32%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP or conjugate is from about 20% to about 30%. In some embodiments, the mol% of the helper

lipid (e.g., DSPC or sphingomyelin) in the LNP or conjugate is from about 22% to about 28%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP or conjugate is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, about 30%, about 31%, or about 32%. It will be understood that mol% of the helper lipid as used herein refers to the mol% of the total lipid component of the LNP, which does not include the therapeutic agent (i.e., payload) or the targeting moiety.

[0035] As set forth in the examples below, in vivo delivery of certain payloads following administration of the disclosed LNPs with these percentages of helper lipids provides enhanced transduction and/or expression of the payloads relative to LNPs with smaller or larger quantities of helper lipid.

[0036] In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) in the conjugate is from about 1:1 to about 7:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 4:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 3:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC) is from about 1:1 to about 2.5:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 2:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC) is from about 1.5:1 to about 2.5:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 2:1 to about 2.5:1.

[0037] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 154 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 154 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0038] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 155 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 155 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0039] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 163 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 163 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0040] In some embodiments, the LNP, e.g., targeted LNP (conjugate), comprises Lipid 173 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 173 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0041] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 209 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 209 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further

comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0042] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 232 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 232 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0043] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 262 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 262 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0044] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 263 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 263 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%. In some embodiments, the LNP further comprises a pegylated lipid comprising at least one C16 alkyl chain (e.g., two C16 alkyl chains). In some such embodiments, the pegylated lipid is DPPE-PEG2000 or DPG-PEG2000.

[0045] In any of the foregoing embodiments, the LNP or conjugate can comprise one or more cholesterol molecules. In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 6:1 to about 0.5:1. In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 3:1 to about 0.5:1.

In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 2:1 to about 0.5:1. In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1.5:1 to about 0.5:1. In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 0.5:1. In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:2 to about 0.8:1.

[0046] In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that targets a receptor on the surface of a HSC selected from CD33, CD34, CD38, CD43, CD46, CD49d, CD50, CD59, CD71, CD105, CD117, CD123, CD164, CD184, and CD338. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD33. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD34. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD38. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD43. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD46. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD49d. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to C503. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD59. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD71. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD105. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD117. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD123. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD164. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD184. In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that binds to CD338.

[0047] In some embodiments, a targeted LNP (conjugate) for delivery of a therapeutic agent to HSCs, as described herein, comprises:

a plurality of targeting moieties conjugated to the LNP, wherein the plurality of targeting moieties bind to CD33, CD34, CD38, CD43, CD59, CD105, CD123, CD164, CD338, CD71, CD117, CD50, CD49d, CD46, or CD184 (CXCR4); and
a therapeutic agent encapsulated within the LNP.

5 **[0048]** In certain embodiments, the targeting moiety of the targeted LNP (conjugate) is a T-cell targeting moiety, for example, an antibody, Fab fragment or scFv that binds to a T-cell antigen selected from the group consisting of CD2, CD3, CD4, CD5, CD7, CD8, CD28, CD137, CD45, T-cell receptor (TCR) β , TCR- α , TCR- α/β , TCR- γ/δ , PD1, CTLA4, TIM3, LAG3, CD18, IL-2 receptor, CD11a, TLR2, TLR4, TLR5, IL-7 receptor, and IL-15 receptor.
10 In some embodiments the CD80 targeting moiety is a CD80 extracellular domain (ECD).

[0049] In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that targets a receptor on the surface of the T cell selected from CD2, CD3, CD4, CD5, CD6, CD7 and CD8. In some embodiments, the targeting moiety targets a CD3 receptor on the surface of the T cell. In some embodiments, the targeting moiety targets a CD7 receptor
15 on the surface of the T cell. In some embodiments, the targeting moiety targets a CD5 receptor on the surface of the T cell.

[0050] In some embodiments, the targeted LNP (conjugate) comprises a targeting moiety that targets CD3 on the surface of the T cell, wherein the targeting moiety is an antibody, Fab fragment or scFv selected from SP34, teclistamab, mosunetuzumab,
20 odronextamab, tebentafusp, tepilizumab, Muromonab and visilizumabm or antigen-binding portions thereof. In certain embodiments, the targeting moiety is SP34 or an antigen-binding portion thereof. In other embodiments, the targeting moiety is teclistamab or an antigen-binding portion thereof. In other embodiments, the targeting moiety is visilizumab or an antigen-binding portion thereof. In other embodiments, the targeting moiety is tepilizumab or
25 an antigen-binding portion thereof.

[0051] In some embodiments, the LNP comprises a GalNAc.

[0052] In some embodiments, the conjugates are formed via a Click reaction between a first Click handle on the targeting moiety (e.g., antibody, Fab fragment, or scFv) and a second Click handle on the LNP, thereby generating a Click product. The second Click handle can be
30 on one or more of the lipids comprising the LNP. In some embodiments, the second Click handle is covalently bonded to a pegylated lipid comprising the LNP.

[0053] In one embodiment, the Click product can be formed using a copper-catalyzed Click reaction. One such copper-catalyzed Click reaction is a Huisgen 1,3-dipolar cycloaddition (CuAAC) between an azide and an alkyne. In some embodiments, the first or second Click handle comprises a cyclic derivative of the alkynyl group. In some embodiments, the cyclic derivative of the alkynyl group is selected from dibenzocyclooctyne, cyclooctyne, and difluorinated cyclooctyne. In some embodiments, the click chemistry involves strain promoted cycloaddition of azides. In some embodiments, the click chemistry is based upon reaction of strained alkenes. In another embodiment, the Click product can be formed using copper-free Click chemistry. For example, the Click product can be formed between an azide and dibenzocyclooctene (DBCO). Alternatively, the Click product can be formed using a Staudinger reaction between an azide and a phosphine, hence producing an aza-ylide. In some embodiments, the Click product can be formed from an inverse electron demand Diels-Alder reaction between a trans-cyclooctene (TCO) moiety on the first or second Click handle and a tetrazine ring on the first or second Click handle. In some embodiments, the first Click handle comprises a tetrazine (Tz) ring and the second Click handle comprises a TCO moiety. In some embodiments, the tetrazine ring is unsubstituted. In some such embodiments, the tetrazine ring is methyltetrazine. In some embodiments, the tetrazine ring is a 6-methyl substituted tetrazine.

[0054] In some embodiments, the Click product is formed by conjugating an LNP to a targeting moiety that has been modified with an enzyme recognition sequence. For instance, an antibody, Fab fragment or single chain variable fragment (ScFv), can be covalently linked to a first Click handle through a linker comprising an enzyme recognition sequence and the LNP is covalently linked to a second Click handle on the LNP, thereby generating a Click product. In some embodiments, the enzyme recognition sequence is a sortase recognition motif or a LplA acceptor peptide. In some embodiments, the antibody Fab fragment or ScFv is directly bonded to the enzyme recognition sequence. In some embodiments, the antibody Fab fragment or ScFv is bonded to the enzyme recognition sequence via one or more amino acid residues. Particular amino acid residues added that can be covalently attached to the C-terminus of the antibody, Fab fragment or ScFv include, but are not limited to (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10.

[0055] In some embodiments, the conjugates are formed by a method comprising:

(i) contacting a Fab fragment with a reducing reagent, wherein the Fab fragment comprises a heavy chain and a light chain and an interchain disulfide bond linking the constant light chain domain (C_L) and the constant heavy chain domain 1 (C_{H1}), whereby the reducing reagent reduces the interchain disulfide bond of the Fab fragment to generate two free cysteine residues; and

(ii) contacting the product of step (i) with a precursor LNP comprising a plurality of thiol-reactive groups covalently bonded to one or more lipids of the precursor LNP, thereby forming a targeted LNP.

[0056] In some embodiments, the conjugates are formed by a method comprising:

(i) contacting a Fab fragment with a reducing reagent, wherein the Fab fragment comprises a heavy chain and a light chain and an interchain disulfide bond linking the constant light chain domain (C_L) and the constant heavy chain domain 1 (C_{H1}), whereby the reducing reagent reduces the interchain disulfide bond of the Fab fragment to generate two free cysteine residues; and

(ii) contacting the product of step (i) with a precursor LNP comprising a plurality of thiol-reactive groups covalently bonded to one or more lipids of the precursor LNP, thereby forming a targeted LNP.

[0057] In some aspects, the targeted delivery of a system for modifying or altering genomic DNA (e.g., a gene modifying system) with an LNP comprising two or more HSC specific binders conjugated to the LNP results in synergistic enhancement of gene editing or modification of the target HSCs, LT-HSCs or immune cells (e.g., T cells), wherein the LNP comprises a lipid of formula (I) (e.g., Lipid 154).

[0058] A targeted LNP in some embodiments comprises two or more different targeting moieties, each of which may bind to the same or a different target. A targeted LNP having two different targeting moieties which bind to the same or different targets may be referred to as a “dual binder” or “dual binding” LNP. For instance, in some embodiments, each of the targeting moieties is conjugated to the lipid nanoparticle through a linker, wherein each linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such embodiments, the targeting moieties are antibodies or antigen binding fragments thereof. In other such

embodiments, the targeting moieties are ScFvs. In some embodiments, the targeting moieties are Fab fragments.

[0059] In some embodiments, the LNPs comprising dual binders are capable of modifying at least 5% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying at least 10% of the genomes of target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying at least 15% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying at least 20% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying at least 25% of the genomes of the HSCs or LT-HSCs. In some embodiments, the LNPs comprising dual binders are capable of modifying at least 30% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying at least 30% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying from about 10% to about 30% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying from about 20% to about 30% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In some embodiments, the LNPs comprising dual binders are capable of modifying from about 10% to about 20% of the genomes of the target HSCs, LT-HSCs or immune cells (e.g., T cells). In any of the foregoing embodiments, the system for modifying or altering genomic DNA elicits a single-stranded break in the genomic DNA. In some such embodiments, system for modifying or altering genomic DNA includes a gene modifying polypeptide and a template RNA for a nucleic acid sequence to be inserted at a specific location of the genomic DNA, hence resulting in modification of the genomic DNA. In some embodiments, the gene modifying system comprises a Cas9 nickase fused to a reverse transcriptase (RT) domain. In some embodiments, system for modifying or altering genomic DNA induces target-primed reverse transcription (TPRT) to insert a heterologous nucleic acid sequence into a DNA sequence, e.g., genomic DNA.

[0060] In some embodiments, the dual binders can be added to the surface of the LNP by methods described in Section III of the Detailed Description. For instance, in some embodiments, each of the targeting moieties is conjugated to the lipid nanoparticle through a linker, wherein each linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such
5 embodiments, the targeting moieties are antibodies or antigen binding fragments thereof. In other such embodiments, the targeting moieties are scFvs. In some embodiments, the targeting moieties are Fab fragments.

[0061] In some aspects, the disclosure provides a method of treating sickle cell disease
10 in a subject (e.g., a human patient) in need of, said method comprising administering to the subject a targeted lipid nanoparticle (LNP) comprising:

- i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);
- ii. a helper lipid;
- 15 iii. one or more targeting moieties conjugated to the LNP, wherein the one or more targeting moieties bind to one or more target on the surface of the HSCs (e.g., LT-HSCs); and
- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene
20 modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the targeted LNP comprises two RNA molecules.

[0062] In some aspects, the disclosure provides a method of treating sickle cell disease
25 in a subject (e.g., a human patient) in need of, said method comprising administering to the subject a targeted lipid nanoparticle (LNP) comprising:

- i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);
- ii. a helper lipid;
- 30 iii. one or more targeting moieties conjugated to the LNP, wherein the one or more targeting moieties bind to one or more target on the surface of the HSCs (e.g., LT-HSCs); and

- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the targeted LNP comprises two RNA molecules.

[0063] In some aspects, the disclosure provides a method of treating sickle cell disease in a subject (e.g., a human patient) in need of, said method comprising administering to the subject a targeted lipid nanoparticle (LNP) comprising:

- i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);
- ii. a helper lipid;
- iii. two different targeting moieties conjugated to the LNP, wherein each of the two different targeting moieties binds to a target on the surface of the HSCs (e.g., LT-HSCs); and
- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the targeted LNP comprises two RNA molecules.

[0064] In some embodiments, the dual binders targeting HSCs can be added to the surface of the LNP by methods described in Section III of the Detailed Description. For instance, in some embodiments, each of the targeting moieties is conjugated to the lipid nanoparticle through a linker, wherein each linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such embodiments, the targeting moieties are antibodies or antigen binding fragments thereof. In other such embodiments, the targeting moieties are ScFvs. In some embodiments, the targeting moieties are Fab fragments.

[0065] In some embodiments, the nucleic acid components do not introduce double stranded breaks into the genomic DNA. In some embodiments, greater than 5% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the

targeted LNP. In some embodiments, greater than 10% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, greater than 15% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, greater than 20% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, greater than 30% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, from about 5% to about 15% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, from about 10% to about 30% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, from about 10% to about 20% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, from about 15% to about 20% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP. In some embodiments, from about 15% to about 25% of the HSCs or LT-HSCs of the patient with sickle cell disease are altered following administration of the targeted LNP.

[0066] In some embodiments, following in vivo administration of the targeted LNPs to the patient with sickle cell disease, the HSCs or LT-HSCs maintain their hematopoietic potential and function to support multilineage development. In some embodiments, the gene modifying system corrects the E6V HBB mutation in the HSCs and LT-HSCs of the patient.

[0067] In some aspects, the targeted delivery of a system for modifying or altering genomic DNA (e.g., a gene modifying system) with an LNP comprising two or more HSC specific binders conjugated to the LNP results in synergistic enhancement of gene editing or modification of the target immune cells (e.g., T cells), wherein the LNP comprises a lipid of formula (I) (e.g., Lipid 154).

[0068] In some aspects, the disclosure provides a targeted lipid nanoparticle (LNP) for ex vivo or in vivo delivery of a therapeutic agent to immune cells (e.g., T cells), wherein the targeted LNP comprises:

- i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);

- ii. a helper lipid;
- iii. one or more targeting moieties conjugated to the LNP, wherein the one or more bind to one or more target on the surface of the immune cells (e.g., T cells); and
- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the targeted LNP comprises two RNA molecules.

10 **[0069]** In some aspects, the disclosure provides a targeted lipid nanoparticle (LNP) for ex vivo or in vivo delivery of a therapeutic agent to immune cells (e.g., T cells), wherein the targeted LNP comprises:

- i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);
- 15 ii. a helper lipid;
- iii. two different targeting moieties conjugated to the LNP, wherein each of the two different targeting moieties binds to a target on the surface of the immune cells (e.g., T cells); and
- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the targeted LNP comprises two RNA molecules.

25 **[0070]** In some embodiments, the dual binders targeting T cells can be added to the surface of the LNP by methods described in Section III of the Detailed Description. For instance, in some embodiments, each of the targeting moieties is conjugated to the lipid nanoparticle through a linker, wherein each linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such embodiments, the targeting moieties are antibodies or antigen binding fragments thereof. In other such embodiments, the targeting moieties are scFvs. In some embodiments, the targeting moieties are Fab fragments.

[0071] In some aspects, the disclosure provides a lipid nanoparticle (LNP) for ex vivo or in vivo delivery of a therapeutic agent to liver cells, wherein the LNP comprises:

i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);

5 ii. a helper lipid; and

iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the LNP comprises two RNA molecules.

10

In some embodiments, the LNP for delivery of a therapeutic agent to liver cells comprises a galNAc.

[0072] In some aspects, the disclosure provides a method of treating a liver disease in a subject (such as, but not limited to, alpha-1 antitrypsin deficiency or phenylketonuria in a human patient) in need of, said method comprising administering to the subject a lipid nanoparticle (LNP) comprising:

15

i. an ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);

20 ii. a helper lipid; and

iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system). In some embodiments, the one or more nucleic acids are RNA molecules. In some embodiments, the LNP comprises two RNA molecules.

25

BRIEF DESCRIPTION OF THE DRAWINGS

[0073] FIG. 1 shows the formation of an LNP conjugated to the C-terminus of a Fab fragment through a sortase mediated ligation followed by a Click reaction.

30 [0074] FIG. 2 shows the formation of an LNP conjugated to the C-terminus of a Fab fragment through a lipoic acid ligase mediated ligation followed by a Click reaction.

[0075] FIG. 3 shows an enzymatic approach of site-specifically introducing a Tz ring onto a sugar moiety of an antibody.

[0076] FIG. 4 shows an example of a light-induced crosslinking approach of site-specifically introducing a Tz ring onto an antibody.

5 [0077] FIG. 5 shows an example of using a non-natural amino acid to site-specifically introduce a Tz ring onto an antibody or Fab fragment.

[0078] FIG. 6 shows an example of using a cysteine-maleimide reaction to site-specifically introduce a first Click handle onto a Fab fragment, which subsequently reacts with an LNP bonded to a second Click handle, thereby generating a targeted LNP (conjugate).

10 [0079] FIG. 7 shows a Fab fragment modified with a free cysteine designed to react with a maleimide group.

[0080] FIG. 8A shows a schematic example for producing a site-specific conjugate through the reaction of an anti-CD117 Fab fragment with site-specific sortase recognition motif and an LNP. In the schematic, the squiggly lines represent PEG bonded to a lipid of the LNP.

15 FIG. 8B shows a schematic of producing a conjugate through the reaction of a Fab fragment that has been modified using a random approach (NHS modification) and an LNP.

[0081] FIGs. 9A-9F show the synthesis of Lipid 154.

[0082] FIGs. 10A-10E show the synthesis of Lipid 155.

20 [0083] FIGs. 11A and 11B show enhanced transduction and/or expression levels of GFP in LSK cells using LNPs containing ionizable Lipid 154 or ionizable Lipid 155 relative to LNPs containing ionizable Lipid V003 in an in vivo mouse model. FIGs. 11C and 11D show enhanced transduction of GFP and/or enhanced GFP expression in in LT-HSCs using LNPs containing ionizable Lipid 154 or ionizable Lipid 155 relative to an LNPs containing ionizable Lipid V003.

25 [0084] FIGs. 12A and 12C show enhanced transduction of GFP in HSPCs or LT-HSPCs, respectively, using LNPs containing ionizable Lipid 154 or ionizable Lipid 155 relative to LNPs containing ionizable Lipid V003 in a humanized in vivo mouse model. FIGs. 12B and 12D show enhanced expression of GFP in HSPCs or LT-HSCs, respectively, using

LNPs containing ionizable Lipid 154 or ionizable Lipid 154 relative to LNPs containing ionizable Lipid V003 in a humanized *in vivo* mouse model.

[0085] **FIG. 13** shows results of screening of targeted LNP (tLNP) in humanized NBSGW mice with increased quantities of helper lipid. **FIGs. 13A** and **13B** show results for transduction and expression in HSPCs, respectively. **FIGs. 13C** and **13D** show results for transduction and expression in LT-HSCs, respectively.

[0086] **FIGs. 14A** and **14B** show that the administered anti-CD117 tLNPs can deliver payloads suitable for gene editing to human HSPCs in an animal model, leading to *in vivo* gene editing in HSPCs. **FIG. 14A** shows that tLNPs formulated with Lipid 154 and the wtCas9-RT mRNA and B2M gRNA payload resulted in 15.7% indels in B2M in HSPCs. **FIG. 14B** shows that the tLNPs formulated with Lipid 154 and the nCas9-RT and B2M template RNA payload generated high TPRT-mediated knockout of B2M in HPSCs (mean 5.2%).

[0087] **FIG. 15** shows that the Lipid 154 tLNP delivers GFP mRNA to liver tissue.

[0088] **FIG. 16** shows that Lipid 154 LNPs dosed at 4mpk causes a significant increase in ALT levels compared to the ALT levels in rats dosed at 2mpk or in the nontreated rat.

[0089] **FIG. 17** shows that the lipid-PEG type as well as mol % influences the number of human CD34+ cells transfected by the tLNPs (%GFP+ cells) and amount of GFP expression (MFI) in those cells. Increasing the percentage of Lipid-PEG in tLNPs results in a decrease of %GFP+ cells (**FIG. 17A**), as well as in MFI (**FIG. 17B**). Additionally, increasing the lipid anchor length (from C14 to C16 to C18) decreases the %GFP+ cells and MFI levels, even at the same mol %, both in the presence and absence of serum.

[0090] **FIG. 18A** shows that administration of tLNPs formulated with smaller anchor length PEG (C14) lead to higher GFP+ MFI levels in human LT-HSCs *in vivo*. Administration of tLNPs formulated with C18 anchor length PEG resulted in significantly lower GFP MFI. The administration of tLNPs formulated with C16 anchor length PEG, however, resulted in only a minimal decrease in MFI levels relative to the MFI levels generated using C14 anchor length PEG in LT-HSCs. **FIG. 18B** shows that the tLNPs formulated with C14-PEG exhibited a higher propensity for liver delivery compared to the tLNPs formulated with either the C16- or C18-PEG, regardless of the mol % tested in this study.

[0091] FIG. 19 shows that Lipid 154 LNP, formulated with 2% DPG-PEG2000 (C16), does not lead to elevated serum ALT levels compared to baseline levels.

[0092] FIG. 20A shows that administration of anti-CD117 tLNP formulated with Lipid 154 and 2 mol % C16-PEG in the macaques (n=2) resulted in high percentages of GFP+ HSPCs and LT-HSCs (96% GFP+ and 91% GFP+ cells, respectively). FIG. 20B shows that the expression of GFP mRNA dosed with the Lipid 154 tLNP.

[0093] FIG. 21 shows that administration of the anti-CD117/anti-CD34 dual tLNP formulated with Lipid 154 dual tLNP formulated with C16-PEG generated a significant number of human HSPCs *in vivo* comprising B2M indels (greater than 45% indels).

10 [0094] FIG. 22 shows that anti-CD117/anti-CD34 dual-targeted tLNP, formulated with Lipid 154 and C16 PEG led to a mean 11.8% perfect gene modification in HSPCs and mean 14.3 perfect gene modification in LT-HSCs *in vivo*.

[0095] FIG. 23A shows that administration of dual tLNP formulated with Lipid 154 to cynomolgus macaques (n=2) resulted in an average of over 90% GFP+ HSPCs and LT-HSCs for the anti-CD117/anti-CD34 dual tLNP, and an average of over 80% GFP+ HSPCs and over 70% GFP+ LT-HSCs for the anti-CD117/anti-CD71 dual tLNP. FIG. 23B shows that HSPCs and LT-HSCs in the NHPs expressed GFP following administration of each of the dual tLNP.

[0096] FIGs. 24A-24E show the synthesis of Lipid 232.

20 [0097] FIG. 25 shows an exemplary conjugate comprising a Fab fragment and a lipid nanoparticle (LNP).

[0098] FIG. 26A shows assembly of an LNP encapsulating a therapeutic payload with a functional group (e.g., maleimide, 2,3-dibromomaleimide, sortase tag, etc.) to be reacted with an antibody or antigen-binding fragment thereof. FIG. 26B shows assembly of an LNP encapsulating a therapeutic payload using a post-insertion technique.

25 [0099] FIG. 27 shows an exemplary schematic of the reduction of an interchain disulfide bond in a Fab fragment. After reduction, each cysteine residue is available to react with a thiol-reactive group conjugated to the surface of the LNP.

[0100] FIG. 28 shows an exemplary schematic of conjugate formation. A Fab fragment comprising an interchain disulfide bond between the heavy and light chain is contacted with a reducing reagent, whereby the reducing reagent reduces the interchain disulfide to generate two free cysteine residues (step (i)). In step (ii), the Fab fragment is contacted with an LNP comprising a plurality of thiol-reactive groups (e.g., a plurality of maleimide or DBM groups) conjugated to the surface of the LNP, whereby a thiol-reactive group of the plurality reacts with at least one of the two free cysteine residues of the Fab fragment.

[0101] FIG. 29 shows an exemplary schematic of conjugate formation. A Fab fragment comprising an interchain disulfide bond between the heavy and light chain is contacted with a reducing reagent, whereby the reducing reagent reduces the interchain disulfide to generate two free cysteine residues (step (i)). In step (ii), the Fab fragment is contacted with an LNP comprising a plurality of maleimide groups conjugated to the surface of the LNP, whereby a maleimide group of the plurality reacts with one of the two free cysteine residues of the Fab fragment.

[0102] FIG. 30 shows an exemplary schematic of conjugate formation using DBM as a bridging agent.

[0103] FIG. 31 shows an exemplary schematic of conjugate formation. In FIG. 31, a first Fab fragment (Fab1) and a second Fab fragment (Fab2) are each contacted with a reducing reagent, whereby the reducing reagent reduces the interchain disulfide between the heavy and light chain of each Fab. After step (i), each Fab comprises two free cysteine residues with one free cysteine residue on the heavy chain and one free cysteine residue on the light chain. In step (ii), the first Fab fragment and the second Fab fragment are contacted with an LNP comprising a plurality of thiol-reactive groups (e.g., maleimide or DBM) conjugated to the surface of the LNP, whereby a thiol-reactive group of the plurality reacts with at least one of the two free cysteine residues of the first Fab Fragment and another thiol-reactive group of the plurality reacts with at least one of the two free cysteine residues of the second Fab Fragment.

[0104] FIG. 32 shows some examples of a pegylated lipid bonded to a maleimide moiety.

[0105] FIG. 33 shows some examples of non-pegylated lipid bonded to a maleimide moiety.

[0106] FIG. 34 shows some examples of ionizable lipid bonded to a maleimide moiety.

[0107] FIG. 35 shows some examples of sterols bonded to a maleimide moiety.

[0108] FIG. 36A provides a bar graph showing that administration of Lipid 232 LNPs comprising an exemplary gene modifying system to transgenic mice comprising a human *SERPINA1* (*hSERPINA1*) gene with the PiZ mutation results an average correction of about 15.9% of the *SERPINA1* alleles at 8 days post-dosing. FIG. 36B shows that administration of each of the LNPs comprising the gene modifying system resulted in very low indel percentages.

[0109] FIG. 37A provides a bar graph showing that an average of 22.7% of the *SERPINA1* alleles in the hepatocytes of the dosed NHPs (2 mg/kg dosage) carried the intended precise edit in the *SERPINA1* gene following administration of the LNPs formulated with Lipid 232 and the exemplary gene modifying system. FIG. 37B shows that increasing the LNP dose to 3 mg/kg leads to a corresponding increase in perfect rewriting of the *SERPINA1* locus in NHP hepatocytes, resulting in an average of 42.7% of *SERPINA1* alleles having the intended precise edit. An average of approximately 76.3 percent of the liver cDNA carried the intended precise edit at *SERPINA1*.

[0110] FIG. 38 provides a bar graph showing that an average of about 36.3% of the *SERPINA1* alleles in the hepatocytes of the dosed NHPs carried the intended precise edit following administration of the GalNAc-conjugated LNPs formulated with Lipid 232 and an exemplary gene modifying system.

20

DETAILED DESCRIPTION

I. DEFINITIONS

[0111] **Antigen binding domain:** The term “antigen binding domain” as used herein refers to that portion of antibody or a chimeric antigen receptor which binds an antigen. In some embodiments, an antigen binding domain binds to a cell surface antigen of a cell. In some embodiments an antigen binding domain binds an antigen characteristic of a cancer, e.g., a tumor associated antigen in a neoplastic cell. In some embodiments, an antigen binding domain binds an antigen characteristic of an infectious disease, e.g. a virus associated antigen in a virus infected cell. In some embodiments, an antigen binding domain binds an antigen characteristic of a cell targeted by a subject’s immune system in an autoimmune disease, e.g., a 0-antigen. In some embodiments, an antigen binding domain is or comprises an antibody or

30

antigen-binding portion thereof. In some embodiments, an antigen binding domain is or comprises an scFv or Fab.

[0112] Expression cassette: The term “expression cassette,” as used herein, refers to a nucleic acid construct comprising nucleic acid elements sufficient for the expression of the nucleic acid molecule of the instant invention.

[0113] gRNA spacer: A “gRNA spacer”, as used herein, refers to a portion of a nucleic acid that has complementarity to a target nucleic acid and can, together with a gRNA scaffold, target a Cas protein to the target nucleic acid.

[0114] gRNA scaffold: A “gRNA scaffold”, as used herein, refers to a portion of a nucleic acid that can bind a Cas protein and can, together with a gRNA spacer, target the Cas protein to the target nucleic acid. In some embodiments, the gRNA scaffold comprises a crRNA sequence, tetraloop, and tracrRNA sequence.

[0115] Gene modifying polypeptide: A “gene modifying polypeptide”, as used herein, , refers to a polypeptide comprising a retroviral reverse transcriptase, or a polypeptide comprising an amino acid sequence having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity to a retroviral reverse transcriptase, which is capable of integrating a nucleic acid sequence (e.g., a sequence provided on a template nucleic acid) into a target DNA molecule (e.g., in a mammalian host cell, such as a genomic DNA molecule in the host cell). In some embodiments, the gene modifying polypeptide is capable of integrating the sequence substantially without relying on host machinery. In some embodiments, the gene modifying polypeptide integrates a sequence into a specific target site. In some embodiments, a gene modifying polypeptide includes one or more domains that, collectively, facilitate 1) binding the template nucleic acid, 2) binding the target DNA molecule, and 3) facilitate integration of the at least a portion of the template nucleic acid into the target DNA. Gene modifying polypeptides include both naturally occurring polypeptides as well as engineered variants of the foregoing, e.g., having one or more amino acid substitutions to the naturally occurring sequence. Gene modifying polypeptides also include heterologous constructs, e.g., where one or more of the domains recited above are heterologous to each other, whether through a heterologous fusion (or other conjugate) of otherwise wild-type domains, as well as fusions of modified domains, e.g., by way of replacement or fusion of a heterologous sub-domain or other substituted domain. Exemplary gene modifying

polypeptides, and systems comprising them and methods of using them, that can be used in the methods provided herein are described, e.g., in PCT/US2021/020948, which is incorporated herein by reference with respect to gene modifying polypeptides that comprise a retroviral reverse transcriptase domain. In some embodiments, a gene modifying polypeptide integrates a
5 a sequence into a gene. In some embodiments, a gene modifying polypeptide integrates a sequence into a sequence outside of a gene. A “gene modifying system,” as used herein, refers to a system comprising a gene modifying polypeptide and a template nucleic acid.

[0116] Gene modifying system: A “gene modifying system,” as used herein, refers to a system comprising a gene modifying polypeptide, or a nucleic acid (e.g., an mRNA) encoding
10 the gene modifying polypeptide, and a template nucleic acid.

[0117] Domain: The term “domain” as used herein refers to a structure of a biomolecule that contributes to a specified function of the biomolecule. A domain may comprise a contiguous region (e.g., a contiguous sequence) or distinct, non-contiguous regions (e.g., non-contiguous sequences) of a biomolecule. Examples of protein domains include, but
15 are not limited to, an endonuclease domain, a DNA binding domain, a reverse transcriptase domain; an example of a domain of a nucleic acid is a regulatory domain, such as a transcription factor binding domain.

[0118] Exogenous: As used herein, the term “exogenous,” when used with reference to a biomolecule (such as a nucleic acid sequence or polypeptide) means that the biomolecule
20 was introduced into a host genome, cell, or organism by the hand of man. For example, a nucleic acid that is as added into an existing genome, cell, tissue, or subject using recombinant DNA techniques or other methods is exogenous to the existing nucleic acid sequence, cell, tissue or subject.

[0119] First Strand and Second Strand: The terms “first strand” and “second strand”,
25 as used herein, describe the individual DNA strands of target DNA, distinguish the two DNA strands based upon which strand the reverse transcriptase domain initiates polymerization, e.g., based upon where target primed synthesis initiates. The first strand refers to the strand of the target DNA upon which the reverse transcriptase domain initiates

[0120] Heterologous: The term “heterologous”, ,” as used herein to describe a first
30 element in reference to a second element means that the first element and second element do not exist in nature disposed as described. For example, a heterologous polypeptide, nucleic acid

molecule, construct or sequence refers to (a) a polypeptide, nucleic acid molecule or portion of a polypeptide or nucleic acid molecule sequence that is not native to a cell in which it is expressed, (b) a polypeptide or nucleic acid molecule or portion of a polypeptide or nucleic acid molecule that has been altered or mutated relative to its native state, or (c) a polypeptide or nucleic acid molecule with an altered expression as compared to the native expression levels under similar conditions. For example, a heterologous regulatory sequence (e.g., promoter, enhancer) may be used to regulate expression of a gene or a nucleic acid molecule in a way that is different than the gene or a nucleic acid molecule is normally expressed in nature. In another example, a heterologous domain of a polypeptide or nucleic acid sequence (e.g., a DNA binding domain of a polypeptide or nucleic acid encoding a DNA binding domain of a polypeptide) may be disposed relative to other domains or may be a different sequence or from a different source, relative to other domains or portions of a polypeptide or its encoding nucleic acid. In certain embodiments, a heterologous nucleic acid molecule may exist in a native host cell genome, but may have an altered expression level or have a different sequence or both. In other embodiments, heterologous nucleic acid molecules may not be endogenous to a host cell or host genome but instead may have been introduced into a host cell by transformation (e.g., transfection, electroporation), wherein the added molecule may integrate into the host genome or can exist as extra-chromosomal genetic material either transiently (e.g., mRNA) or semi-stably for more than one generation (e.g., episomal viral vector, plasmid or other self-replicating vector).

[0121] Mutation or Mutated: The term “mutated” when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference (e.g., native) nucleic acid sequence. A single alteration may be made at a locus (a point mutation) or multiple nucleotides may be inserted, deleted, or changed at a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. A nucleic acid sequence may be mutated by any method known in the art. In some embodiments a mutation occurs naturally. In some embodiments a desired mutation can be produced by a system described herein.

[0122] Mutation Region: The term “mutation region,” as used herein, refers to a region in a template RNA having one or more sequence difference relative to the corresponding sequence in a target nucleic acid. The sequence difference may comprise, for example, a substitution, insertion, frameshift, or deletion.

[0123] **Nucleic acid molecule:** “Nucleic acid molecule” refers to both RNA and DNA molecules including, without limitation, complementary DNA (“cDNA”), genomic DNA (“gDNA”), and messenger RNA (“mRNA”), and also includes synthetic nucleic acid molecules, such as those that are chemically synthesized or recombinantly produced, such as RNA templates, as described herein. The nucleic acid molecule can be double-stranded or single-stranded, circular, or linear. If single-stranded, the nucleic acid molecule can be the sense strand or the antisense strand. Unless otherwise indicated, and as an example for all sequences described herein under the general format “SEQ ID NO:,” or “nucleic acid comprising SEQ ID NO:1” refers to a nucleic acid, at least a portion which has either (i) the sequence of SEQ ID NO:1, or (ii) a sequence complimentary to SEQ ID NO:1. The choice between the two is dictated by the context in which SEQ ID NO:1 is used. For instance, if the nucleic acid is used as a probe, the choice between the two is dictated by the requirement that the probe be complementary to the desired target. Nucleic acid sequences of the present disclosure may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more naturally occurring nucleotides with an analog, inter-nucleotide modifications such as uncharged linkages (for example, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (for example, phosphorothioates, phosphorodithioates, etc.), pendant moieties, (for example, polypeptides), intercalators (for example, acridine, psoralen, etc.), chelators, alkylators, and modified linkages (for example, alpha anomeric nucleic acids, etc.). Also included are chemically modified bases (see, for example, Table 13), backbones (see, for example 14, **Table 16**), and modified caps. Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of a molecule, e.g., peptide nucleic acids (PNAs). Other modifications can include, for example, analogs in which the ribose ring contains a bridging moiety or other structure such as modifications found in “locked” nucleic acids (LNAs). In various embodiments, the nucleic acids are in operative association with additional genetic elements, such as tissue-specific expression-control sequence(s) (e.g., tissue-specific promoters and tissue-specific microRNA recognition sequences), as well as additional elements, such as inverted repeats (e.g., inverted terminal repeats, such as elements from or derived from viruses, e.g., AAV ITRs) and tandem

repeats, inverted repeats/direct repeats, homology regions (segments with various degrees of homology to a target DNA), untranslated regions (UTRs) (5', 3', or both 5' and 3' UTRs), and various combinations of the foregoing. The nucleic acid elements of the systems provided by the invention can be provided in a variety of topologies, including single-stranded, double-stranded, circular, linear, linear with open ends, linear with closed ends, and particular versions of these, such as doggybone DNA (dbDNA), closed-ended DNA (ceDNA).

[0124] Primer Binding Sequence: The term “primer binding site sequence” or “PBS sequence,” as used herein, refers to a portion of a template RNA capable of binding to a region comprised in a target nucleic acid sequence. In some instances, a PBS sequence is a nucleic acid sequence comprising at least 3, 4, 5, 6, 7, or 8 bases with 100% identity to the region comprised in the target nucleic acid sequence. In some embodiments the primer region comprises at least 5, 6, 7, 8 bases with 100% identity to the region comprised in the target nucleic acid sequence. Without wishing to be bound by theory, in some embodiments when a template RNA comprises a PBS sequence and a heterologous object sequence, the PBS sequence binds to a region comprised in a target nucleic acid sequence, allowing a reverse transcriptase domain to use that region as a primer for reverse transcription, and to use the heterologous object sequence as a template for reverse transcription.

[0125] It is understood that aspects and embodiments described herein as “comprising” include “consisting of” and “consisting essentially of” embodiments.

[0126] Reference to “about” a value or parameter herein includes (and describes) embodiments that are directed to that value or parameter per se. For example, description referring to “about X” includes description of “X”.

[0127] In some embodiments, reference to about a measurable parameter value encompasses variations of $\pm 20\%$, $\pm 10\%$, $\pm 5\%$, $\pm 1\%$, or $\pm 0.1\%$ from the specified value.

25

II. LIPID NANOPARTICLES

[0128] Lipid nanoparticles, in some embodiments, comprise one or more ionic lipids, such as non-cationic lipids (e.g., neutral or anionic, or zwitterionic lipids), also referred to herein as helper lipids; one or more conjugated lipids (such as PEG-conjugated lipids or lipids conjugated to polymers described in Table 5 of WO2019217941; incorporated herein by

30

reference in its entirety); one or more sterols (e.g., cholesterol); and, optionally, one or more targeting molecules (e.g., conjugated receptors, receptor ligands, antibodies); or combinations of the foregoing.

[0129] Lipids that can be used in nanoparticle formations (e.g., lipid nanoparticles) include, for example, those described in Table 4 of WO2019217941, which is incorporated by reference—e.g., a lipid-containing nanoparticle can comprise one or more of the lipids in table 4 of WO2019217941. Lipid nanoparticles can include additional elements, such as polymers, such as the polymers described in table 5 of WO2019217941, incorporated by reference.

[0130] In some embodiments, conjugated lipids, when present, can include one or more of PEG-diacylglycerol (DAG) (such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG)), PEG-dialkylloxypropyl (DAA), PEG-phospholipid, PEG-ceramide (Cer), a pegylated phosphatidylethanolamine (PEG-PE), PEG succinate diacylglycerol (PEGS-DAG) (such as 4-0-(2',3'-di(tetradecanoyloxy)propyl-l-0-(w-methoxy(polyethoxy)ethyl) butanedioate (PEG-S-DMG)), PEG dialkoxypropylcarbam, N-(carbonyl-methoxypoly ethylene glycol 2000)- 1,2-distearoyl-sn-glycero-3-phosphoethanolamine sodium salt, and those described in Table 2 of WO2019051289 (incorporated by reference), and combinations of the foregoing.

[0131] In some embodiments, sterols that can be incorporated into lipid nanoparticles include one or more of cholesterol or cholesterol derivatives, such as those in W02009/127060 or US2010/0130588, which are incorporated by reference. Additional exemplary sterols include phytosterols, including those described in Eygeris et al (2020), dx.doi.org/10.1021/acs.nanolett.0c01386, incorporated herein by reference.

[0132] In some embodiments, the lipid particle comprises an ionizable lipid, a non-cationic lipid, a conjugated lipid that inhibits aggregation of particles, and a sterol. The amounts of these components can be varied independently and to achieve desired properties. For example, in some embodiments, the lipid nanoparticle comprises an ionizable lipid is in an amount from about 20 mol % to about 90 mol % of the total lipids (in other embodiments it may be 20-70% (mol), 30-60% (mol) or 40-50% (mol); about 50 mol % to about 90 mol % of the total lipid present in the lipid nanoparticle), a non-cationic lipid in an amount from about 5 mol % to about 35 mol % of the total lipids, a conjugated lipid in an amount from about 0.5 mol % to about 20 mol % of the total lipids, and a sterol in an amount from about 20 mol % to

about 50 mol % of the total lipids. The ratio of total lipid to nucleic acid (e.g., comprising the therapeutic agent and/or encoding the gene modifying polypeptide, template nucleic acid, or gene modifying system) can be varied as desired. For example, the total lipid to nucleic acid (mass or weight) ratio can be from about 10: 1 to about 30: 1.

5 **[0133]** In some embodiments, the average LNP diameter of the targeted LNP formulation may be between 10s of nm and 100s of nm, e.g., measured by dynamic light scattering (DLS). In some embodiments, the average LNP diameter of the targeted LNP formulation may be from about 40 nm to about 150 nm, such as about 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 10
115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, or 150 nm. In some embodiments, the average LNP diameter of the targeted LNP formulation may be from about 50 nm to about 100 nm, from about 50 nm to about 90 nm, from about 50 nm to about 80 nm, from about 50 nm to about 70 nm, from about 50 nm to about 60 nm, from about 60 nm to about 100 nm, from about 60 nm to about 90 nm, from about 60 nm to about 80 nm, from about 60 nm to about 70 nm, from about 70 nm to about 100 nm, from about 70 nm to about 90 nm, from about 70 nm to about 80 nm, from about 80 nm to about 100 nm, from about 80 nm to about 90 nm, or from about 90 nm to about 100 nm. In some embodiments, the average LNP diameter of the targeted LNP formulation may be from about 70 nm to about 100 nm. In a particular embodiment, the average LNP diameter of the targeted LNP formulation may be about 80 nm.
15
20 In some embodiments, the average LNP diameter of the targeted LNP formulation may be about 100 nm. In some embodiments, the average LNP diameter of the targeted LNP formulation ranges from about 1 mm to about 500 mm, from about 5 mm to about 200 mm, from about 10 mm to about 100 mm, from about 20 mm to about 80 mm, from about 25 mm to about 60 mm, from about 30 mm to about 55 mm, from about 35 mm to about 50 mm, or
25 from about 38 mm to about 42 mm.

[0134] An LNP described herein, e.g., a targeted LNP, may, in some instances, be relatively homogenous. A polydispersity index may be used to indicate the homogeneity of a LNP, e.g., the particle size distribution of the lipid nanoparticles. A small (e.g., less than 0.3) polydispersity index generally indicates a narrow particle size distribution. A LNP may have
30 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 LNP may be from about 0.10 to about 0.20.

[0135] The zeta potential of an LNP may be used to indicate the electrokinetic potential of the composition. In some embodiments, the zeta potential may describe the surface charge of a LNP. Lipid nanoparticles with relatively low charges, positive or negative, are generally desirable, as more highly charged species may interact undesirably with cells, tissues, and other elements in the body. In some embodiments, the zeta potential of a LNP may 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.

[0136] The efficiency of encapsulation of a protein and/or nucleic acid (e.g., an mRNA encoding a polypeptide), describes the amount of protein and/or nucleic acid that is encapsulated or otherwise associated with an LNP after preparation, relative to the initial amount provided. The encapsulation efficiency is desirably high (e.g., close to 100%). The encapsulation efficiency may be measured, for example, by comparing the amount of protein or nucleic acid in a solution containing the lipid nanoparticle before and after breaking up the lipid nanoparticle with one or more organic solvents or detergents. An anion exchange resin may be used to measure the amount of free protein or nucleic acid (e.g., RNA) in a solution. Fluorescence may be used to measure the amount of free protein and/or nucleic acid (e.g., RNA) in a solution. For the lipid nanoparticles described herein, the encapsulation efficiency of a protein and/or nucleic acid may be at least 50%, for example 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the encapsulation efficiency may be at least 80%. In some embodiments, the encapsulation efficiency may be at least 90%. In some embodiments, the encapsulation efficiency may be at least 95%.

[0137] An LNP of the disclosure may optionally comprise one or more coatings. In some embodiments, an LNP may be formulated in a capsule, film, or tablet having a coating.

A capsule, film, or tablet including a composition described herein may have any useful size, tensile strength, hardness or density.

[0138] Additional exemplary lipids, formulations, methods, and characterization of LNPs are taught by WO2020061457, which is incorporated herein by reference in its entirety.

5 [0139] In some embodiments, in vitro or ex vivo cell lipofections are performed using Lipofectamine MessengerMax (Thermo Fisher) or TransIT-mRNA Transfection Reagent (Mirus Bio). In certain embodiments, LNPs are formulated using the GenVoy_ILM ionizable lipid mix (Precision NanoSystems). In certain embodiments, targeted LNPs of the disclosure are formulated using 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (dLin-KC2-DMA)
10 or dilinoleylmethyl-4-dimethylaminobutyrate (dLin-MC3-DMA or MC3), the formulation and in vivo use of which are taught in Jayaraman et al. Angew Chem Int Ed Engl 51(34):8529-8533 (2012), incorporated herein by reference in its entirety.

[0140] Additional specific LNP formulations useful for delivery of nucleic acids are described in US8158601 and US8168775, both incorporated by reference, which include
15 formulations used in patisiran, sold under the name ONPATPRO.

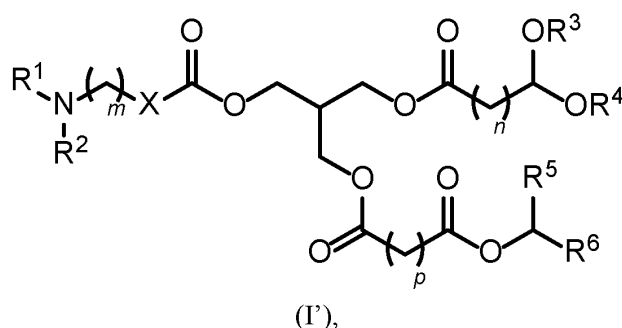
Ionizable Lipids

[0141] The LNPs of the disclosure (e.g., targeted LNPs) comprise one or more ionizable lipids as set forth below. The LNPs are capable of encapsulating a therapeutic agent (e.g., one or more nucleic acids). The therapeutic agent (e.g., one or more nucleic acids) may
20 be adsorbed to the surface of an LNP, e.g., an LNP comprising the ionizable lipid. . It will be understood that any reference to an ionizable lipid in an LNP as set forth herein also includes pharmaceutically acceptable salts of the lipids.

[0142] In some embodiments, the therapeutic agent (e.g., one or more nucleic acids) may be encapsulated in an LNP. In some embodiments, the lipid nanoparticle may comprise a
25 targeting moiety, e.g., coated with a targeting agent. In embodiments, the LNP formulation is biodegradable. In some embodiments, a lipid nanoparticle comprising one or more lipid described herein encapsulates at least 1%, at least 5%, 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%, at least 92%, at least 95%, at least 97%, at least 98% or 100% of an RNA molecule (e.g., an mRNA
30 molecule).

[0143] In some embodiments, the lipid to nucleic acid ratio (mass/mass ratio; w/w ratio) can be in the range of from about 1 : 1 to about 25: 1, from about 10: 1 to about 14: 1, from about 3 :1 to about 15: 1, from about 4: 1 to about 10: 1, from about 5: 1 to about 9: 1, or about 6: 1 to about 9: 1. The amounts of lipids and nucleic acid can be adjusted to provide a
 5 desired N/P ratio, for example, N/P ratio of 3, 4, 5, 6, 7, 8, 9, 10 or higher. Generally, the lipid nanoparticle formulation's overall lipid content can range from about 5 mg/ml to about 30 mg/mL.

[0144] In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid having a structure of formula (I'):



or a pharmaceutically acceptable salt thereof, wherein:

X is -O- or -CH₂-;

m is 0, 1, 2, or 3;

R¹ is C₁₋₄alkyl;

R² is C₁₋₄alkyl;

n is 1, 2, 3, or 4;

R³ is C₄₋₁₀alkyl, wherein one or more alkylene units of the C₄₋₁₀alkyl of R⁵ is optionally replaced with a group selected from -O-, -NH-, -N(C₁₋₄alkyl)-, and -S-;

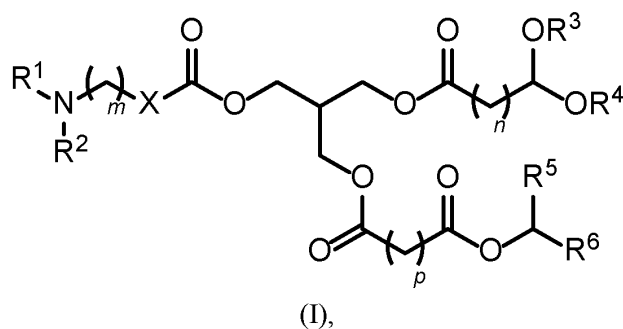
R⁴ is C₄₋₁₀alkyl, wherein one or more alkylene units of the C₄₋₁₀alkyl of R⁵ is optionally replaced with a group selected from -O-, -NH-, -N(C₁₋₄alkyl)-, and -S-;

p is 2, 3, 4, 5, or 6;

R⁵ is C₄₋₁₀alkyl, wherein one or more alkylene units of the C₄₋₁₀alkyl of R⁵ is optionally replaced with a group selected from -O-, -NH-, -N(C₁₋₄alkyl)-, and -S-; and

R^6 is C_{4-10} alkyl, wherein one or more alkylene units of the C_{4-10} alkyl of R^6 is optionally replaced with a group selected from -O-, -NH-, -N(C_{1-4} alkyl)-, and -S-.

- 5 **[0145]** In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid having a structure of formula (I):



- 10 or a pharmaceutically acceptable salt thereof, wherein:

X is -O- or -CH₂-;

m is 0, 1, 2, or 3;

R^1 is C_{1-4} alkyl;

R^2 is C_{1-4} alkyl;

- 15 n is 1, 2, 3, or 4;

R^3 is C_{4-10} alkyl;

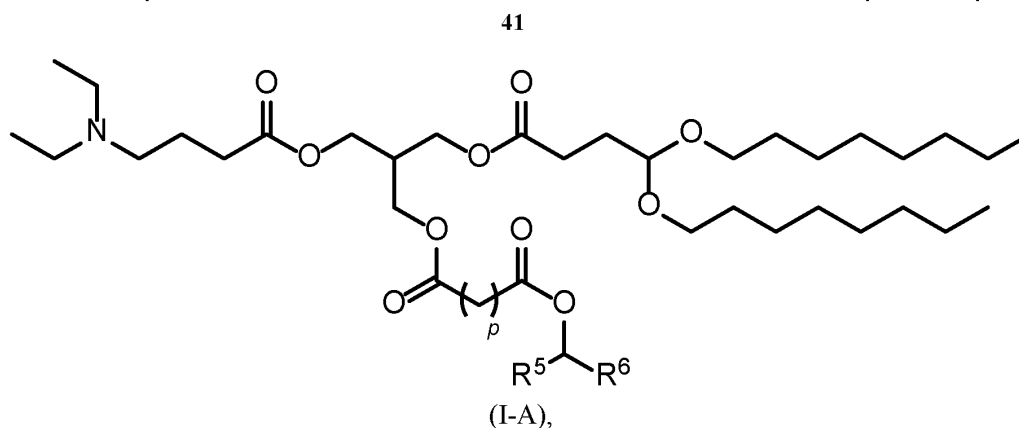
R^4 is C_{4-10} alkyl;

p is 2, 3, 4, 5, or 6;

R^5 is C_{4-10} alkyl; and

- 20 R^6 is C_{4-10} alkyl.

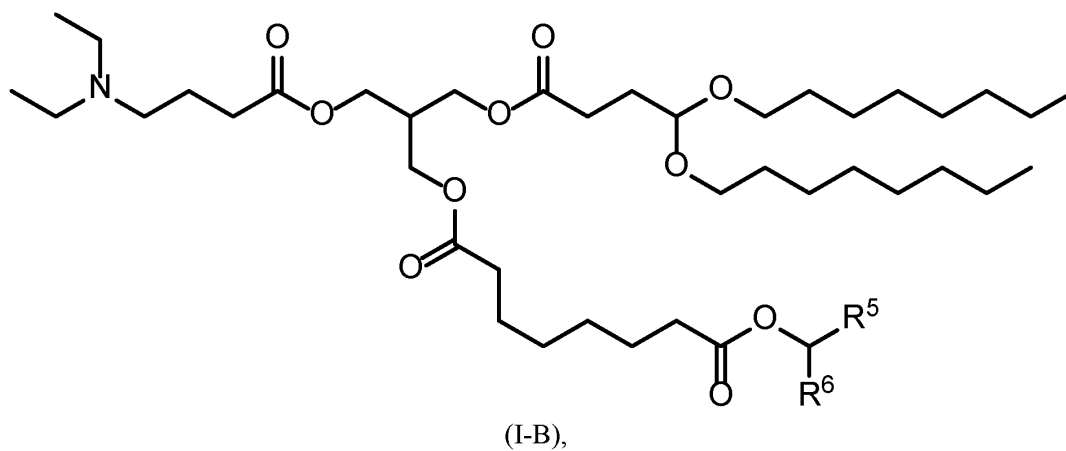
[0146] In some embodiments, compounds of formula (I) are compounds of formula (I-A):



or a pharmaceutically acceptable salt thereof, wherein:

5 p is 2, 3, 4, 5, or 6;
 R⁵ is C₄₋₁₀alkyl; and
 R⁶ is C₄₋₁₀alkyl.

[0147] In some embodiments, compounds of formula (I) are compounds of formula (I-10 B):

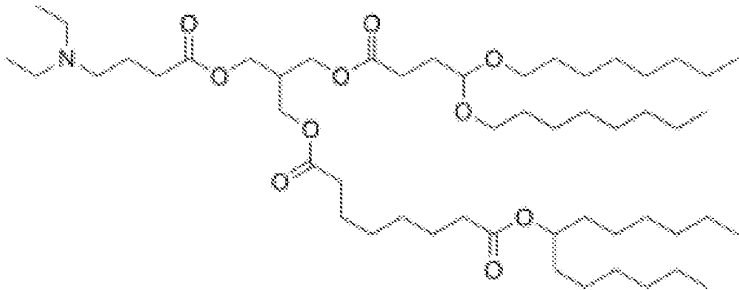
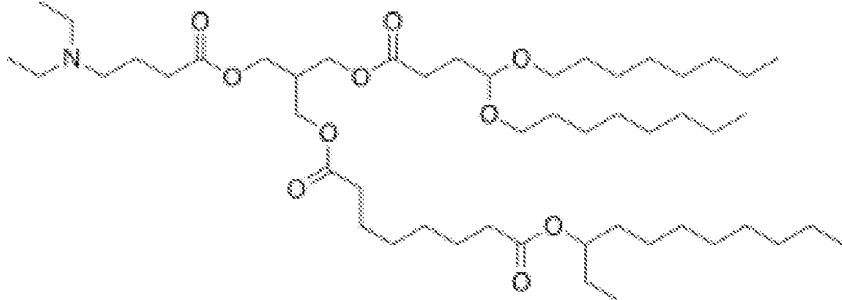
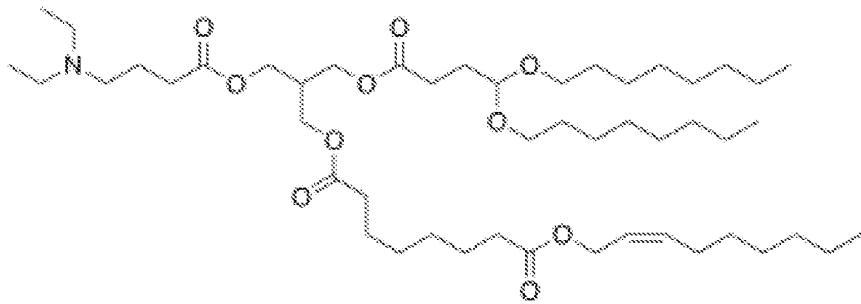


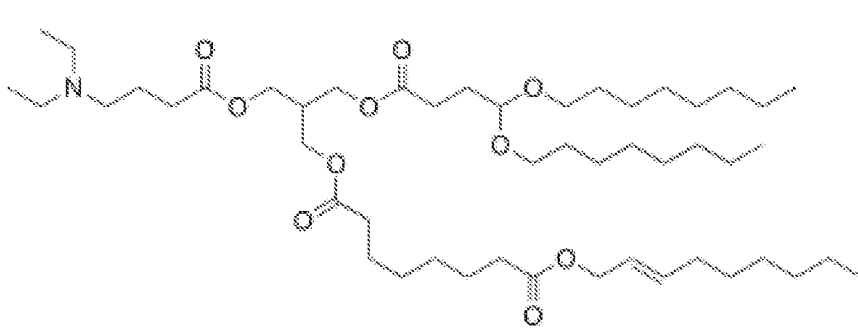
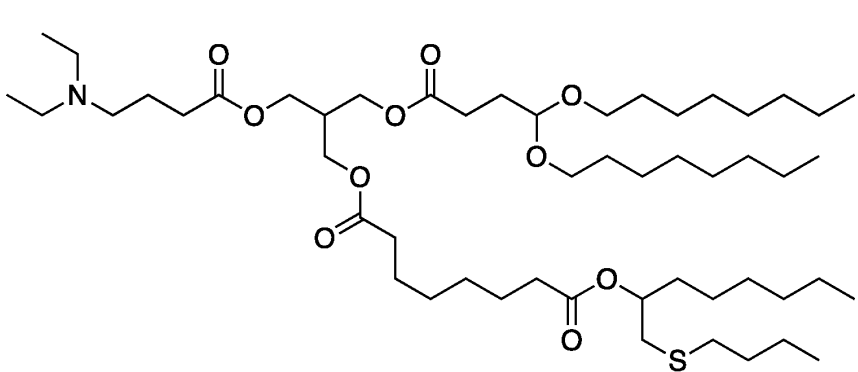
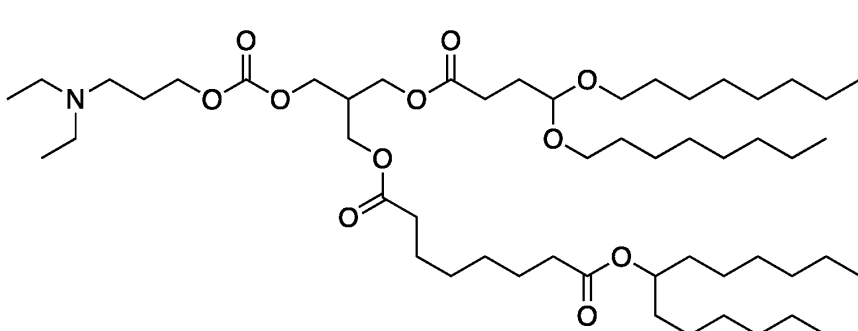
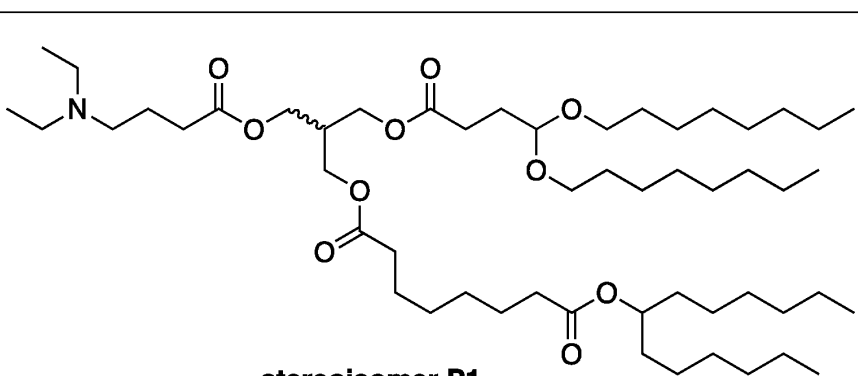
or a pharmaceutically acceptable salt thereof, wherein:

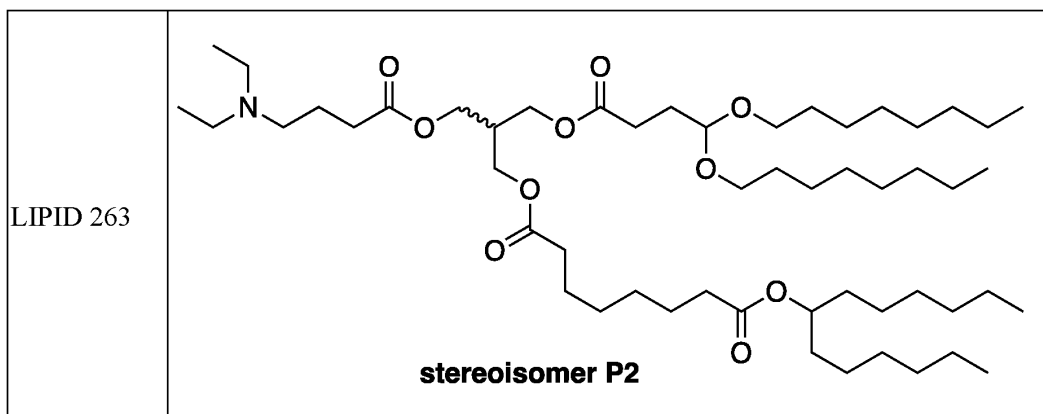
15 R⁵ is C₄₋₆alkyl; and
 R⁶ is C₄₋₆alkyl.

[0148] In some embodiments, the ionizable lipid is a compound selected from the exemplary compounds of **Table 1**.

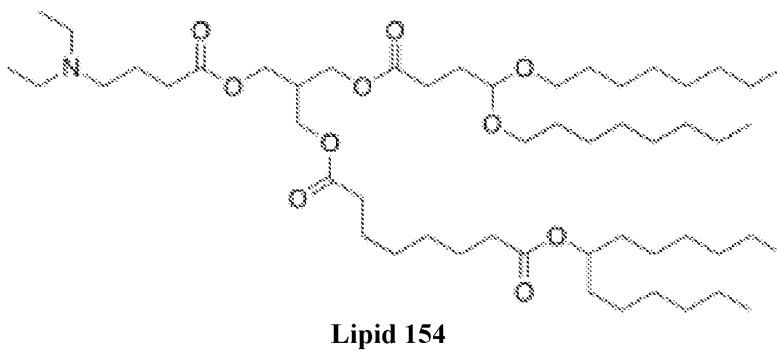
Table 1: Exemplary Ionizable Lipids

Compound	Structure
LIPID 154	
LIPID 155	
LIPID 163	

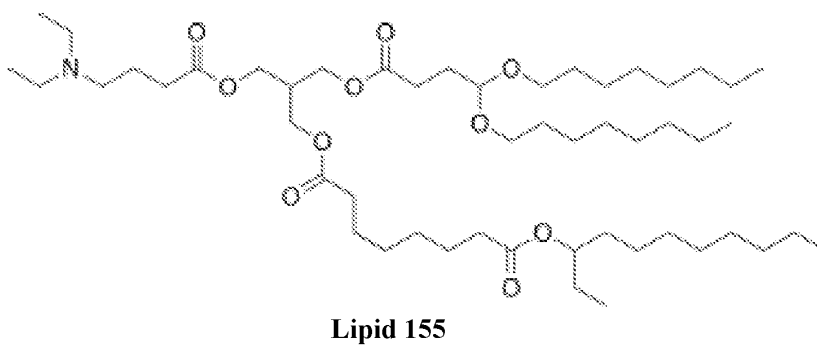
LIPID 173	 Chemical structure of Lipid 173: A complex molecule featuring a central glycerol backbone. The left side is substituted with a diethylammonium group (N(CH₂CH₃)₂) and a long alkyl chain. The right side is substituted with two long alkyl chains. A third long alkyl chain is attached to the central carbon via an ester linkage.
LIPID 209	 Chemical structure of Lipid 209: A complex molecule featuring a central glycerol backbone. The left side is substituted with a diethylammonium group (N(CH₂CH₃)₂) and a long alkyl chain. The right side is substituted with two long alkyl chains. A third long alkyl chain is attached to the central carbon via an ester linkage, and a sulfur atom (S) is present in the side chain.
LIPID 232	 Chemical structure of Lipid 232: A complex molecule featuring a central glycerol backbone. The left side is substituted with a diethylammonium group (N(CH₂CH₃)₂) and a long alkyl chain. The right side is substituted with two long alkyl chains. A third long alkyl chain is attached to the central carbon via an ester linkage.
LIPID 262	 Chemical structure of Lipid 262: A complex molecule featuring a central glycerol backbone. The left side is substituted with a diethylammonium group (N(CH₂CH₃)₂) and a long alkyl chain. The right side is substituted with two long alkyl chains. A third long alkyl chain is attached to the central carbon via an ester linkage. stereoisomer P1



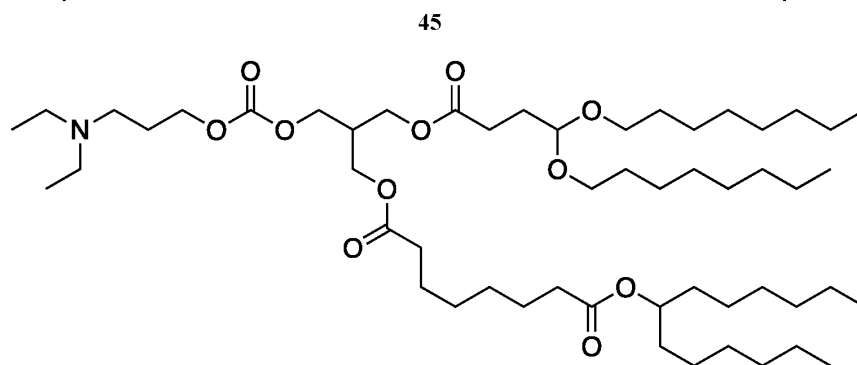
[0149] In some embodiments, the ionizable lipid has one of the structures depicted below:



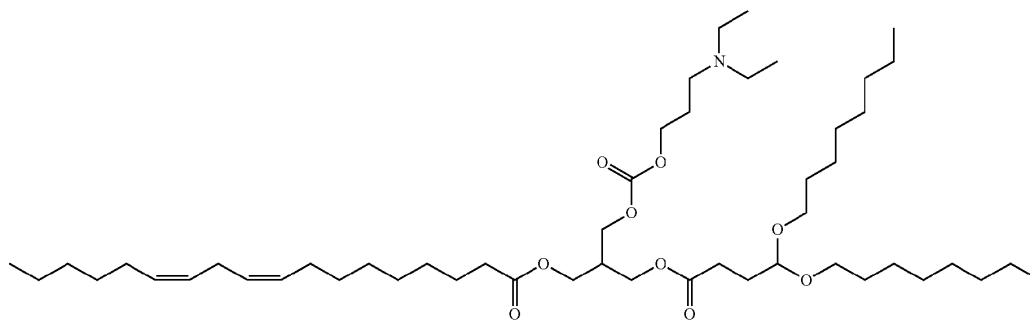
5 [0150] In some embodiments, the ionizable lipid has one of the structures depicted below:



[0151] In some embodiments, the ionizable lipid has one of the structures depicted
10 below:

**Lipid 232**

[0152] In some cases, an LNP containing an ionizable lipid of formula (I) or an ionizable lipid shown in Table 1 exhibits higher levels of transduction in HSCs and/or higher expression of a payload protein in HSCs relative to an LNP that contains V003 as the ionizable lipid. In some cases, an LNP containing an ionizable lipid of formula (I) or an ionizable lipid shown in Table 1 exhibits higher levels of transduction in T cells and/or higher expression of a payload protein in T cells relative to an LNP that contains V003 as the ionizable lipid. In some cases, an LNP containing an ionizable lipid of formula (I) or an ionizable lipid shown in Table 1 exhibits higher levels of transduction in liver cells and/or higher expression of a payload protein in liver cells relative to an LNP that contains V003 as the ionizable lipid. V003, depicted below. V003, described below, is described in U.S. Patent No. 10,059,655.

**Lipid V003**

15

[0153] In some embodiments, Lipid 154 is used as an ionizable lipid to generate LNPs for delivery to cells (e.g., T cells, HSCs, or liver cells). In some embodiments, Lipid 155 is used as an ionizable lipid to generate LNPs for delivery to cells (e.g., T cells, HSCs, or liver cells). In some embodiments, Lipid 163 is used as an ionizable lipid to generate LNPs for delivery to cells (e.g., T cells, HSCs, or liver cells). In some embodiments, Lipid 173 is used

20

as an ionizable lipid to generate LNPs for delivery to cells (e.g., T cells, HSCs, or liver cells). In some embodiments, Lipid 232 is used as an ionizable lipid to generate LNPs for delivery to cells (e.g., T cells, HSCs, or liver cells).

5 [0154] In some embodiments, the mol% of the ionizable lipid in the LNP, e.g., targeted LNP, is from about 25% to about 65%. In some embodiments, the mol% of the ionizable lipid in the LNP, e.g., targeted LNP, is from about 35% to about 60%. In some embodiments, the mol% of the ionizable lipid in the LNP, e.g., targeted LNP, is from about 40% to about 50%. In some embodiments, the mol% of the ionizable lipid in the LNP, e.g., targeted LNP, is from about 45% to about 50%.

10 [0155] The compounds disclosed herein (e.g., lipids in **Table 1**), or their pharmaceutically acceptable salts, may include an asymmetric center and may thus give rise to enantiomers, diastereomers, and other stereoisomeric forms. The present disclosure contemplates various stereoisomers and mixtures thereof and includes “enantiomers,” which refers to two stereoisomers whose molecules are nonsuperimposable mirror images of one
15 another and “diastereomers,” which refers to stereoisomers that have at least two asymmetric atoms, but which are not mirror-images of each other. The present disclosure is meant to include all such possible isomers, as well as their racemic and optically pure forms. Optically active (+) and (-), or (*R*)- and (*S*)- isomers may be prepared using chiral synthons or chiral reagents, or resolved using conventional techniques, for example, chromatography and
20 fractional crystallization. Conventional techniques for the preparation/isolation of individual enantiomers include chiral synthesis from a suitable optically pure precursor or resolution of the racemate using, for example, chiral high pressure liquid chromatography (HPLC). When the compounds described herein contain olefinic double bonds or other centers of geometric asymmetry, and unless specified otherwise, it is intended that the compounds include both E
25 and Z geometric isomers.

Helper Lipids

30 [0156] The LNPs, e.g., targeted LNPs, of the disclosure comprise one or more ionizable lipids. Exemplary helper lipids include, but are not limited to, distearoyl-sn-glycero-phosphoethanolamine, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylglycerol (DOPG), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), dipalmitoylphosphatidylglycerol (DPPG), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyl-oleoylphosphatidylcholine

(POPC), palmitoyl-oleoylphosphatidylethanolamine (POPE), dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidyl-ethanolamine (DSPE), monomethyl-phosphatidylethanolamine (such as 16-O-monomethyl PE), dimethyl-phosphatidylethanolamine (such as 16-O-dimethyl PE), 18-l-trans PE, 1-stearoyl-2-oleoyl-phosphatidylethanolamine (SOPE), hydrogenated soy phosphatidylcholine, egg phosphatidylcholine (EPC), dioleoylphosphatidylserine (DOPS), sphingomyelin (SM), dimyristoyl phosphatidylcholine (DMPC), dimyristoyl phosphatidylglycerol (DMPG), distearoylphosphatidylglycerol (DSPG), dierucoylphosphatidylcholine (DEPC), palmitoyl-oleoylphosphatidylglycerol (POPG), dielaidoyl-phosphatidylethanolamine (DEPE), lecithin, phosphatidylethanolamine, lysolecithin, lysophosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, sphingomyelin, egg sphingomyelin (ESM), cephalin, cardiolipin, phosphatidic acid, cerebroside, dicetylphosphate, lysophosphatidylcholine, dilinoleoylphosphatidylcholine, or mixtures thereof. It is understood that other diacylphosphatidylcholine and diacylphosphatidylethanolamine phospholipids can also be used. The acyl groups in these lipids are preferably acyl groups derived from fatty acids having C10-C24 carbon chains, e.g., lauroyl, myristoyl, palmitoyl, stearoyl, or oleoyl. Additional exemplary lipids, in certain embodiments, include, without limitation, those described in Kim et al. (2020) [dx.doi.org/10.1021/acs.nanolett.0c01386](https://doi.org/10.1021/acs.nanolett.0c01386), incorporated herein by reference.

[0157] Other examples of non-cationic lipids suitable for use in the lipid nanoparticles include, without limitation, nonphosphorous lipids such as, e.g., stearylamine, dodecylamine, hexadecylamine, acetyl palmitate, glycerol ricinoleate, hexadecyl stearate, isopropyl myristate, amphoteric acrylic polymers, triethanolamine-lauryl sulfate, alkyl-aryl sulfate, polyethoxylated fatty acid amides, dioctadecyl dimethyl ammonium bromide, ceramide, sphingomyelin, and the like. Other non-cationic lipids are described in WO2017/099823 or US patent publication US2018/0028664, the contents of which are incorporated herein by reference in their entirety. In some embodiments, the non-cationic lipid is oleic acid or a compound of Formula I, II, or IV of US2018/0028664, incorporated herein by reference in its entirety.

[0158] In some embodiments, the helper lipid is a sphingolipid. In some such embodiments, the non-pegylated lipid is a sphingomyelin. In some embodiments, the

sphingomyelin has a head group selected from, phosphocholine, phosphoethanolamine or ceramide. In some embodiments, the sphingomyelin is egg sphingomyelin.

[0159] In some embodiments, the helper lipid comprises 5-40% (mol), 8%-30%, 10%-28%, 20%-36%, 22%-32%, or 10-15% (mol) of the total lipid present in the lipid nanoparticle.

5 In embodiments, the molar ratio of ionizable lipid to the neutral lipid ranges from about 2:1 to about 8:1 (e.g., about 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, or 8:1).

[0160] In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP, e.g., targeted LNP, is from about 18% to about 32%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP, e.g.,
10 targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP, e.g., targeted LNP, is from about 22% to about 32%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the LNP, e.g., targeted LNP, is from about 22% to about 28%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin) in the targeted LNP is from about 21% to about 23%. In some embodiments, the mol% of the helper lipid (e.g., DSPC or sphingomyelin)
15 in the targeted LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, about 30%, about 31%, or about 32%. As set forth in the examples below, in vivo delivery of certain payloads following administration of the disclosed LNPs (e.g., targeted LNPs) with these percentages of helper lipids provides
20 enhanced transduction and expression of the payloads relative to targeted LNPs with smaller or larger quantities of helper lipid.

[0161] In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) in the LNP, e.g., targeted LNP, is from about 1:1 to about 7:1. In some embodiments, the molar ratio between the ionizable lipid and
25 the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 4:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 3:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC) is from about 1:1 to about 2.5:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about
30 2:1. In some embodiments, the molar ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC) is from about 1.5:1 to about 2.5:1. In some embodiments, the molar

ratio between the ionizable lipid and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 2:1 to about 2.5:1.

[0162] In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid in **Table 1**, and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid **Table 1** and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 22% to about 28%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0163] In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid of Formula I and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises an ionizable lipid of Formula I and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP is from about 22% to about 28%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0164] In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 154 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 154 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 22% to about 28%. In some embodiments, the mol% of DSPC or sphingomyelin in the targeted LNP is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0165] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 155 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 155 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0166] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 163 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 163 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0167] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 173 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 173 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0168] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 209 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 209 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In

some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

5 **[0169]** In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 232 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 232 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In
10 some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0170] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 262 and
15 DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 262 and a sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP,
20 is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

[0171] In some embodiments, the LNP, e.g., targeted LNP comprises Lipid 263 and DSPC. In some embodiments, the LNP, e.g., targeted LNP, comprises Lipid 263 and a
25 sphingomyelin. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 25%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of the DSPC or sphingomyelin in the LNP, e.g., targeted LNP, is from about 20% to about 30%. In some embodiments, the mol% of DSPC or sphingomyelin
30 in the LNP, e.g., targeted LNP, is about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 25%, about 27%, about 28%, about 29%, or about 30%.

Sterols

[0172] In some embodiments, the LNPs, e.g., targeted LNPs, of the disclosure can comprise a component, such as a sterol, to provide membrane integrity. One exemplary sterol that can be used in the lipid nanoparticle is cholesterol and derivatives thereof. Non-limiting
5 examples of cholesterol derivatives include polar analogues such as 5 α -cholestanol, 5 β -coprostanol, choiesterolyl-(2'-hydroxy)-ethyl ether, choiesterolyl-(4'-hydroxy)-butyl ether, and 6-ketocholestanol; non-polar analogues such as 5 α -cholestane, cholestenone, 5 α -cholestanone, 5 β -cholestanone, and choiesterolyl decanoate; and mixtures thereof. In some embodiments, the cholesterol derivative is a polar analogue, e.g., choiesterolyl-(4'-hydroxy)-butyl ether.
10 Exemplary cholesterol derivatives are described in PCT publication W02009/127060 and US patent publication US2010/0130588, each of which is incorporated herein by reference in its entirety.

[0173] In some embodiments, the component providing membrane integrity, such as a sterol, can comprise 0-50% (mol) (e.g., 0-10%, 10-20%, 20-30%, 30-40%, or 40-50%) of the
15 total lipid present in the lipid nanoparticle. In some embodiments, such a component is 20-50% (mol) 30-40% (mol) of the total lipid content of the lipid nanoparticle.

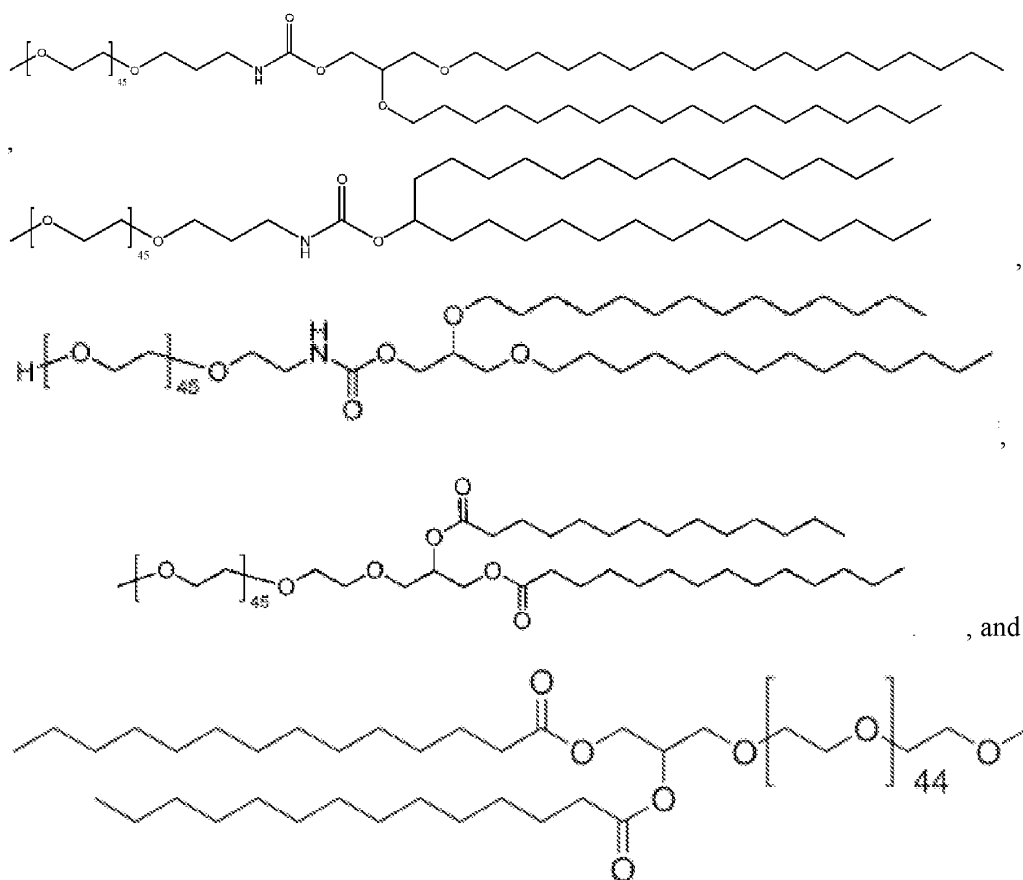
[0174] In some embodiments, the molar ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 6:1 to about 0.5:1. In some embodiments, the ratio between the cholesterol molecule and the non-pegylated
20 helper lipid (e.g., DSPC or sphingomyelin) is from about 3:1 to about 0.5:1. In some embodiments, the ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 2:1 to about 0.5:1. In some embodiments, the ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1.5:1 to about 0.5:1. In some embodiments, the ratio between
25 the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:1 to about 0.5:1. In some embodiments, the ratio between the cholesterol molecule and the non-pegylated helper lipid (e.g., DSPC or sphingomyelin) is from about 1:2 to about 0.8:1.

Pegylated Lipids

30 [0175] In some embodiments, the LNPs, e.g., targeted LNPs, of the disclosure can comprise a polyethylene glycol (PEG) or a conjugated lipid molecule. Generally, these are used to inhibit aggregation of lipid nanoparticles and/or provide steric stabilization. Exemplary

conjugated lipids include, but are not limited to, PEG-lipid conjugates, polyoxazoline (POZ)-lipid conjugates, polyamide-lipid conjugates (such as ATTA-lipid conjugates), cationic-polymer lipid (CPL) conjugates, and mixtures thereof. In some embodiments, the conjugated lipid molecule is a PEG-lipid conjugate, for example, a (methoxy polyethylene glycol)-conjugated lipid.

[0176] Exemplary PEG-lipid conjugates include, but are not limited to, PEG-diacylglycerol (DAG) (such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG)), PEG-dialkylxypropyl (DAA), PEG-phospholipid, PEG-ceramide (Cer), a pegylated phosphatidylethanolamine (PEG-PE), 1,2-dimyristoyl-sn-glycerol, methoxypolyethylene glycol (DMG-PEG-2K), PEG succinate diacylglycerol (PEGS-DAG) (such as 4-0-(2',3'-di(tetradecanoyloxy)propyl-1-0-(w-methoxy(polyethoxy)ethyl) butanedioate (PEG-S-DMG)), PEG dialkoxypropylcarbam, N-(carbonyl-methoxypolyethylene glycol 2000)-1,2-distearoyl-sn-glycero-3-phosphoethanolamine sodium salt, or a mixture thereof. Additional exemplary PEG-lipid conjugates are described, for example, in US5,885,613, US6,287,591, US2003/0077829, US2003/0077829, US2005/0175682, US2008/0020058, US2011/0117125, US2010/0130588, US2016/0376224, US2017/0119904, and US/099823, the contents of all of which are incorporated herein by reference in their entirety. In some embodiments, a PEG-lipid is a compound of Formula III, III-a-I, III-a-2, III-b-1, III-b-2, or V of US2018/0028664, the content of which is incorporated herein by reference in its entirety. In some embodiments, a PEG-lipid is of Formula II of US20150376115 or US2016/0376224, the content of both of which is incorporated herein by reference in its entirety. In some embodiments, the PEG-DAA conjugate can be, for example, PEG-dilauryloxypropyl, PEG-dimyristyloxypropyl, PEG-dipalmityloxypropyl, or PEG-distearoyloxypropyl. The PEG-lipid can be one or more of PEG-DMG, PEG-dilaurylglycerol, PEG-dipalmitoylglycerol, PEG-disteryl-glycerol, PEG-dilaurylglycamide, PEG-dimyristylglycamide, PEG-dipalmitoylglycamide, PEG-disteryl-glycamide, PEG-cholesterol (1-[8'-(Cholest-5-en-3[beta]-oxy)carboxamido-3',6'-dioxaoctanyl] carbamoyl-[omega]-methyl-poly(ethylene glycol), PEG-DMB (3,4-Ditetradecoxybenzyl-[omega]-methyl-poly(ethylene glycol) ether), and 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000]. In some embodiments, the PEG-lipid comprises PEG-DMG, 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000]. In some embodiments, the PEG-lipid comprises a structure selected from:



5

[0177] In some embodiments, lipids conjugated with a molecule other than a PEG can also be used in place of PEG-lipid. For example, polyoxazoline (POZ)-lipid conjugates, polyamide-lipid conjugates (such as ATTA-lipid conjugates), and cationic-polymer lipid (GPL) conjugates can be used in place of or in addition to the PEG-lipid.

10 **[0178]** Exemplary conjugated lipids, i.e., PEG-lipids, (POZ)-lipid conjugates, ATTA-lipid conjugates and cationic polymer-lipids are described in the PCT and LIS patent applications listed in Table 2 of WO2019051289A9 and in WO2020106946A1, the contents of all of which are incorporated herein by reference in their entirety.

[0179] In some embodiments, the pegylated lipid has at least one C16 (palmitoyl) PEG lipid anchor. In some embodiments, the pegylated lipid has two C16 PEG lipid anchors (i.e., dialkyl chains of 16 carbons long). In some embodiments, the pegylated lipid is 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (DPPE-PEG2000.). In some embodiments, the pegylated lipid is 1,2-Dipalmitoyl-rac-glycero-

3-methylpolyoxyethylene (DPG-PEG2000). In some embodiments, the pegylated lipid is C16 PEG ceramide. In some embodiments, the targeted LNPs comprising the C16 pegylated lipids show reduced liver uptake than otherwise identical LNPs comprising C14 pegylated lipids.

5 **[0180]** In some embodiments, the pegylated lipid has at least one C18 PEG lipid anchor. In some embodiments, the pegylated lipid has two C18 PEG lipid anchors (i.e., dialkyl chains of 18 carbons long). In some embodiments, the C18 pegylated lipid is 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (DSPE-PEG2000). In some embodiments, the C18 pegylated lipid is distearoyl-rac-glycerol-PEG2000 (DSG-PEG2000).

10 **[0181]** In some embodiments, the PEG or the conjugated lipid can comprise 0-20% (mol) of the total lipid present in the lipid nanoparticle. In some embodiments, PEG or the conjugated lipid content is 0.5- 10% or 2-5% (mol) of the total lipid present in the lipid nanoparticle.

Additional components

15 **[0182]** In some embodiments, one or more additional compounds can also be included in the LNPs, e.g., targeted LNPs, of the disclosure. Those compounds can be administered separately or the additional compounds can be included in the lipid nanoparticles of the invention. In other words, the lipid nanoparticles can contain other compounds, such as a payload, e.g., a therapeutic agent, as described herein. In some embodiments, the lipid
20 nanoparticles can contain one or more nucleic acids. In some embodiments, the lipid nanoparticles can contain at least a first nucleic acid and a second nucleic acid, wherein the second nucleic acid is different than the first nucleic acid.. Without limitations, other additional compounds can be selected from the group consisting of small or large organic or inorganic molecules, monosaccharides, disaccharides, trisaccharides, oligosaccharides, polysaccharides,
25 peptides, proteins, peptide analogs and derivatives thereof, peptidomimetics, nucleic acids, nucleic acid analogs and derivatives, an extract made from biological materials, or any combinations thereof.

30 **[0183]** In some embodiments, a lipid nanoparticle (or a formulation comprising lipid nanoparticles) lacks reactive impurities (e.g., aldehydes or ketones), or comprises less than a preselected level of reactive impurities (e.g., aldehydes or ketones). While not wishing to be bound by theory, in some embodiments, a lipid reagent is used to make a lipid nanoparticle

formulation, and the lipid reagent may comprise a contaminating reactive impurity (e.g., an aldehyde or ketone). A lipid reagent may be selected for manufacturing based on having less than a preselected level of reactive impurities (e.g., aldehydes or ketones). Without wishing to be bound by theory, in some embodiments, aldehydes can cause modification and damage of RNA, e.g., cross-linking between bases and/or covalently conjugating lipid to RNA (e.g., forming lipid-RNA adducts). This may, in some instances, lead to failure of a reverse transcriptase reaction and/or incorporation of inappropriate bases, e.g., at the site(s) of lesion(s), e.g., a mutation in a newly synthesized target DNA.

[0184] In some embodiments, a lipid nanoparticle formulation is produced using a lipid reagent comprising less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content. In some embodiments, a lipid nanoparticle formulation is produced using a lipid reagent comprising less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species. In some embodiments, a lipid nanoparticle formulation is produced using a lipid reagent comprising: (i) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content; and (ii) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species. In some embodiments, the lipid nanoparticle formulation is produced using a plurality of lipid reagents, and each lipid reagent of the plurality independently meets one or more criterion described in this paragraph. In some embodiments, each lipid reagent of the plurality meets the same criterion, e.g., a criterion of this paragraph.

[0185] In some embodiments, the lipid nanoparticle formulation comprises less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content. In some embodiments, the lipid nanoparticle formulation comprises less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species. In some embodiments, the lipid nanoparticle formulation comprises: (i) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content; and (ii) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species.

[0186] In some embodiments, one or more, or optionally all, of the lipid reagents used for a lipid nanoparticle as described herein or a formulation thereof comprise less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content. In some embodiments, one or more, or optionally all, of the lipid reagents used for a lipid nanoparticle as described herein or a formulation thereof comprise less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species. In some embodiments, one or more, or optionally all, of the lipid reagents used for a lipid nanoparticle as described herein or a formulation thereof comprise: (i) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% total reactive impurity (e.g., aldehyde) content; and (ii) less than 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, or 0.1% of any single reactive impurity (e.g., aldehyde) species.

[0187] In some embodiments, total aldehyde content and/or quantity of any single reactive impurity (e.g., aldehyde) species is determined by liquid chromatography (LC), e.g., coupled with tandem mass spectrometry (MS/MS), e.g., according to the method described in Example 34. In some embodiments, reactive impurity (e.g., aldehyde) content and/or quantity of reactive impurity (e.g., aldehyde) species is determined by detecting one or more chemical modifications of a nucleic acid molecule (e.g., an RNA molecule, e.g., as described herein) associated with the presence of reactive impurities (e.g., aldehydes), e.g., in the lipid reagents. In some embodiments, reactive impurity (e.g., aldehyde) content and/or quantity of reactive impurity (e.g., aldehyde) species is determined by detecting one or more chemical modifications of a nucleotide or nucleoside (e.g., a ribonucleotide or ribonucleoside, e.g., comprised in or isolated from a template nucleic acid, e.g., as described herein) associated with the presence of reactive impurities (e.g., aldehydes), e.g., in the lipid reagents. In some embodiments, chemical modifications of a nucleic acid molecule, nucleotide, or nucleoside are detected by determining the presence of one or more modified nucleotides or nucleosides, e.g., using LC-MS/MS analysis.

III. CONJUGATES

[0188] In one aspect, the disclosure provides an LNP (conjugate) comprising an ionizable lipid as described herein (e.g., in Table 1), wherein the LNP can deliver a payload, such as a therapeutic agent (e.g., a gene modifying system) to a cell. In some embodiments,

the cell is an HSC (e.g., a LT-HSC) or HPC. In other embodiments, the cell is a T cell. In other embodiments, the cell is a liver cell.

[0189] In another aspect, the LNP (conjugate) comprises a targeting moiety that binds to a protein (e.g., a protein receptor) on a cell such as an HSC (e.g., a LT-HSC) HPC, T cell, or liver cell, as described herein. In some embodiments, an LNP (conjugate) comprises both an ionizable lipid and a targeting moiety, as described herein.

[0190] In one aspect, the disclosure provides targeted LNPs (conjugates) comprising a targeting moiety and a lipid nanoparticle (LNP) encapsulating a payload (e.g., a therapeutic agent, as described herein, such as a gene modifying polypeptide or a gene modifying system), wherein the targeting moiety binds to a protein (e.g., protein receptor) on an HSC (e.g., LT-HSC), HPC, T cell, or liver cell. In some embodiments, the targeting moiety is an antibody or antigen binding fragment thereof. In some instances, the targeting moiety is an antibody, a Fab fragment, a scFv, a DARPIN, a VHH domain antibody, a FN3 domain, a nanobody, a single domain antibody or a Centyrin. In other embodiments, the targeting moiety is a folate moiety, an antibiotic mimetic, a polynucleotide (such as a DNA or RNA aptamer), a carbohydrate, a vitamin or a N-Acetylgalactosamine (GalNac). In some embodiments, the payload (e.g., a therapeutic agent, as described herein, such as a gene modifying polypeptide or a gene modifying system) is capable of modifying one or more genes of the target HSC.

[0191] In some embodiments, the targeting moiety component of a conjugate of the disclosure can target a cell surface antigen or receptor. In some embodiments, the targeting moiety component of a conjugate of the disclosure targets T cell receptors including, but not limited to, CD2, CD3, CD4, CD5, CD6 or CD7. In other embodiments, the targeting moiety component of a conjugate of the disclosure targets hematopoietic stem cells (HSCs). In some embodiments, the targeting moiety component of a conjugate of the disclosure targets HSC receptors including, but not limited to, CD90 or CD117. In other embodiments, the targeting moiety component of a conjugate binds to CD8, TCR alpha, TCR beta, CD10, CD33, CD34, CD68, CD19, CCR7, CD62L, CD25, CXCR3, CCR2, CCR3, CCR4, CCR5, CCR, or CCR7.

[0192] In certain embodiments, the targeting moiety binds to CD4⁺ or CD8⁺ T cell. In other embodiments, the targeting moiety binds to a natural killer (NK) cell. In other embodiments, the targeting moiety binds to a hematopoietic stem cell. In other embodiments,

the targeting moiety binds to a lymphoid progenitor cell. In other embodiments, the targeting moiety binds to a myeloid cell. In other embodiments, the targeting moiety binds to a macrophage.

[0193] In some embodiments, the LNP comprises a targeting moiety that targets
5 a receptor on the surface of the HSC selected from CD33, CD34, CD38, CD43, CD46, CD49d, CD50, CD59, CD71, CD105, CD117, CD123, CD164, CD184, and CD338. In some embodiments, the targeting moiety on the LNP binds to CD33. In some embodiments, the targeting moiety on the LNP binds to CD34. In some embodiments, the targeting moiety on the LNP binds to CD38. In some embodiments, the targeting moiety on the LNP binds to CD43.
10 CD43. In some embodiments, the targeting moiety on the LNP binds to CD46. In some embodiments, the targeting moiety on the LNP binds to CD49d. In some embodiments, the targeting moiety on the LNP binds to CD50. In some embodiments, the targeting moiety on the LNP binds to CD59. In some embodiments, the targeting moiety on the LNP binds to CD71. In some embodiments, the targeting moiety on the LNP binds to CD105. In some embodiments, the targeting moiety on the LNP binds to CD117. In some embodiments, the targeting moiety on the LNP binds to CD123. In some embodiments, the targeting moiety on the LNP binds to CD184. In some embodiments, the targeting moiety on the LNP binds to CD164. In some
15 embodiments, the targeting moiety on the LNP binds to CD338.

[0194] In some embodiments, a lipid nanoparticle (LNP) for delivery of a therapeutic
20 agent to HSCs, as described herein, comprises:

an ionizable lipid and a helper lipid, wherein the ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154, Lipid 155, Lipid 209, Lipid 232, Lipid 262, or Lipid 263);

a plurality of targeting moieties conjugated to the LNP, wherein the plurality of targeting moieties bind to at least one of CD33, CD34, CD38, CD43, CD59, CD105, CD123,
25 CD164, CD338, CD71, CD117, CD50, CD49d, CD46, or CD184 (CXCR4); and
a therapeutic agent encapsulated within the LNP.

[0195] In some embodiments, the dual binders can be added to the surface of the LNP by methods described herein. For instance, in some embodiments, each of the targeting
30 moieties is conjugated to the lipid nanoparticle through a linker, wherein each linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such embodiments, the targeting

moieties are antibodies or antigen binding fragments thereof. In other such embodiments, the targeting moieties are ScFvs. In some embodiments, the targeting moieties are Fab fragments.

[0196] In some embodiments, the plurality of targeting moieties bind to two or more of CD33, CD34, CD38, CD43, CD59, CD105, CD123, CD164, CD338, CD71, CD117, CD50, CD49d, CD46, or CD184 (CXCR4). In some embodiments, the plurality of targeting moieties bind to CD117 and CD34. In some embodiments, the plurality of targeting moieties bind to CD117 and CD71. In some embodiments, the plurality of targeting moieties bind to CD34 and CD71. In some embodiments, the plurality of targeting moieties bind to CD117 and CD50. In some embodiments, the plurality of targeting moieties bind to CD34 and CD50.

[0197] In some embodiments, two different targeting moieties can bind to the same receptor on the surface of the HSC. For instance, in some embodiments, both targeting moieties can bind to CD117. In other embodiments, both targeting moieties can bind to CD34.

[0198] In some embodiments, a lipid nanoparticle (LNP) for delivery of a therapeutic agent to HSCs, as described herein, comprises:

a plurality of targeting moieties conjugated to the LNP, wherein the plurality of targeting moieties bind to CD33, CD34, CD38, CD43, CD59, CD105, CD123, CD164, CD338, CD71, CD117, CD50, CD49d, CD46, or CD184 (CXCR4); and
a therapeutic agent encapsulated within the LNP.

[0199] In some aspects, the disclosure provides a targeted lipid nanoparticle (LNP) for in vivo delivery of a therapeutic agent to HSCs (e.g., LT-HSCs), wherein the LNP comprises:

- i. An ionizable lipid of formula (I), or a salt thereof (e.g., Lipid 154);
- ii. a helper lipid;
- iii. two different targeting moieties conjugated to the LNP, wherein each of the two different targeting moieties binds to a target on the surface of the HSCs (e.g., LT-HSCs); and
- iv. one or more nucleic acids (e.g., two or more nucleic acids) encoding components of a system for modifying or altering genomic DNA (e.g., a gene modifying system), wherein the one or more nucleic acids (e.g., two or more nucleic acids) are RNA molecules.

[0200] Methods of Making Targeted LNPs (Conjugates)

[0201] Different approaches can be used to introduce a targeting moiety onto the surface of an LNP. For example, one approach relies on functionalizing a preformed LNP with a targeting moiety. The LNP generally includes a lipid that has polyethylene glycol (PEG) spacer functionalized with a reactive moiety such as a thiol, amine, maleimide or carboxylic acid group. The functionalized lipid of the LNP reacts with a complementary group that is covalently bonded to a targeting moiety, hence generating a conjugate of the LNP and the targeting moiety.

[0202] In some embodiments, the targeting moiety is conjugated to the lipid nanoparticle through a linker, and wherein the linker comprises a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some such embodiments, the targeting moiety is an antibody or antigen binding fragment thereof. In other such embodiments, the targeting moiety is a ScFv. In some embodiments, the targeting moiety is a Fab fragment.

[0203] In one embodiment, the Click product can be formed using a copper-catalyzed Click reaction. One such copper-catalyzed Click reaction is a Huisgen 1,3-dipolar cycloaddition (CuAAC) between an azide and an alkyne. In some embodiments, the first or second Click handle comprises a cyclic derivative of the alkynyl group. In some embodiments, the cyclic derivative of the alkynyl group is selected from dibenzocyclooctyne, cyclooctyne, and difluorinated cyclooctyne. In some embodiments, the click chemistry involves strain promoted cycloaddition of azides. In some embodiments, the click chemistry is based upon reaction of strained alkenes.

[0204] In another embodiment, the Click product can be formed using copper-free Click chemistry. For example, the Click product can be formed between an azide and dibenzocyclooctene (DBCO). Alternatively, the Click product can be formed using a Staudinger reaction between an azide and a phosphine, hence producing an aza-ylide.

[0205] In some embodiments, the Click product can be formed from an inverse electron demand Diels-Alder reaction between a trans-cyclooctene (TCO) moiety on the first or second Click handle and a tetrazine ring on the first or second Click handle. In some embodiments, the first Click handle comprises a tetrazine (Tz) ring and the second Click handle comprises a TCO moiety. In some embodiments, the tetrazine ring is unsubstituted. In some

such embodiments, the tetrazine rung is methyltetrazine. In some embodiments, the tetrazine ring is a 6-methyl substituted tetrazine.

[0206] In another embodiment, the targeting moiety (e.g., antibody, Fab fragment or ScFv) is first selectively modified with an enzyme recognition sequence. An enzyme recognizing the enzyme recognition sequence can site-specifically introduce the first Click handle onto the targeting moiety through covalent attachment. The first Click handle can next react with the second Click handle on the LNP to produce the targeted LNP. Hence, in one embodiment, an antibody, Fab fragment or single chain variable fragment (scFv) that is covalently linked to a first Click handle through a linker comprising an enzyme recognition sequence is reacted with an LNP comprising a second Click handle, thereby forming a Click reaction product that conjugates the antibody, Fab fragment or ScFv to the LNP. In some embodiments, the antibody Fab fragment or ScFv is directly bonded to the enzyme recognition sequence. In some embodiments, the targeting moiety (e.g., antibody Fab fragment or scFv) is bonded to the enzyme recognition sequence via one or more amino acid residues. Particular amino acid residues added that can be covalently attached to the C-terminus of the targeting moiety (e.g., antibody, Fab fragment or ScFv) include, but are not limited to (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10.

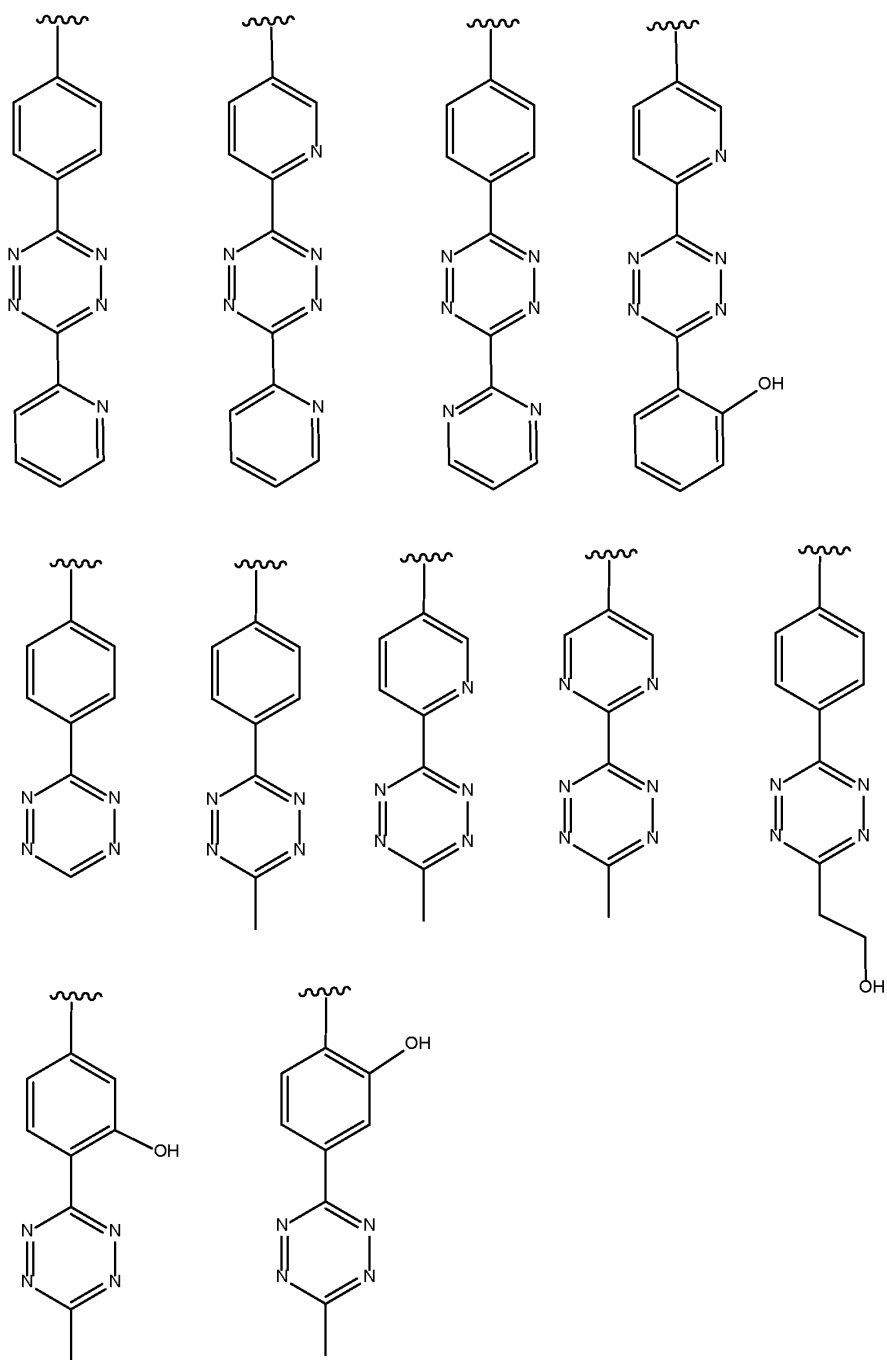
[0207] In some embodiments, the enzyme recognition sequence is a sortase recognition motif or an LplA acceptor peptide. In some embodiments where the LNP is conjugated to an antibody, the C-terminus of one or more of the heavy or light chains of the antibody is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide) either directly or through a linker comprising one or more amino acid residues, as set forth herein. In some embodiments where the LNP is conjugated to a Fab fragment, the C-terminus of the heavy or light chain of the Fab fragment is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide). In some embodiments where the LNP is conjugated to a scFv, the C-terminus of the scFv is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide). In some of the foregoing embodiments, the first Click handle comprises a tetrazine ring or TCO moiety and the second Click handle comprises a tetrazine ring or TCO moiety. In some such embodiments, the first Click handle comprises a tetrazine ring and the second Click handle comprises a TCO moiety. In some embodiments, the tetrazine

ring is a methyltetrazine. In some embodiments, the conjugation efficiency achieved by the disclosed method is greater than 50%, greater than 60%, greater than 70%, greater than 80%, or greater than 90%. In some embodiments, the conjugation efficiency achieved by the disclosed method is from about 60% to about 95%. In some embodiments, the conjugation efficiency achieved by the disclosed methods are from about 70% to about 85%.

[0208] In some embodiments, the linker further comprises a spacer between the targeting moiety and the Click product. The spacer can include additional functional groups that indirectly link the targeting moiety to the Tz group. The spacer may also include additional amino acid residues that indirectly links the targeting moiety to the Tz group.

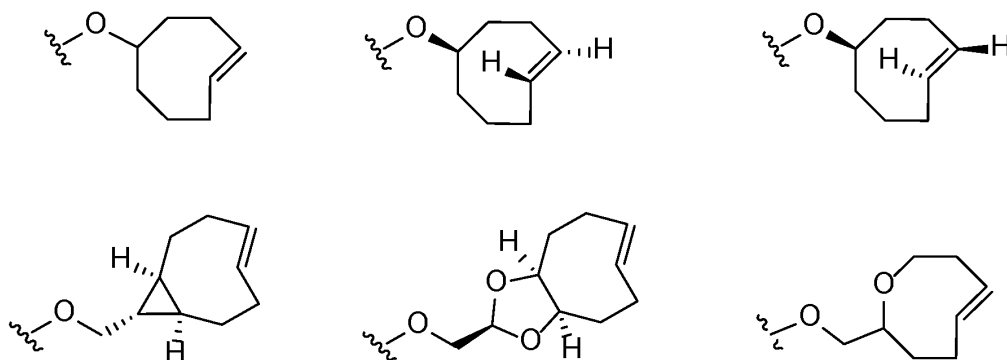
[0209] In some embodiments, the first or second Click handle is a tetrazine derivative having one of the following structures:

64



wherein  represents the point of attachment to the linker of the conjugate or to the LNP.

- 5 **[0210]** In some embodiments, the first or second Click handle is a TCO derivative having one of the following structures:



[0211] In some embodiments, the spacer that links the targeting moiety to the Tz ring is an enzyme recognition sequence. Accordingly, the disclosure provides methods of conjugating an LNP to a targeting moiety that has been modified with an enzyme recognition sequence, wherein said conjugating is accomplished via a Click reaction between a Tz ring covalently bound to the targeting moiety and a TCO moiety bound to the LNP. For instance, the disclosure provides methods of conjugating an LNP to an antibody, Fab fragment or single chain variable fragment (ScFv), wherein the antibody, Fab fragment or ScFv is covalently linked to a first Click handle (Tz ring) through a linker comprising an enzyme recognition sequence and the LNP is covalently linked to a second Click handle (TCO moiety), said method comprising contacting the LNP with an antibody, Fab fragment or ScFv such that first Click handle reacts with the second Click handle to form a Click reaction product (dihydropyridazine) that conjugates the antibody, Fab fragment or ScFv to the LNP. In some embodiments, the antibody, Fab fragment or ScFv is directly bonded to the enzyme recognition sequence. In some embodiments, the antibody Fab fragment or ScFv is bonded to the enzyme recognition sequence via one or more amino acid residues. Particular amino acid residues added that can be covalently attached to the C-terminus of the antibody, Fab fragment or ScFv include, but are not limited to (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10.

[0212] In some embodiments, the enzyme recognition sequence is a sortase recognition motif or a LplA acceptor peptide. In some embodiments where the LNP is conjugated to an

antibody, the C-terminus of one or more of the heavy or light chains of the antibody is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide) either directly or through a linker comprising one or more amino acid residues, as set forth herein. In some embodiments where the LNP is conjugated to a Fab fragment, the C-terminus of the heavy or light chain of the Fab fragment is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide). In some embodiments where the LNP is conjugated to s ScFv, the C-terminus of the ScFv is covalently bonded to the enzyme recognition sequence (e.g., sortase recognition motif or LplA acceptor peptide). In some embodiments, the conjugation efficiency achieved by the disclosed method is greater than 50%, greater than 60%, greater than 70%, greater than 80%, or greater than 90%. In some embodiments, the conjugation efficiency achieved by the disclosed method is from about 60% to about 95%. In some embodiments, the conjugation efficiency achieved by the disclosed methods are from about 70% to about 85%.

[0213] In one embodiment, lipid nanoparticles conjugated to an antibody, Fab fragment or ScFv in a site-specific manner are produced according to the following steps:

- (i) covalently bonding a peptide with a sortase recognition site to one or more C-terminus of the antibody or Fab fragment, or to the C-terminus of the ScFv;
- (ii) contacting the product of step (i) with a first Click handle comprising a Tz ring covalently bound to three or more glycine residues in the presence of a sortase enzyme under conditions suitable for the sortase enzyme to ligate the antibody, Fab fragment or ScFv to the first Click handle;
- (iii) contacting the product of step (ii) with a lipid nanoparticle (LNP), wherein one or more lipids comprising the LNP include a second Click handle capable of reacting with the first Click handle, wherein the second Click handle comprises a TCO moiety, thereby generating a Click product.

[0214] In some embodiments, step (i) involves covalently bonding the peptide with the sortase recognition site to a linker covalently bonded to one or more C-terminus of the antibody or Fab fragment, or to the C-terminus of the ScFv. In some embodiments, the linker comprises one or more amino acid residues. Particular amino acid residues added that can be covalently attached to the C-terminus of the antibody, Fab fragment or ScFv include, but are not limited to (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO:

3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10.

5 [0215] In some embodiments, the sortase recognition motif bonded to the C-terminus of the antibody, Fab fragment or ScFv is bonded to one or more amino acid residues. For instance, in some embodiments, the glycine (G) of the LPXTG (SEQ ID NO: 5) motif can be bonded to one or more (e.g., between 1-10) histidine (H) residues (see **FIG. 1**). The glycine and the histidine residues are removed during the transpeptidation reaction.

10 [0216] In some embodiments, step (ii) (i.e., the sortase-catalyzed reaction) is carried out in the presence of sortase A5. Sortase A5 is an engineered pentamutant variant of the wild-type sortase from *Staphylococcus aureus* that is significantly more active than the wild-type sortase. See U.S. patent No. 9,267,227. In such embodiments, the sortase recognition site bond to the antibody, Fab fragment or ScFv in step (i) has the sequence LPXTG (SEQ ID NO: 5), where X is any amino acid residue. In step (ii), sortase A catalyzes the transpeptidation reaction between the LPXTG (SEQ ID NO: 5) recognition motif and a glycine residue bound to the first Click handle by cleaving the sortase recognition site between the threonine and the glycine residues.

20 [0217] In other embodiments, step (ii) is carried out in the presence of the sortase enzyme *Streptococcus pyogenes* sortase A (SpSrtA WT). In such embodiments, the sortase recognition site bond to the antibody, Fab fragment or ScFv in step (i) has the sequence LPXTA (SEQ ID NO: 6), where X is any amino acid residue. In step (ii), SpSrtA catalyzes the transpeptidation reaction between the LPXTA (SEQ ID NO: 6) recognition motif and a glycine residue bound to the first Click handle by cleaving the sortase recognition site between the threonine and the alanine residues.

25 [0218] **FIG. 1** shows the formation of an LNP conjugated to the C-terminus of a Fab fragment through a sortase mediated ligation followed by a Click reaction. In the schematic, a C-terminus of the Fab fragment is first covalently modified with a sortase recognition motif covalently bonded to six histidine (H) residues. The sortase modified Fab fragment is then reacted with a first Click handle (Click handle 1) covalently bonded to glycine (G) residues. The resultant Fab fragment modified with the first Click handle is then reacted with an LNP
30 that comprises at least one lipid that is covalently bonded to a second Click handle (Click handle 2), thus affording an LNP site-specifically conjugated to the Fab fragment.

[0219] In one embodiment, the conjugate comprises a protein targeting moiety as set forth above (e.g., antibody, Fab fragment or ScFv) and a lipid nanoparticle (LNP), wherein a C-terminus of the targeting moiety is conjugated to the lipid nanoparticle through a linker, and wherein the linker comprises a sortase recognition motif and a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some embodiments, the C-terminus of one or more of the heavy or light chains of the antibody, the C-terminus of the heavy or light chain of the Fab fragment, or C-terminus of the ScFv is conjugated to the lipid nanoparticle through a linker, and wherein the linker comprises a sortase recognition motif and a Click product formed from a Click reaction between a first Click handle on the antibody, Fab fragment, or ScFv and a second Click handle on the LNP. In particular embodiments, the sortase recognition motif comprises an LPXT motif, wherein X is any amino acid residue. In some such embodiments, the sortase recognition motif comprises an LPET motif. In some embodiments, the linker comprises 3 or more glycine residues between the sortase recognition motif and the Click product. In some embodiments, the linker comprises between 3 and 50 glycine residues. In some embodiments, the linker comprises between 3 and 10 glycine residues. In some such embodiments, the linker comprises 3, 4, 5, or 6 glycine residues.

[0220] In some embodiments, the conjugates have the structure Antibody-LPXT(G)_n-Click product-LNP, Fab fragment- LPXT(G)_n-Click product-LNP, or ScFv- LPXT(G)_n-Click product-LNP, wherein X is any amino acid residue and n is between 1 and 20 (e.g., between 3 and 10 or between 3 and 6). In these embodiments, the leucine (L) residue of the sortase recognition motif is bonded to at least one C –terminus of the antibody or Fab fragment or to the C-terminus of the ScFv. In some such embodiments, X is E. In some embodiments, the linker between the antibody (or Fab fragment or ScFv) and the Click includes the following amino acid residues: LPETGGG (SEQ ID NO: 7), LPETGGGG (SEQ ID NO: 8), LPETGGGGG (SEQ ID NO: 9), LPETAGGG (SEQ ID NO: 10) or LPETGGGGGG (SEQ ID NO: 11).

[0221] In some embodiments, the linker comprises additional amino acid residues between the antibody, Fab fragment or ScFv and the sortase recognition motif. In such embodiments, a C-terminus of the antibody, Fab Fragment or ScFv can be covalently modified with one or more amino acid residues prior to covalently linking the sortase recognition motif. For instance, in particular embodiments, the conjugates have the structure Antibody-Z-

LPXT(G)_n-Click product-LNP, Fab fragment-Z-LPXT(G)_n-Click product-LNP, or ScFv-Z-LPXT(G)_n-Click product-LNP, wherein Z is a linker between the antibody (or Fab fragment or ScFv) and the leucine of the sortase recognition motif. In some embodiments, Z comprises one or more amino acid residues. In some embodiments Z is (GGGGS)_v (SEQ ID NO: 1), (G)_v,
 5 (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10. In some embodiments, Z is GG, GGG, GGGG (SEQ ID NO: 12), GGGGG (SEQ ID NO: 13), GGGGGG (SEQ ID NO: 14), and GGGGGGG (SEQ ID NO: 15) or GGGGGGS (SEQ ID NO: 16).

[0222] In some embodiments, the conjugates have the structure Antibody-LPXT(G)_n-
 10 dihydropyridazine-LNP, Fab fragment- LPXT(G)_n- dihydropyridazine -LNP, or ScFv-LPXT(G)_n- dihydropyridazine -LNP, Antibody-Z-LPXT(G)_n- dihydropyridazine-LNP, Fab fragment-Z-LPXT(G)_n- dihydropyridazine -LNP, or ScFv-Z-LPXT(G)_n- dihydropyridazine -LNP, wherein variables n and Z are defined as above. In particular embodiments, the dihydropyridazine moiety is a 1,4- dihydropyridazine.

15 [0223] In another embodiment, the LNP lipid is conjugated to an antibody, Fab fragment or ScFv in a site-specific manner, comprising

- (i) covalently bonding a Lp1A acceptor peptide site to the antibody, Fab fragment or ScFv;
- (ii) contacting the product of step (i) with a carboxylic acid compound that includes
 20 a first Click handle comprising a Tz ring in the presence of a lipoic acid ligase under conditions suitable for lipoic acid ligase to ligate the antibody, Fab fragment or ScFv to the first Click handle; and
- (iii) contacting the product of step (ii) with a lipid nanoparticle (LNP), wherein one or more lipids comprising the LNP includes a second Click handle capable of reacting with the
 25 first Click handle, wherein the second Click handle comprises a TCO moiety, thereby generating a Click product.

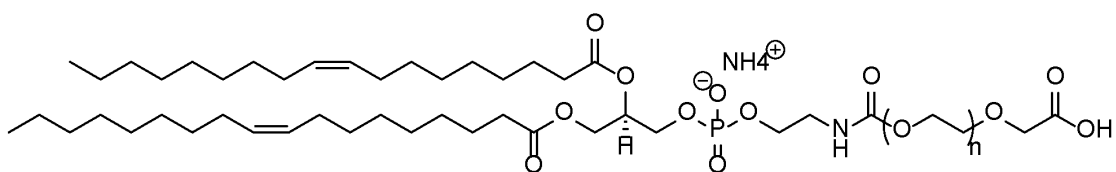
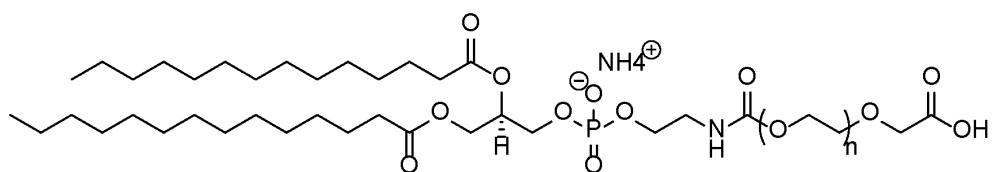
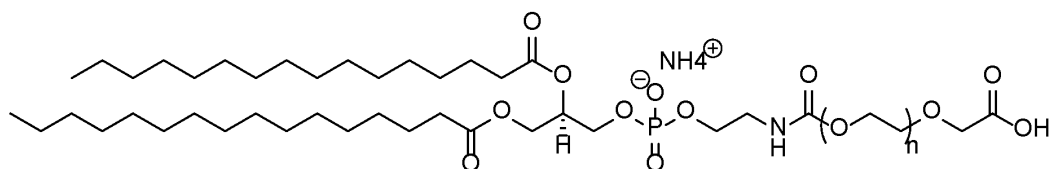
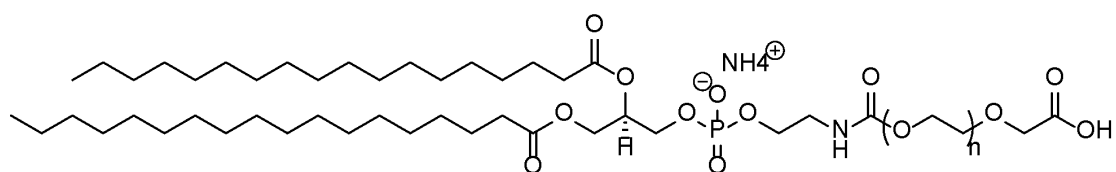
[0224] In some embodiments of step (i) the Lp1A acceptor peptide is covalently bonded to one or more C-terminus of the antibody or Fab fragment. In other embodiments of
 30 step i) the Lp1A acceptor peptide is covalently bonded to the C-terminus of the ScFv.

[0225] In some embodiments, step (i) involves covalently bonding the peptide with the Lp1A acceptor peptide to a linker covalently bonded to one or more C-terminus of the antibody or Fab fragment, or to the C-terminus of the ScFv. In some embodiments, the linker comprises one or more amino acid residues. Particular amino acid residues added that can be covalently
5 attached to the C-terminus of the antibody, Fab fragment or ScFv include, but are not limited to (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10.

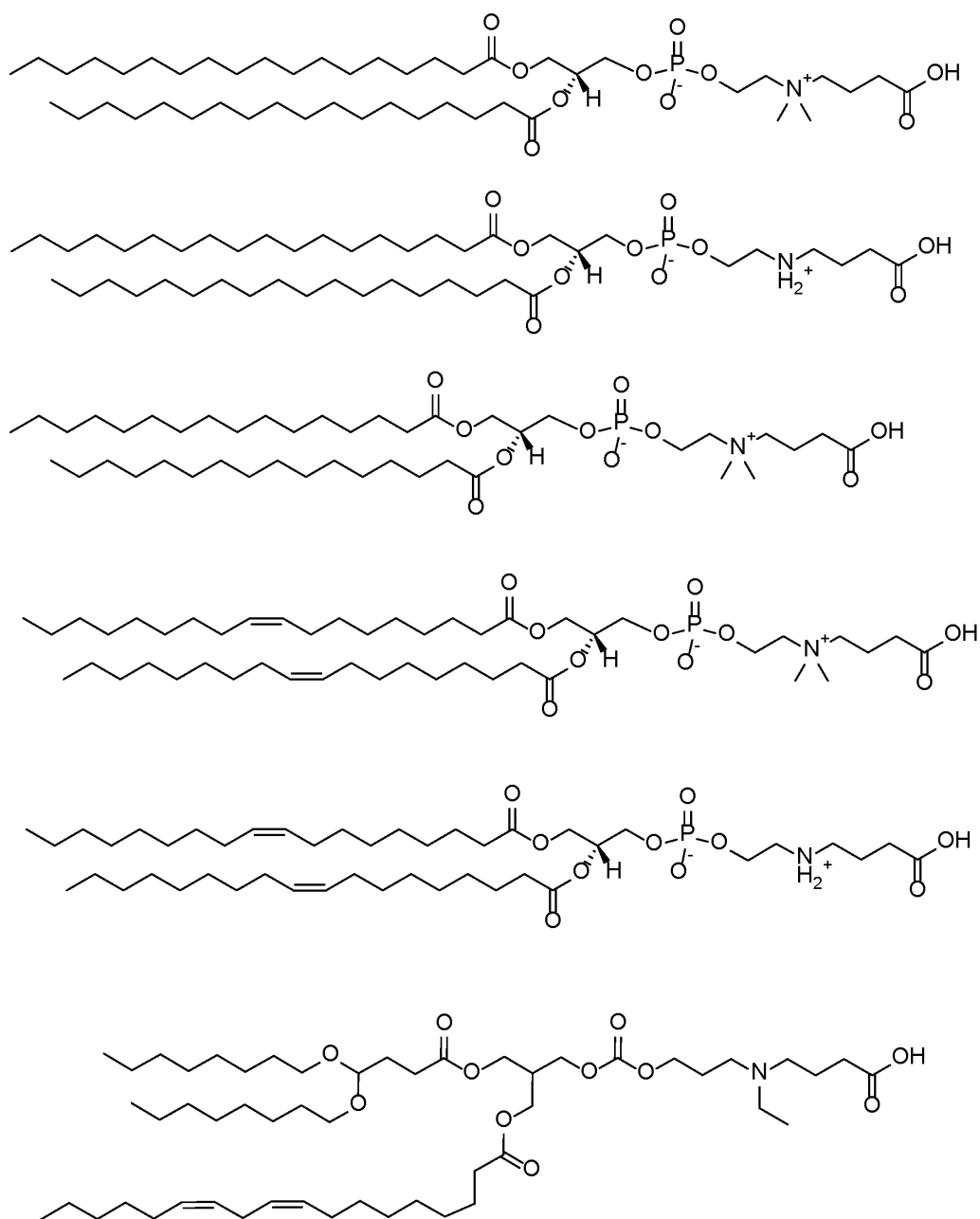
[0226] In some embodiments, step (ii) (i.e., the lipoic acid ligase catalyzed reaction) is
10 carried out in the presence of a mutated lipoic acid ligase, wherein the tryptophan residue (W) at position of the lipoic acid ligase is mutated. Specific W37 mutants are described in Cohen et al., 2012, ChemBioChem, 13, 888-894. In some embodiments, the mutant lipoic acid ligase is selected from W37V, W37I, W37T, W37L, W37C, W37A, and W37G.

[0227] In some embodiments, the carboxylic acid in step (ii) (i.e., the lipoic acid ligase
15 catalyzed reaction) is a C3-C20 carboxylic acid. In some embodiments, the carboxylic acid in step (ii) is a fatty acid, for instance a C7-C19 fatty acid. In some such embodiments, the fatty acid is selected from decanoic acid, palmitic acid, lauric acid or octanoic acid.

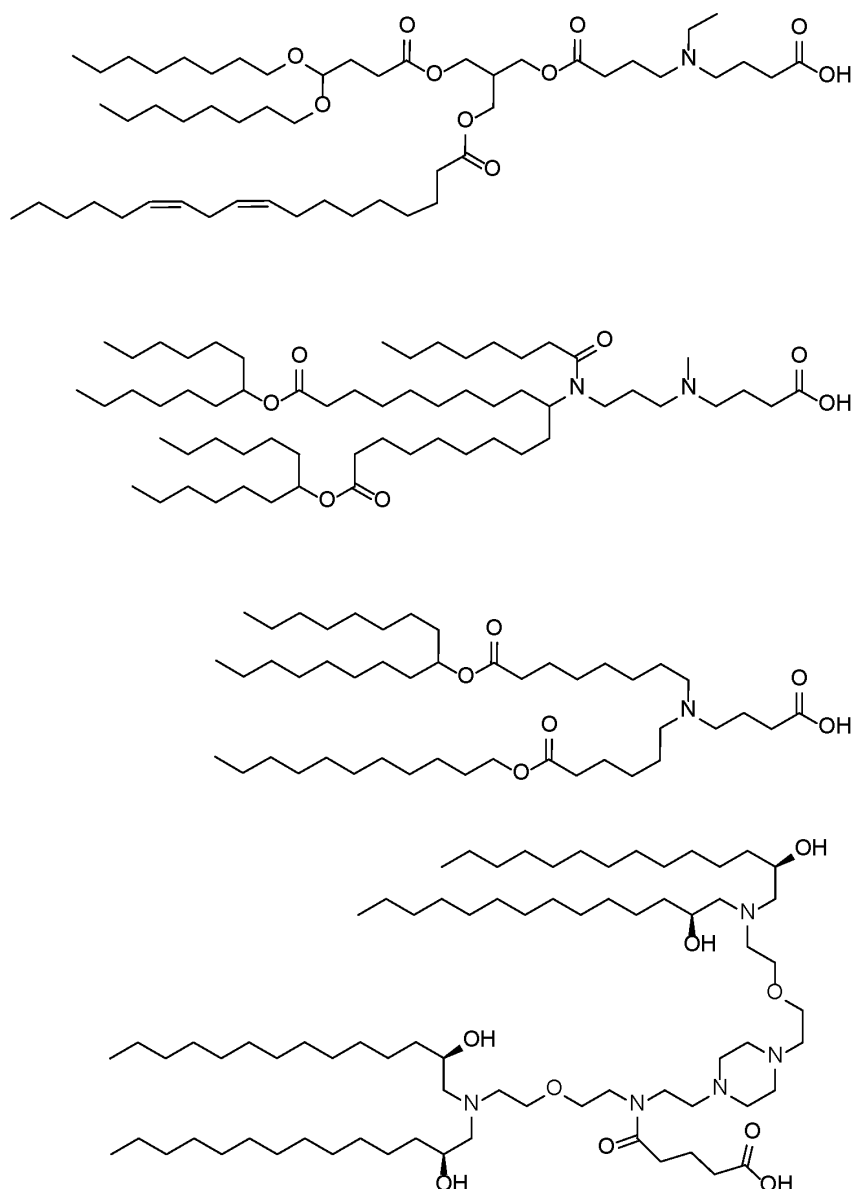
[0228] In other embodiments, the carboxylic acid in step (ii) has the formula COOH-
R-Click handle, where R includes two lipid tails. In some such embodiments, the carboxylic
20 acid as shown below, wherein n = 1—120.



72



73



- [0229] FIG. 2 depicts the formation of an LNP conjugated to the C-terminus of a Fab fragment through a LplA ligation followed by a Click reaction. In the schematic, a C-terminus of the Fab fragment is first covalently modified with a LplA acceptor peptide. Specifically, a first Click handle (Click handle 1) comprising a Tz ring is first introduced by chemically reacting a carboxylic acid to the lysine residue of the LplA acceptor peptide. The resultant Fab fragment modified with the first Click handle is then reacted with an LNP that comprises at

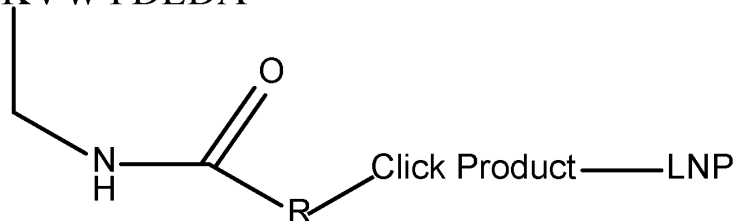
least one lipid that is covalently bonded to a second Click handle comprising a TCO moiety (Click handle 2), thus affording an LNP site-specifically conjugated to the Fab fragment.

[0230] In some embodiments, the conjugate produced by methods disclosed herein comprises a protein targeting moiety as set forth above (e.g., antibody, Fab fragment or ScFv),
5 conjugated to the lipid nanoparticle through a linker, and wherein the linker comprises a lipopeptide (LplA) acceptor peptide and a Click product formed from a Click reaction between a first Click handle on the targeting moiety and a second Click handle on the LNP. In some
10 embodiments, the linker further comprises one or more additional amino acid residues between the protein targeting moiety (e.g., antibody, Fab fragment or ScFv) and the LplA acceptor peptide. In some embodiments, the LplA acceptor peptide has the sequence GFEDKVWYDLDA (SEQ ID NO: 17). In some embodiments, the conjugate comprises an antibody, wherein the C-terminus of one or more of the heavy or light chains of the antibody is bonded to the linker. In some embodiments, the conjugate comprises a Fab fragment, the C-terminus of the heavy or light chain of the Fab fragment is bonded to the linker. In some
15 embodiments, the conjugate comprises a ScFv, wherein the C-terminus of the ScFv is bonded to the linker.

[0231] In some embodiments, the linker comprises additional amino acid residues between the targeting moiety (e.g., antibody, Fab fragment or ScFv) and the LplA acceptor peptide. In such embodiments, a C-terminus of the targeting moiety (e.g., antibody, Fab
20 Fragment or ScFv) can be covalently modified with one or more amino acid residues prior to covalently linking the LplA acceptor peptide. For instance, in particular embodiments, the conjugates have the structure targeting moiety-Z-LplA acceptor peptide-Click product-LNP (e.g., Antibody-Z- LplA acceptor peptide -Click product-LNP, Fab fragment-Z- LplA acceptor peptide -Click product-LNP, or ScFv-Z- LplA acceptor peptide -Click product-LNP), wherein
25 Z is a linker between the antibody (or Fab fragment or ScFv) and the glycine residue of the LplA acceptor peptide. In some embodiments, Z comprises one or more amino acid residues. In some embodiments Z is (GGGGS)_v (SEQ ID NO: 1), (G)_v, (EAAAK)_v (SEQ ID NO: 2), (PAPAP)_v (SEQ ID NO: 3), (AP)_v and A(EAAAK)_u ALEA(EAAAK)_vA (SEQ ID NO: 4), wherein u is 1-10 and v is 1-10. In some embodiments, Z is GG, GGG, GGGG (SEQ ID NO:
30 12), GGGGG (SEQ ID NO: 13), GGGGGG (SEQ ID NO: 14), and GGGGGGG (SEQ ID NO: 15) or GGGGGGS (SEQ ID NO: 16).

[0232] It will be understood that in the forgoing embodiments, the lysine (K) residue of the LplA acceptor peptide is covalently linked to Click product, which is covalently linked to the LNP. Specifically, to generate conjugates, the side chain lysyl group reacts with the carboxylic acid compound that includes the first Click handle. The resultant modified targeting moieties (antibodies or Fab fragments or ScFvs) are reacted with an LNP that has been modified with a second Click handle, as disclosed herein, thereby generating a Click product. An LNP surface modified with a Fab fragment is depicted below, where R is a lipid group (e.g., C2-C30 alkyl group).

Fab-Z-GFEIDKVWYDLDA



(SEQ ID NO: 18)

[0233] In other embodiments, the enzyme recognition sequence is a transglutaminase enzyme recognition sequence (LLQG (SEQ ID NO: 19)). The transglutaminase enzyme recognition sequence (LLQG (SEQ ID NO: 19)) is also referred to as Q-tag. The Q-tag may be present on or can be inserted at one or more locations of targeting moiety, (e.g., antibody, Fab fragment or ScFv), for instance at a C-terminus. The transglutamine enzyme catalyzes the reaction between a side-chain amide group on the Q-tag and an alkyl-primary amine on a component of the LNP (e.g., a lipid), thus linking the antibody to the LNP through an amide bond.

[0234] In other embodiments, the enzyme recognition sequence is a sequence recognized by formylglycine generating enzyme, specifically CXPXR, wherein each X is any amino acid. In such embodiments, the CXPXR sequence can be inserted at one or more locations of the targeting moiety (e.g., antibody, Fab fragment or ScFv), for instance at a C-terminus. The formylglycine generating enzyme converts the cysteine thiol of CXPXR into an

aldehyde group, which can be reacted with an aminooxy or hydrazine group covalently bonded to a component of the LNP.

[0235] In another embodiment, the lipid nanoparticle is conjugated to the targeting moiety, e.g., antibody, Fab fragment or ScFv, in a site-specific manner through a sugar moiety on a glycosylated antibody. **FIG. 3** shows one particular embodiment of introducing a Tz Click handle through a sugar moiety on an antibody. In step 1, an azide moiety is site-specifically introduced into a sugar moiety of the targeting antibody catalyzed by the enzymes galactosidase and transferase. The azide functionality then reacts with a dibenzocyclooctene (DBCO)-MeTz via a Click reaction. The Click reaction is used to introduce the Tz ring, in this case MeTz onto the antibody, which can then be reacted with a TCO moiety on an LNP (not shown) in a second Click reaction, hence affording a conjugate.

[0236] In another embodiment, the LNP is conjugated to a targeting moiety (e.g., antibody, Fab fragment or ScFv) in a site-specific manner through a light-induced crosslinking reaction. **FIG. 4** shows one particular embodiment using oYo-Link Tz. In **FIG. 4**, a Tz ring (e.g., MeTz) is covalently bound to a site specific antibody label (oYo link) that specifically reacts with residues on the heavy chain of the antibody. The oYo-Link tetrazine covalently bound to the antibody reacts with TCO (not shown) bound to an LNP, hence generating the surface-modified LNP.

[0237] In another embodiment, the lipid nanoparticle is conjugated to the targeting moiety (e.g., antibody, Fab fragment or ScFv) in a site-specific manner by mutating an amino acid residue on the targeting moiety and subsequently reacting the mutated targeting moiety with a compound that includes a Tz ring. **FIG. 5** shows particular embodiments where an azide functionality is introduced during antibody or Fab fragment production. The azide group then undergoes a Click reaction with a DBCO group that is linked to a tetrazine ring. The tetrazine ring bound to the antibody reacts with TCO (not shown) bound to an LNP, hence generating the surface modified LNP.

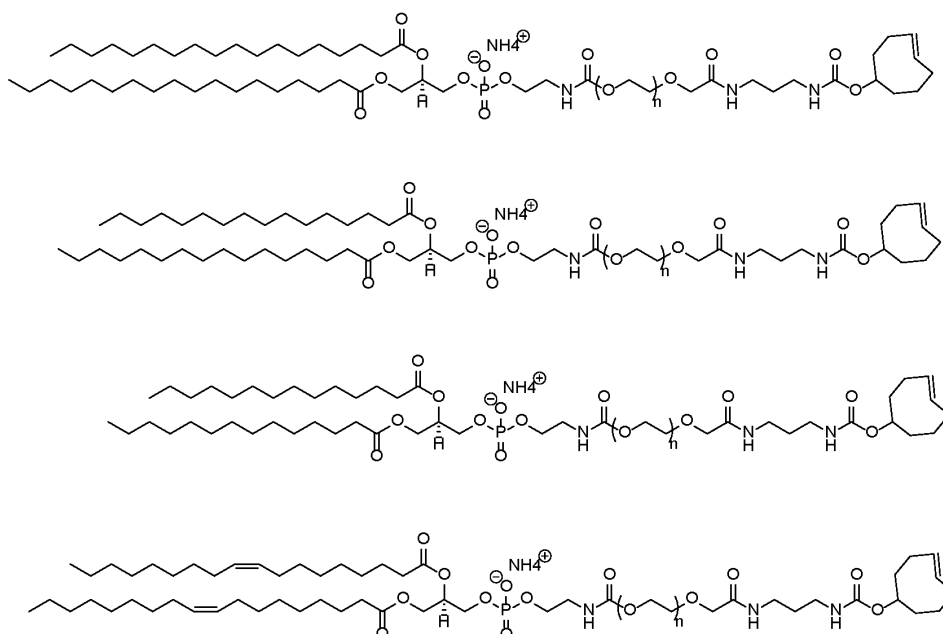
[0238] In another embodiment, the lipid nanoparticle is conjugated to the targeting moiety (e.g., antibody, Fab fragment or ScFv) in a site-specific manner by using a cysteine-maleimide reaction to introduce the first Click handle onto the targeting moiety (antibody, Fab fragment or ScFv). **FIG. 6** shows one particular embodiment of introducing a first Click handle (e.g., Tz) onto a Fab fragment. In **FIG. 6**, a hinge with a free cysteine is first introduced at a

C-terminus of the Fab fragment. In some such embodiments, an engineered inter-chain disulfide bound away from the C-terminus and buried in the CL-CH1 interface is introduced for improved stability (**FIG. 7**). The resultant modified Fab fragment is reacted with a maleimide moiety with a first Click handle (Click handle 1) covalently bonded to the maleimide nitrogen. The first Click handle then reacts with a second Click handle (Click handle 2) on the LNP, hence generating the surface modified targeted LNP. In some embodiments, Click handle 1 is a TZ and Click handle 2 is a TCO moiety,

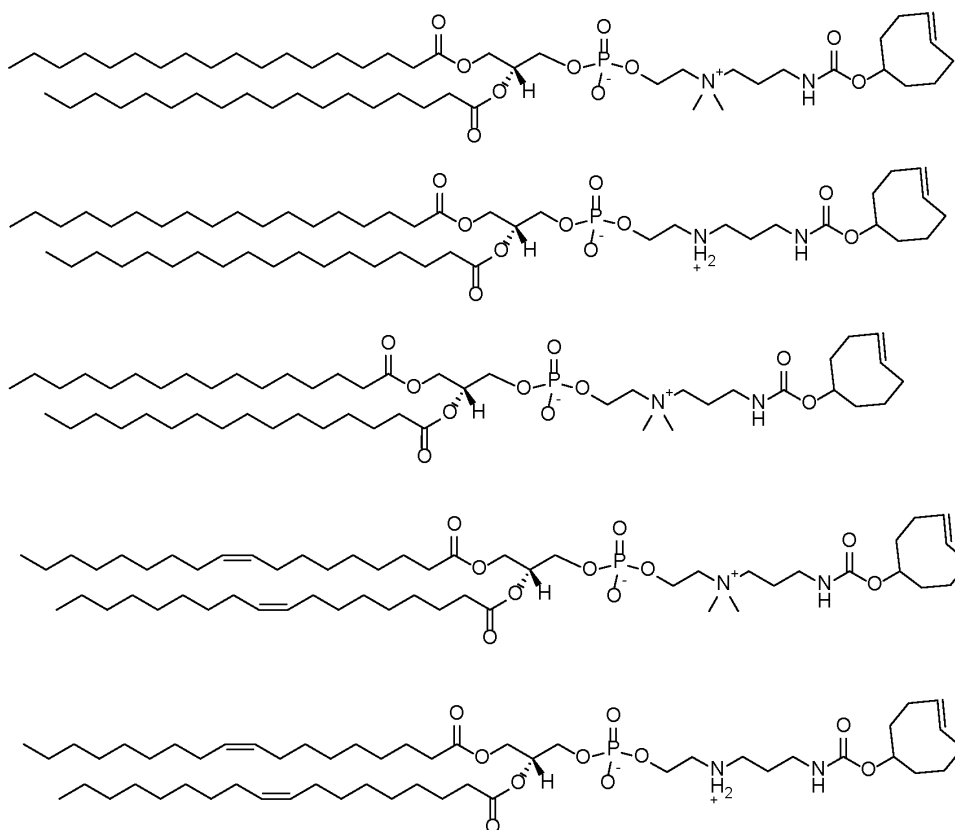
[0239] In some embodiments, the Click product can be formed using any suitable photo-induced Click chemistry reaction. In some embodiments, the Click product can be formed using photoinducible 1,3-dipolar cycloaddition reaction between a tetrazole and an alkene (*see, e.g., Song et al., Angew. Chem., Int. Ed.* 2008, 47 (15), 2832-2835).

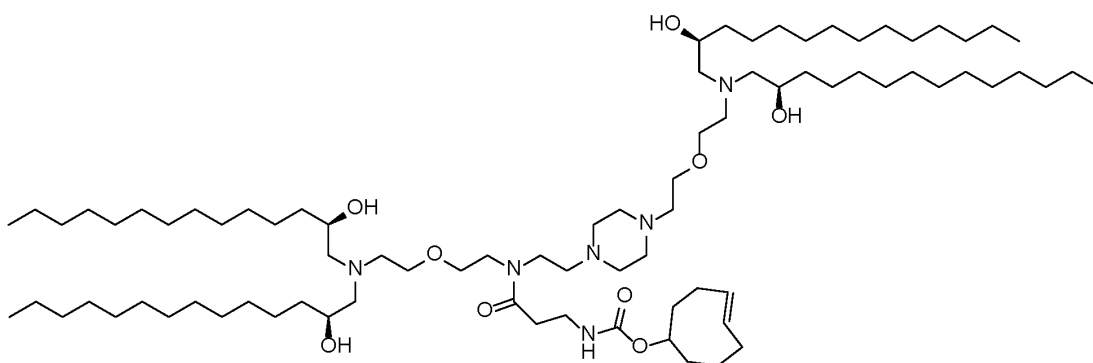
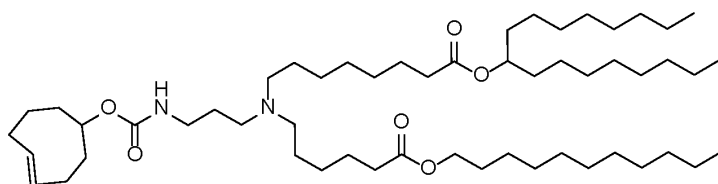
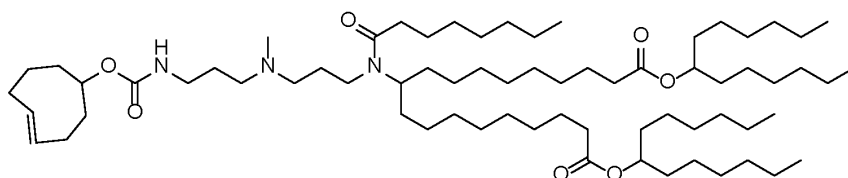
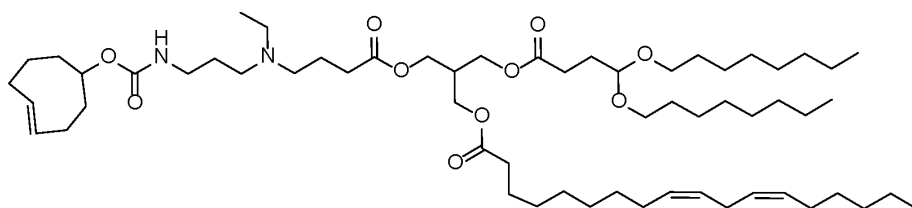
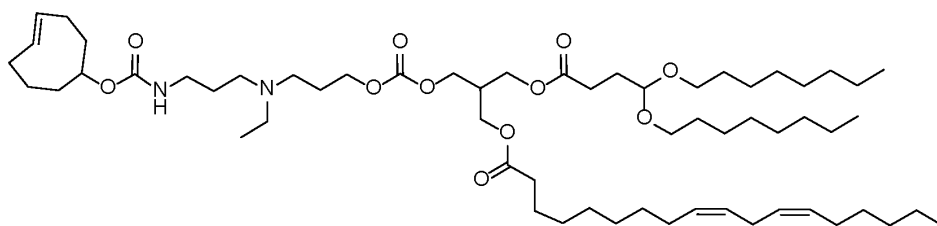
[0240] In some embodiments, the Click product can be formed using oxime and hydrazone ligations. In some embodiments, a ketone or aldehyde can react with α effect amine, such as hydroxylamine, hydrazine and hydrazide (*see, e.g., Agten et al., ChemBioChem* 2013, 14 (18), 2431-2434 and Dirksen et al., *J. Am. Chem. Soc.* 2006, 128 (49), 15602-15603).

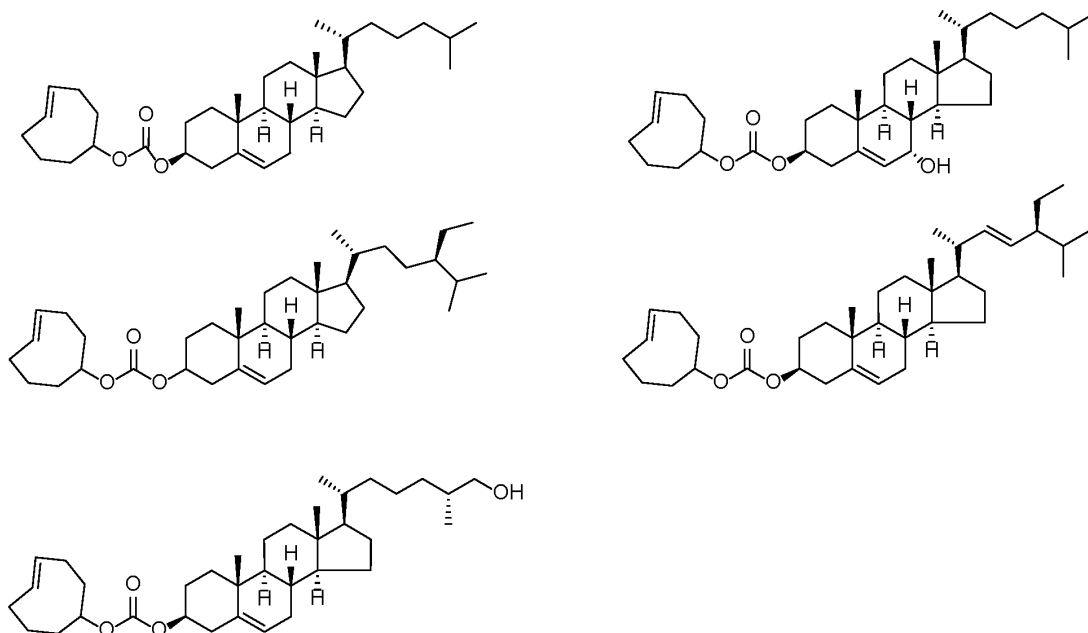
[0241] Exemplary lipids and cholesterol molecules bonded to Click handles that can be used to make targeted LNPs are depicted below.



78







[0242] Conjugates prepared by site-specific methods (e.g., **FIGs. 1-6**) have a high density of the targeting moiety on the surface of the LNP. For instance, the conjugate can comprise more than one targeting moiety (e.g., antibody, Fab fragment or ScFv) per LNP. In some embodiments, the conjugate can comprise more than 10 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise more than 20 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise more than 30 targeting moieties (e.g., antibodies, Fab fragments or ScFvs). In some embodiments, the conjugate can comprise more than 50 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise more than 75 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise more than 100 targeting moieties (e.g., antibodies, Fab fragments or ScFvs). In some embodiments, the conjugate can comprise from about 50 to about 200 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise from about 100 to about 200 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise from about 100 to about 230 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP. In some embodiments, the conjugate can comprise from about 10 to about 150 targeting moieties (e.g., antibodies, Fab fragments or

ScFvs) per LNP. In some embodiments, the conjugate can comprise from about 10 to about 30 targeting moieties (e.g., antibodies, Fab fragments or ScFvs) per LNP.

[0243] In some embodiments, the weight ratio between a targeting moiety on the surface of the LNP and the payload (e.g., RNA) encapsulated in the LNP can be 1:20, 1:15, 1:10, 1:9, 1:8, 1:7, 1:6, 1:5, 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 15:1, 20:1.

Interchain C_L - C_{H1} disulfide reduction followed by conjugation

[0244] IgG antibodies consist of four polypeptide chains linked by disulfide bonds. The two polypeptide chains of low molecular weight are called light chains (L). The light chains consist of a variable light chain domain (VL) and a constant light chain domain (CL). The heavy chains consist of a variable heavy light domain (VH) and three constant heavy chain domains (CH1, CH2, and CH3). The Fab region of the antibody includes the VL, CL, VH, and CH1 domains. The Fc region includes the constant heavy chain domains CH2, and CH3. A hinge region of the IgG antibody covalently links the CH1 domain to the CH2 domain. The two heavy chains of IgG antibodies are connected in the hinge region by a variable number of disulfide bonds depending on the IgG subclass. Different subclasses of IgG antibodies have varying numbers of interchain disulfide bonds. Additionally, the light chain is covalently linked to the heavy chain via a disulfide bond between the light chain and the heavy chain. Using standard IgG nomenclature, this natural interchain disulfide bond is also referred to as the CL-CH1 disulfide bond to distinguish it from disulfide bonds present in the hinge region. Therapeutic antibodies of type IgG1 possess an intermolecular disulfide bond between Cys233 (Kabat numbering) of the heavy domain and Cys214 (Kabat numbering) of the light domain. Therapeutic antibodies of type IgG4 possess an intermolecular disulfide bond between Cys127 (Kabat numbering) of the heavy domain and Cys214 (Kabat numbering) of the light domain. Therapeutic antibodies of type IgG2 possess an intermolecular disulfide bond between Cys135 (Kabat numbering) of the heavy domain and Cys214 (Kabat numbering) of the light domain.

[0245] Proteolytic cleavage of an IgG antibody results in the formation of a Fab fragment known as a $F(ab')_2$ fragment. The $F(ab')_2$ fragment does not include the CH2 domain or the CH3 domain. However, the hinge region of the antibody is retained in a $F(ab')_2$ fragment. The $F(ab')_2$ fragment includes disulfide bonds that covalently link two Fab fragments. Reduction of the disulfide bond in the $F(ab')_2$ generates two $F(ab')$ fragments. The

sulfhydryl (thiol) groups of the F(ab') could potentially react with a thiol-reactive group on the surface of an LNP, hence generating a conjugate. However, owing to the presence of multiple sulfhydryl groups in the hinge region of the F(ab') fragment, site-specific conjugation is challenging. Moreover, the reduction of the F(ab')₂ to the F(ab') fragments could also disrupt the natural interchain disulfide bonds between the CL and CH1 regions of the Fab fragments, hence further compromising site-specific conjugation.

[0246] As set forth herein, Fab fragment can be site-selectively conjugated to the surface of a precursor lipid nanoparticle (LNP) through the natural interchain disulfide bond between the heavy chain and the light chain (i.e., the CL-CH1 disulfide bond) of the Fab fragment to make a targeted LNP (conjugate).

[0247] The term "precursor LNP" or "base LNP" refers to an LNP that has been functionalized with a reactive moiety (e.g., thiol-reactive group or polyglycine) prior to reacting with the Fab fragment. The process for conjugating a targeting moiety, as disclosed herein, involves reducing the natural interchain disulfide bond between the CL and CH1 domains of a Fab fragment, and reacting the reduced Fab fragment with a thiol-reactive group (e.g., a maleimide or DBM group) covalently bonded to the surface of a precursor LNP, thus forming a conjugate. Alternatively, the reduced Fab fragment can be reacted with a lipid that has been chemically modified (functionalized) with a thiol-reactive group (e.g., maleimide or DBM group). The resultant lipid can then be inserted into a preexisting LNP, thus generating a conjugate. As described herein, despite the removal of the natural interchain disulfide bond linking the heavy and light chains of the Fab fragment, the resulting conjugates are able to effectively target specific cell types depending on the nature of the Fab targeting moiety. For instance, specific Fab fragments for targeting immune cells or hematopoietic stem cells (HSCs) as disclosed herein. A schematic of an LNP site-specifically conjugated to a Fab fragment is shown in FIG. 25.

[0248] In one embodiment, Fab fragments used for conjugation may be used by recombinant methods. In particular embodiments, the Fab fragments generated recombinantly are designed not to include a hinge region at the C-terminus. Therefore, the recombinantly generated Fab fragments include only one disulfide bond between the CL-CH1 and domains. As set forth herein, the CL-CH1 can then be reduced and the resultant free thiol groups can be used as anchors to conjugate the Fab fragment to the surface of an LNP.

[0249] In some embodiments, the Fab fragment is of the IgG class, the IgM class, or the IgA class. In some embodiments, the Fab fragment is of the IgG class and has an IgG1, IgG2, IgG3, or IgG4 isotype. In some embodiments, the Fab fragment is a native protein. In some embodiments, the Fab fragment is an engineered protein.

5 [0250] In one aspect, the disclosure provides methods of making a targeted LNP, said method comprising:

(i) contacting a composition comprising a Fab fragment with a reducing reagent, wherein the Fab fragment comprises a heavy chain and a light chain and an interchain disulfide bond linking the constant light chain domain (C_L) and the constant heavy chain
10 domain 1 (C_{H1}), whereby the reducing reagent reduces the interchain disulfide bond of the Fab fragment to generate two free cysteine residues; and

(ii) contacting the product of step (i) with a precursor LNP comprising a plurality of thiol-reactive groups covalently bonded to one or more lipids of the precursor LNP, thereby forming a targeted LNP.

15 [0251] In some aspects, the disclosure provides a conjugate produced by a method comprising:

(i) contacting a composition comprising a Fab fragment with a reducing reagent, wherein the Fab fragment comprises a heavy chain and a light chain and an interchain disulfide bond linking the constant light chain domain (C_L) and the constant heavy chain
20 domain 1 (C_{H1}), whereby the reducing reagent reduces the interchain disulfide bond of the Fab fragment to generate two free cysteine residues; and

(ii) contacting the product of step (i) with a precursor LNP comprising a plurality of thiol-reactive groups covalently bonded to one or more lipids of the precursor LNP, thereby forming a targeted LNP.

25 [0252] In some embodiments, the thiol-reactive group (e.g., maleimide, pyridyl disulfide, 2,3-dibromomaleimide, or haloacetyl) is chemically reacted with a lipid molecule to create a modified lipid wherein the thiol-reactive group is covalently attached to the lipid where it is capable of reacting with at least one free cysteine residue of the reduced Fab fragment (either on the heavy or light chain of the Fab fragment). The reaction between the
30 thiol-reactive group and the at least one free cysteine residue can be completed prior to or after formation of the LNP with the modified lipid. For instance, as shown in **FIG. 26A**, the various

components (e.g., lipids) comprising the LNP and a therapeutic payload can be mixed with lipid molecules, including one or more lipids that comprise a thiol-reactive group, thus generating an LNP that comprises a plurality of thiol-reactive groups (each thiol-reactive group shown schematically as a “functional group” in **FIG. 26A**). The thiol-reactive group can then be reacted with at least one free cysteine residue of the Fab fragment, hence generating a conjugate. Alternatively, a lipid that has been modified with the thiol-reactive group can be directly reacted with at least one free cysteine residue of a Fab fragment. As depicted in **FIG. 26B**, the resultant modified lipid attached to the Fab fragment can then be inserted into a pre-formed LNP that has not yet been surface modified. This procedure allows for the reaction to be performed on an individual lipid molecule rather than on the surface of the LNP.

[0253] Any suitable reducing reagent can be used to reduce the interchain disulfide bond of the Fab fragment. Examples of reducing reagents include, but are not limited to, 2-mercaptoethanol, 2-mercaptoethylamine, dithiothreitol (DTT), dithioerythritol (DTE), and tris(carboxyethyl)phosphine (TCEP), and combinations thereof. In some embodiments, the reducing reagent is a mild reducing reagent. Examples of mild reducing reagents include, e.g., DTT, TCEP, and DTE. In some embodiments, the reducing reagent is TCEP. Any suitable reaction conditions can be used for the reduction of the interchain disulfide bond in step (i). In some embodiments, the reduction reaction can occur in water, aqueous buffer, or cell culture media. In some embodiments, the reduction reaction is performed at physiological pH (e.g., about 7.4). In some embodiments, the reduction reaction is performed at physiological temperature (e.g., about 37° C). In some embodiments, the reduction reaction is performed between 0°C and 40°C, e.g., between 10°C and 35°C, between 15°C and 30°C, between 20°C and 30°C, or between 20°C and 25°C. In some embodiments, the reduction reaction is performed at ambient temperature (e.g., about 23 to about 25° C). In some embodiments, the reduction reaction is performed at about 0° C to about 4° C.

[0254] In some embodiments, excess reducing agent is removed following step (i), prior to conjugation to the precursor LNP. In some embodiments, excess reducing agent is not removed following step (i), prior to conjugation to the precursor LNP.

[0255] In some embodiments of the method or process, the thiol-reactive groups on the LNP (or lipid to be post-inserted into an LNP) comprises any suitable reactive group, including but not limited to, maleimide, pyridyl disulfide, 2,3-dibromomaleimide, or haloacetyl.

[0256] In some embodiments, the thiol-reactive group is maleimide. In some embodiments, maleimide reacts with one of the two free cysteine residues of the Fab fragment (either on the heavy or light chain) to form a thiosuccinimide moiety. In some embodiments, maleimide reacts with a free cysteine residue on the heavy chain of the Fab fragment. In some
5 embodiments, maleimide reacts with a free cysteine residue on the light chain of the Fab fragment. In some embodiments, two maleimide groups each react with the Fab fragment, wherein one maleimide reacts with a free cysteine residue on the light chain and the other maleimide reacts with a free cysteine residue on the heavy chain.

[0257] Any suitable conditions can be used for the reaction between the thiol-reactive
10 group and at least one of the two free cysteine residues of the Fab fragment in step (ii). In some embodiments, the reaction can occur in water, aqueous buffer, or cell culture media. In some embodiments, the reaction is performed at physiological pH (*e.g.*, about 7.4). In some embodiments, the reaction is performed at physiological temperature (*e.g.*, about 37° C). In some embodiments, the reduction reaction is performed between 0°C and 40°C, *e.g.*, between
15 10°C and 35°C, between 15°C and 30°C, between 20°C and 30°C, or between 20°C and 25°C. In some embodiments, the reaction is performed at ambient temperature (*e.g.*, about 23 to about 25° C). In some embodiments, the reaction is performed at about 0° C to about 4° C.

[0258] A schematic showing the reduction of an interchain disulfide bond in a Fab fragment is depicted in **FIG. 27**. In some embodiments, the Fab fragment depicted in **FIG. 27**
20 can be produced recombinantly without a hinge region at the C-terminus. Accordingly, in such embodiments, the Fab fragment only comprises a single interchain disulfide bond, which is located in the C_L-C_{H1} interface of the Fab fragment. As shown in **FIG. 27**, the interchain disulfide bond is located between the heavy chain and the light chain of the Fab fragment. The reduction reaction breaks the covalent linkage forming the disulfide bond, thereby generating
25 two free cysteine residues that can react in a subsequent step with a thiol-reactive group.

[0259] **FIG. 28** shows an exemplary schematic of conjugate formation, as described herein. In **FIG. 28**, a Fab fragment comprising an interchain disulfide bond between the heavy and light chain is contacted with a reducing reagent, whereby the reducing reagent reduces the interchain disulfide to generate two free cysteine residues (step (i)). In step (ii), the reduced
30 Fab fragment is reacted with an LNP comprising a plurality of thiol-reactive groups (*e.g.*, maleimide or DBM) conjugated to the surface of the LNP, whereby the thiol-reactive groups react with the free cysteine residues of the reduced Fab fragment. The Fab fragment is site-

specifically conjugated to the surface of the LNP via a linkage through at least one of the free cysteine residues of the Fab fragment.

[0260] Following reaction of the reduced Fab fragment with the thiol-reactive group of the LNP, either the heavy chain, light chain or both the heavy chain and light chain of the Fab fragment are conjugated to the surface of the LNP. The various orientations are depicted in **FIG. 29**. In some embodiments, the conjugates formed by methods disclosed herein include at least one, at least two or all three orientations shown in **FIG. 29**. The concentration of the thiol-reactive group (e.g., maleimide) will likely determine which orientation is dominant. In some embodiments, increasing the number of thiol-reactive groups on the LNP increases the number of Fab fragments conjugated to two thiol-reactive groups. In some embodiments, decreasing the number of thiol-reactive groups on the LNP decreases the number of Fab fragments conjugated to two thiol-reactive groups. Regardless of the orientation, the heavy chain and the light chain remain intact on the surface of the LNP, thus forming a functional Fab fragment that is capable of engaging with a receptor on a targeted cell.

[0261] In some embodiments, the thiol-reactive group is maleimide, as shown in **FIG. 29**. In some embodiments, maleimide reacts with one of the two free cysteine residues of the antibody or antigen-binding fragment thereof. In some embodiments, maleimide reacts with one of the two free cysteine residues of the Fab fragment to form a thiosuccinimide moiety. In some embodiments, maleimide reacts with a free cysteine residue on the heavy chain of the Fab fragment. In some embodiments, maleimide reacts with a free cysteine residue on the light chain of the Fab fragment. In some embodiments, two maleimide groups on the LNP each react with the Fab fragment, wherein one maleimide reacts with a free cysteine residue on the light chain and the other maleimide reacts with a free cysteine residue on the heavy chain.

[0262] In some embodiments, the thiol-reactive group is 2,3-dibromomaleimide (DBM) as shown in **FIG. 30**. Following reduction of the disulfide bond, the reduced Fab fragment is added to DBM covalently bonded to a lipid (represented by squiggly line in **FIG. 30**). As set forth above, the lipid may be part of an LNP or may be post-inserted into an LNP following reaction with the Fab fragment. Both of the free cysteine residues displace the two bromine groups of DBM, hence generating a dithiomaleimide. The dithiomaleimide can be converted to the corresponding maleamic acid via hydrolysis. DBM reacts with a free cysteine residue on the heavy chain and a free cysteine residue on the light chain of the Fab fragment to

form a bridge between the cysteine residues. Accordingly, the heavy and light chain of the Fab fragment are effectively bridged together following reaction with DBM.

[0263] In all embodiments discussed above, the thiol-reactive group can be introduced onto any of the lipids comprising the LNP. In some embodiments, the conjugate can comprise one or more pegylated lipid molecules. In some embodiments, the thiol-reactive group is covalently bonded to at least one of the pegylated lipid molecules, hence generating the structure Lipid-PEG_x-thiol-reactive group, wherein x is 2-120 ethylene glycol units. In such embodiments, at least one free cysteine residue of an Fab fragment reacts with the thiol-reactive group bonded to the one or more of the pegylated lipids comprising the LNP. In some embodiments, the LNP comprises from about 0.05 mol % to about 2 mol % of the pegylated lipid bonded to the thiol-reactive group. In some embodiments, the PEG spacer between the lipid and the thiol-reactive group comprises at least about 5, 10, 20, 30, 50, 60, 70, 80, 90, 200, or 110 ethylene glycol units. In some embodiments, the PEG spacer comprises about 10-120 ethylene glycol units. In some embodiments, the molecular weight of the pegylated lipid bonded to the thiol-reactive group is from about 500 (i.e., PEG500) to about 5,000 (i.e., PEG5000). In some embodiments, the molecular weight of the pegylated lipid bonded to the thiol-reactive group is from about 1,000 (i.e., PEG1000) to about 3,000 (i.e., PEG3000). In some embodiments, the thiol-reactive group is bonded to at least one of the non-pegylated phospholipids comprising the LNP. In some embodiments, the thiol-reactive group is bonded to at least one of the ionizable lipids comprising the LNP. In some embodiments, the thiol-reactive group is bonded to at least one of the sterol molecules comprising the LNP. In some embodiments, the lipid portion of the pegylated lipid bonded to the thiol-reactive group is selected from DMG, DPG, DSG, DTA, DOPE, DPPE, DMPE, DSPE, sphingosine, sphingomyelin, and stearic acid. **FIG. 32** shows some examples of a pegylated lipid bonded to a thiol-reactive group, in this case a maleimide moiety.

[0264] In some embodiments, the thiol-reactive group is bonded to at least one of the non-pegylated phospholipids (helper lipids) comprising the LNP. **FIG. 33** shows some examples of a non-pegylated lipid bonded to a maleimide moiety.

[0265] In some embodiments, the thiol-reactive group is bonded to at least one of the ionizable lipids comprising the LNP. **FIG. 34** shows some examples of an ionizable lipid bonded to a maleimide moiety.

[0266] In some embodiments, the thiol-reactive group is bonded to at least one of the sterol molecules comprising the LNP. **FIG. 35** shows some examples of sterols bonded to a maleimide moiety.

[0267] The disclosed methods provide stable conjugates that display excellent ability to transduce specific targeted cells. In some aspects, the disclosure provides a conjugate comprising an LNP and a Fab fragment, wherein the LNP is covalently bonded to either or both a first cysteine residue in the constant region of the heavy chain of the Fab fragment and a second cysteine residue in the constant region of light chain of the Fab fragment. In some embodiments, the Fab fragment does not comprise a disulfide bond linking the constant region of the heavy chain of the Fab fragment and the constant region of the light chain of the Fab fragment. In some embodiments, both the constant region of the heavy chain constant region of the heavy chain and the constant region of the light chain of the Fab fragment are covalently bonded to the LNP. In some embodiments, only the constant region of the heavy chain of the Fab fragment is covalently bonded to the LNP. In some such embodiments, the light chain remains associated with the covalently bond heavy chain on the surface of the LNP. In some embodiments, only the constant region of the light chain of the Fab fragment is covalently bonded to the LNP. In some such embodiments, the heavy chain remains associated with the covalently bond light chain on the surface of the LNP. In some of the foregoing embodiments, the Fab fragment is linked to the LNP through a thiosuccinimide moiety. In other of the foregoing embodiments, the Fab fragment is linked to the LNP through a dithiomaleimide moiety (see **FIG. 30**). In still other of the foregoing embodiments, the Fab fragment is linked to the LNP through a maleamic acid moiety (see **FIG. 30**).

[0268] In some embodiments, the Fab fragment conjugated to the LNP is an IgG1 Fab fragment. In some such embodiments, the cysteine at position 214 (Kabat numbering) of the light chain of the IgG1 Fab fragment is covalently bonded to the LNP. In other such embodiments, the cysteine at position 233 (Kabat numbering) of the heavy chain of the IgG1 Fab fragment is covalently bonded to the LNP. In other such embodiments, the cysteine at position 233 (Kabat numbering) of the heavy chain of the IgG1 Fab fragment and the cysteine at position 214 (Kabat numbering) of the light chain of the IgG1 Fab fragment are covalently bonded to the LNP.

[0269] In some embodiments, the Fab fragment conjugated to the LNP is an IgG2 Fab fragment. In some such embodiments, the cysteine at position 214 (Kabat numbering) of the

light chain of the IgG2 Fab fragment is covalently bonded to the LNP. In other such embodiments, the cysteine at position 127 (Kabat numbering) of the heavy chain of the IgG2 Fab fragment is covalently bonded to the LNP. In other such embodiments, the cysteine at position 127 (Kabat numbering) of the heavy chain of the IgG2 Fab fragment and the cysteine at position 214 (Kabat numbering) of the light chain of the IgG2 Fab fragment are covalently bonded to the LNP.

[0270] In some embodiments, the Fab fragment conjugated to the LNP is an IgG4 Fab fragment. In some such embodiments, the cysteine at position 214 (Kabat numbering) of the light chain of the IgG4 Fab fragment is covalently to the LNP. In other such embodiments, the cysteine at position 127 (Kabat numbering) of the heavy chain of the IgG4 Fab fragment is covalently to the LNP. In other such embodiments, the cysteine at position 127 (Kabat numbering) of the heavy chain of the IgG4 Fab fragment and the cysteine at position 214 (Kabat numbering) of the light chain of the IgG4 Fab fragment is covalently to the LNP.

[0271] The processes described herein also enable the ability to conjugate two or more different Fab fragments to the surface of an LNP. An embodiment showing the conjugation of two different Fab fragments (which can bind to the same or different antigen) to the surface of an LNP is shown in **FIG. 31**. In **FIG. 31**, two Fab fragments (Fab1 and Fab2) are reduced (step (i)) and reacted (step (ii)) with a precursor LNP comprising a thiol-reactive group (e.g., maleimide or DBM). Following the reaction in step (ii), both Fab1 and Fab2 are conjugated to the surface of the LNP. Despite reduction of the disulfide bonds in Fab1 and Fab2, the heavy chain and light chain in Fab1 and the heavy and light chain in Fab2 remain together on the surface of the LNP. In other words, neither the heavy chain or light chain of Fab1 associate with the heavy or light chain of Fab2 on the surface of the LNP.

[0272] In some embodiments involving conjugating two Fab fragments (i.e., a first Fab fragment and a second Fab fragment) to the surface of an LNP, the first Fab fragment and the second Fab fragment can be reduced in the same reaction (e.g., the first and second Fab fragments are mixed in a reaction vessel and contacted with the same reducing reagent). In some embodiments, the first Fab fragment and the second Fab fragment are reduced separately (e.g., the first and second Fab fragments thereof are each contacted with a reducing reagent in separate reaction vessels). In some embodiments, the first Fab fragment is contacted with the reducing reagent prior to step (ii) (wherein the reduced first Fab fragment is conjugated to the LNP surface). In some embodiments, the second Fab fragment is contacted with the reducing

reagent after step (ii) (wherein the reduced second Fab fragment is conjugated to the LNP surface). In some embodiments, the reduced first Fab fragment and the reduced second Fab fragment are contacted with the LNP simultaneously. In some embodiments, the reduced first Fab fragment thereof and the reduced second Fab fragment are contacted with the LNP sequentially (in either order). It can be contemplated that any number of Fab fragments thereof can be implemented in the method or process (e.g., a third, fourth, fifth, etc. Fab fragment). In some embodiments, a total of three different Fab fragments can be conjugated to the surface of the LNP. In some embodiments, a total of four different Fab fragments can be conjugated to the surface of the LNP.

10 **[0273]** In some embodiments, the reaction between at least one of the two free cysteine residues of the Fab fragment and the thiol-reactive group of the LNP forms at least one covalent bond. In some embodiments, the formation of at least one covalent bond between at least one of the two free cysteine residues of the Fab fragment and the thiol-reactive group of the LNP is reversible. In some embodiments, the formation of at least one covalent bond between at least one of the two free cysteine residues of the Fab fragment and the thiol-reactive group of the LNP is irreversible. In some embodiments, the reaction efficiency between at least one of the two free cysteine residues of the Fab fragment and the thiol-reactive group of the LNP is greater than 5%, greater than 10%, greater than 25%, greater than 50%, greater than 60%, greater than 70%, greater than 80%, or greater than 90%. In some embodiments, the reaction efficiency between at least one of the two free cysteine residues of the Fab fragment and the thiol-reactive group of the LNP is from about 5% to about 30%, about 10% to about 20%, about 25% to about 50%, about 30% to about 40%, about 50% to about 80%, about 60% to about 70%, about 70% to about 95%, or about 80% to about 90%. In some embodiments, the conjugate product of the disclosed method can be purified from remaining intermediate product using any suitable technique such as, but not limited to, ultrafiltration and diafiltration.

[0274] In some embodiments, conjugates prepared by the method or process disclosed herein have a high density of the Fab fragment on the surface of the LNP. For instance, the conjugate can comprise a plurality of Fab fragments conjugated to the LNP surface. In some embodiments, the conjugate can comprise more than 10 Fab fragments per LNP. In some
30 embodiments, the conjugate can comprise more than 20 Fab fragments per LNP. In some embodiments, the conjugate can comprise more than 30 Fab fragments. In some embodiments, the conjugate can comprise more than 50 Fab fragments per LNP. In some embodiments, the

conjugate can comprise more than 75 Fab fragments per LNP. In some embodiments, the conjugate can comprise more than 100 Fab fragments. In some embodiments, the conjugate can comprise from about 50 to about 200 Fab fragments per LNP. In some embodiments, the conjugate can comprise from about 100 to about 200 Fab fragments per LNP. In some
5 embodiments, the conjugate can comprise from about 100 to about 230 Fab fragments per LNP. In some embodiments, the conjugate can comprise from about 10 to about 150 Fab fragments per LNP. In some embodiments, the conjugate can comprise from about 10 to about 30 Fab fragments per LNP.

IV. PAYLOADS

10 [0275] The LNPs and conjugates (targeted LNPs), of the disclosure can be used to deliver payloads to cells, e.g., T cells, HSCs, HSC progenitors, or liver cells. In some embodiments, the LNPs can be used to deliver a payload into long-term HSC (LT-HSC) subpopulations. In some embodiments, a targeted LNP as described herein can be used to deliver a payload to cells expressing cell-surface receptors targeted by the targeting moiety,
15 e.g., antibody, Fab fragment, scFv or nanobody component, of the targeted LNP (conjugate).

[0276] The LNPs and conjugates disclosed herein can be used to deliver a therapeutic of interest to a cell. For instance some embodiments, the conjugates disclosed herein can be used to deliver a nucleic acid (e.g., mRNA) encoding a vaccine. In other embodiments, the conjugates disclosed herein can be used to deliver a nucleic acid (e.g., mRNA) encoding an
20 enzyme. In other embodiments, the LNPs and conjugates disclosed herein can be used to deliver a nucleic acid encoding a chimeric antigen receptor (CAR) to T cells.

[0277] In some embodiments, the payload is one or more nucleic acids. In some embodiments, the payload is one or more RNA molecules. In some embodiments, the payload is an mRNA (e.g., an mRNA encoding an enzyme). In some embodiments, the RNA molecule
25 is a non-coding RNA (ncRNA). In some embodiments, the payload is an RNA template (for example, an RNA template for reverse transcription, e.g., Target Primed Reverse Transcription (TPRT)). In other embodiments the payload is a siRNA or a microRNA (miRNA). In other embodiments, the payload is a guide RNA. In other embodiments, the payload is a tRNA. In other embodiments, the payload is an antisense oligonucleotide (ASO). In other embodiments,
30 the payload is a DNA molecule, for example, a DNA plasmid, closed-ended DNA (ceDNA), or a small circular DNA (e.g., a minicircle or nanoplasmid). Nucleic acid payloads can be linear, circular, covalently closed, single-stranded, double-stranded, or hybrid RNA/DNA

molecules. other embodiments, the payload is a small molecule. In some embodiments, the payload is a peptide or protein. In some embodiments, the conjugates disclosed herein can include two or more payloads, for example, selected from mRNA, ncRNA, guide RNA, siRNA, miRNA, ASO, DNA vector, small molecule, peptide, or protein.

5 **[0278]** The LNPs described herein, e.g., targeted LNPs, can be used to deliver a therapeutic of interest to a cell, e.g., an HSC, HSC progenitor or liver cell. In some embodiments, the LNP, e.g., targeted LNP, contains a payload that is a therapeutic agent. In some embodiments, the therapeutic agent can be a therapeutic peptide or protein, a nucleic acid comprising a therapeutic agent, or a nucleic acid encoding a therapeutic agent. In some
10 embodiments, the therapeutic agent can be a genetic medicine (e.g., for gene therapy or gene editing), wherein the therapeutic agent is capable of modifying, altering or effecting a change in the genomic DNA of a cell (e.g., an HSC (e.g., LT-HSC), HSC progenitor, or liver cell in the subject). In some embodiments, the therapeutic agent is a gene therapy agent or gene editing agent. In some embodiments, the therapeutic agent is a gene modifying polypeptide,
15 as described herein. In some embodiments, the therapeutic agent is a gene modifying system, as described herein.

[0279] In some embodiments, the therapeutic agent can be a peptide or protein, such as an enzyme, or a nucleic acid (e.g., mRNA or DNA) encoding the peptide or protein (e.g., an enzyme). In some embodiments, the enzyme can be a nuclease, recombinase, integrase,
20 transposase, retrotransposase, helicase, transcriptase, polymerase, reverse transcriptase, deaminase, methylase, demethylase, or ligase, or can have a combination of enzymatic activities thereof. In some embodiments, the therapeutic agent can be a peptide or protein, or a nucleic acid encoding the peptide or protein, for use as a replacement gene therapy. In some
25 embodiments, the therapeutic agent can be a peptide or protein, or a nucleic acid encoding the peptide or protein, for use in modifying or altering the genome or epigenome of a cell, e.g., a T cell, an HSC, HSC progenitor, or liver cell of a subject. In some embodiments, the therapeutic agent can comprise one or more components of a system for modifying or altering the genome or epigenome of a cell, e.g., a T cell, an HSC, HSC progenitor, or liver cell of a
30 subject. In some embodiments, the system for modifying or altering the genome or epigenome of a cell of a subject comprises one or more proteins, one or more nucleic acids (e.g., RNA and/or DNA), or combinations thereof.

[0280] In some embodiments, the therapeutic agent can be one or more components of a ribonucleoprotein (RNP) complex for modifying or altering the genome or epigenome of a cell, e.g., a T cell, an HSC, HSC progenitor, or liver cell. For example, in some cases the therapeutic agent can be a protein, or a nucleic acid (e.g., mRNA) encoding the protein, and/or an RNA molecule (e.g., a gRNA) for guiding the protein to a particular location in the genome or epigenome, wherein the protein is capable of modifying or altering the genome or epigenome as a nuclease, recombinase, integrase, transposase, helicase, reverse transcriptase, deaminase, methylase, demethylase, or ligase, or combinations thereof. In certain embodiments, the therapeutic agent is a nuclease. In some embodiments, the nuclease cleaves DNA (e.g., both strands of the DNA), thereby introducing insertion and/or deletion (indel) mutations in DNA, e.g., genomic DNA. In certain embodiments, the nuclease is a nickase (i.e., it cleaves a single stand of DNA). In certain embodiments, the nuclease is mutated such that it is inactive or comprises reduced nuclease activity. In certain embodiments, the therapeutic agent can be used to introduce a substitution in DNA, e.g., genomic DNA, via base editing. In certain embodiments, the therapeutic agent can be used for epigenome editing. In certain embodiments, the therapeutic agent can be used to introduce an indel or a substitution in DNA, e.g., genomic DNA, by inducing target-primed reverse transcription (TPRT). In certain embodiments, the therapeutic agent (i.e., payload) delivered by the LNP, e.g., targeted LNP, can be a gene modifying protein, a nucleic acid encoding a gene modifying protein, or a gene modifying system, as described herein. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a gene modifying protein, a nucleic acid encoding a gene modifying protein, or a gene modifying system, as described herein, into HSPCs, HSCs (e.g., LT-HSCs), or liver cells. In certain embodiments, the therapeutic agent can be a small molecule. In certain embodiments, the therapeutic agent can be an siRNA or miRNA. In some embodiments, the LNPs, e.g., targeted LNPs, of the disclosure can be used to deliver gene editing components into T cells, HSCs, HSC progenitors, or liver cells. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas system into T cells or HSCs. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a Class 1 (type I, type III, or type IV) CRISPR system into T cells or HSCs. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a Class 2 (type II, type V, or type VI) CRISPR system into T cells or HSCs. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas9 system, or a nucleic acid encoding one or more components of the CRISPR-Cas9 system, into T cells or HSCs (e.g., LT-HSCs). In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas12 system (e.g., a Cas12a system),

or a nucleic acid encoding one or more components of the CRISPR-Cas12 system, into T cells or HSCs (e.g., LT-HSCs). In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas system into liver cells. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a Class 1 (type I, type III, or type IV) CRISPR system into liver cells. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a Class 2 (type II, type V, or type VI) CRISPR system into liver cells. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas9 system, or a nucleic acid encoding one or more components of the CRISPR-Cas9 system, into liver cells. In some embodiments, the LNPs, e.g., targeted LNPs, can be used to deliver a CRISPR-Cas12 system (e.g., a Cas12a system), or a nucleic acid encoding one or more components of the CRISPR-Cas12 system, into liver cells. In some such embodiments, the LNPs, e.g., targeted LNPs, can comprise two RNA molecules, such as an RNA comprising a guide RNA (gRNA) and an mRNA encoding the Cas protein. In some embodiments, the Cas is Cas9 or Cas12a. In some embodiments, the Cas is Cas9. In some embodiments, the Cas is Cas12a. In some embodiments, the gRNA is a single guide RNA (sgRNA). In some embodiments, the LNPs, e.g., targeted LNPs, comprise a payload consisting of or comprising a Cas9 or an mRNA encoding a Cas9. In some embodiments, the LNPs, e.g., targeted LNPs, comprise a payload consisting of or comprising a Cas12 (e.g., Cas12a) or an mRNA encoding a Cas12 (e.g., Cas12a). In some embodiments, the payload further consists of or comprises a gRNA.

[0281] In some embodiments, the mRNA component of a gene modifying system comprises a recombinant nuclease, or a nucleic acid encoding the nuclease, for example a restriction endonuclease, meganuclease, homing endonuclease, zinc finger nuclease (ZFN), or a transcription activator-like effector nuclease (TALEN).

[0282] In certain embodiments, the therapeutic agent (i.e., payload) delivered by the LNP, e.g., targeted LNP, can be a gene modifying system, or one or more nucleic acids (e.g., RNA molecules) comprising the gene modifying system (e.g., an mRNA encoding a gene modifying polypeptide and a template RNA), as described herein.

Gene modifying systems

[0283] In some embodiments, a gene modifying system described herein comprises:

(A) a gene modifying polypeptide or a nucleic acid encoding the gene modifying polypeptide, wherein the gene modifying polypeptide comprises (i) a reverse transcriptase domain, and an

endonuclease domain that contains DNA binding functionality; and (B) a template RNA. A gene modifying polypeptide, in some embodiments, acts as a substantially autonomous protein machine capable of integrating a template nucleic acid sequence into a target DNA molecule (e.g., in a mammalian host cell, such as a genomic DNA molecule in the host cell), substantially
5 without relying on host machinery. For example, the gene modifying protein may comprise a DNA-binding domain, a reverse transcriptase domain, and an endonuclease domain. In some embodiments, the DNA-binding function may involve an RNA component that directs the protein to a DNA sequence, e.g., a gRNA spacer. In some embodiments, the gene modifying polypeptide may comprise a reverse transcriptase domain and an endonuclease domain. The
10 RNA template element of a gene modifying system is typically heterologous to the gene modifying polypeptide element and provides an object sequence to be inserted (reverse transcribed) into the host genome. In some embodiments, the gene modifying polypeptide is capable of target primed reverse transcription.

[0284] In some embodiments, a gene modifying polypeptide includes one or more
15 domains that, collectively, facilitate 1) binding the template nucleic acid, 2) binding the target DNA molecule, and 3) facilitate integration of the at least a portion of the template nucleic acid into the target DNA. In some embodiments, the gene modifying polypeptide is an engineered polypeptide that comprises one or more amino acid substitutions to a corresponding naturally occurring sequence. In some embodiments, the gene modifying polypeptide comprises two or
20 more domains that are heterologous relative to each other, e.g., through a heterologous fusion (or other conjugate) of otherwise wild-type domains, or well as fusions of modified domains, e.g., by way of replacement or fusion of a heterologous sub-domain or other substituted domain. For instance, in some embodiments, one or more of: the RT domain is heterologous to the DBD; the DBD is heterologous to the endonuclease domain; or the RT domain is
25 heterologous to the endonuclease domain.

Polypeptide components of gene modifying systems

[0285] In some embodiments, the gene modifying polypeptide possesses the functions of DNA target site binding, template nucleic acid (e.g., RNA) binding, DNA target site cleavage, and template nucleic acid (e.g., RNA) writing, e.g., reverse transcription. In some
30 embodiments, each functions is contained within a distinct domain. In some embodiments, a function may be attributed to two or more domains (e.g., two or more domains, together, exhibit the functionality). In some embodiments, two or more domains may have the same or similar

function (e.g., two or more domains each independently have DNA-binding functionality, e.g., for two different DNA sequences). In other embodiments, one or more domains may be capable of enabling one or more functions, e.g., a Cas9 domain enabling both DNA binding and target site cleavage. In some embodiments, the domains are all located within a single polypeptide. In some aspects, a gene modifying polypeptide described herein comprises (e.g., a system described herein comprises a gene modifying polypeptide that comprises): 1) a Cas domain (e.g., a Cas nickase domain, e.g., a Cas9 nickase domain); 2) a reverse transcriptase (RT) domain, wherein the RT domain is C-terminal of the Cas domain; and a linker disposed between the RT domain and the Cas domain. In some embodiments, the endonuclease domain is N-terminal relative to the RT domain. In some embodiments, the endonuclease domain is C-terminal relative to the RT domain.

[0286] In some embodiments, the gene modifying polypeptide comprises a GG amino acid sequence between the Cas domain and the linker, an AG amino acid sequence between the RT domain and the second NLS, and/or a GG amino acid sequence between the linker and the RT domain. In some embodiments, the gene modifying polypeptide comprises a sequence of SEQ ID NO: 21 which comprises the first NLS and the Cas domain, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto. In some embodiments, the gene modifying polypeptide comprises a sequence of SEQ ID NO: 22 which comprises the second NLS, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto.

Exemplary N-terminal NLS-Cas9 domain

MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK
 NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLLEE
 SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH
 25 MIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
 SRRLLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDL
 DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
 LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
 KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPY
 30 YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEK
 VLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK
 QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVL
 TLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
 TILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICK
 35 GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELG
 SQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
 DDSIDNKVLTRSDKARGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA

ERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSK
 LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYD
 VRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
 GRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG
 5 GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEV
 KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG
 SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQ
 AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQL
 GGDGG (SEQ ID NO: 21)

10

Exemplary C-terminal sequence comprising an NLS

AGKRTADGSEFEKRTADGSEFESPKKKAKVE (SEQ ID NO: 22)

[0287] In some embodiments, the gene modifying polypeptide comprises a template
 15 nucleic acid binding domain (e.g., RNA binding domain) is contained within the target DNA
 binding domain. For example, in some embodiments, the DNA binding domain is a CRISPR-
 associated protein that recognizes the structure of a template nucleic acid (e.g., template RNA)
 comprising a gRNA. In some embodiments, a gene modifying polypeptide comprises a DNA-
 binding domain comprising a CRISPR-associated protein that associates with a gRNA scaffold
 20 that allows the DNA-binding domain to bind a target genomic DNA sequence. In some
 embodiments, the gRNA scaffold and gRNA spacer is comprised within the template nucleic
 acid (e.g., template RNA), thus the DNA-binding domain is also the template nucleic acid
 binding domain. In some embodiments, the polypeptide possesses RNA binding function in
 multiple domains, e.g., can bind a gRNA structure in a CRISPR-associated DNA binding
 25 domain and an additional sequence, or structure in a reverse transcriptase domain.

Writing domain (RT Domain)

[0288] In certain aspects of the present invention, the writing domain of the gene
 modifying system possesses reverse transcriptase activity and is also referred to as a reverse
 30 transcriptase domain (a RT domain). In some embodiments, the RT domain comprises an RT
 catalytic portion and RNA-binding region (e.g., a region that binds the template RNA).

[0289] In some embodiments, the RT domain forms a dimer (e.g., a heterodimer or
 homodimer). In some embodiments, the RT domain is monomeric.

[0290] In certain embodiments, a gene modifying polypeptide comprises the amino
 35 acid sequence of an RT domain sequence from a family selected from: AVIRE, BAEVM, FFV,
 FLV, FOAMV, GALV, KORV, MLVAV, MLVBM, MLVCB, MLVFF, MLVMS, PERV,

SFV1, SFV3L, WMSV, XMRV6, BLVAU, BLVJ, HTL1A, HTL1C, HTL1L, HTL32, HTL3P, HTLV2, JSRV, MLVF5, MLVRD, MMTVB, MPMV, SFVCP, SMRVH, SRV1, SRV2, and WDSV.

[0291] In some embodiments, a gene modifying polypeptide described herein
5 comprises an RT domain having an amino acid sequence according to Table 6 of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety), or a sequence having at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity thereto. In some embodiments, a nucleic acid described herein encodes an RT domain having an amino acid sequence according to Table 6 of International Application WO/2023/039440, or a
10 sequence having at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity thereto.

[0292] In some embodiments, an RT domain (e.g., as listed in Table 6 of International Application WO/2023/039440) comprises one or more mutations as listed in Table 2A of International Application WO/2023/039440, which Table is herein incorporated by reference in its entirety. In some embodiment, an RT domain as listed in Table 6 of International
15 Application WO/2023/039440 comprises one, two, three, four, five, or six of the mutations listed in the corresponding row of Table 2A of International Application WO/2023/039440.

[0293] In some embodiments, a gene modifying polypeptide comprises: (i) a linker comprising a linker sequence as listed in a row of Table T1, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto; and (ii) an RT domain
20 comprising an RT domain sequence as listed in the same row of Table 2, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto. In some embodiments, a gene modifying polypeptide comprises an amino acid sequence according to Table 3, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto.

25

Table 2. Selection of exemplary gene modifying polypeptides

SEQ ID NO: for Full Polypeptide Sequence	Linker Sequence	RT name and sequence

23	AEAAAKEAAA KEAAAKEAAA KALEAEAAK EAAAKEAAK EAAKA (SEQ ID NO: 34)	AVIRE_P03360_3mutA TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEIN PPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLR ETIRKFRAAGILRPVHSPWNTPLLVRKSGTSEYRMV QDLREVNKRVEDIHPTVPNPYTLLSLLPPDRIWYSVL DLKDFAFFCIPLAPESQLIFAFEWADAEEGESGQLTWT RLPQGFKNSPTLTFNEALNRDLQGFRDLHPSVSLQYV DDLIIAADTQAACLSATRDLLMTLAEGLYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTK RQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDP LVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQFLV EETSGAAKGVLTQALGPWKRVPVAYLSKRLDPVAA WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESL LRSPDPKWL TNARITQYQVLLDPPRVRFKQTAALNP ATLLPETDDTLPIHHCLDTLDSLSTRPDLTDQPLAQA EATLFTDGSSYIRDGKRYAGAAVVTLDSEVIWAEPLPI GTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATL HVGMIYRERGWLTAGGKAIKNAPEILALLTAVWLP KRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIR PLSTQATIS
24	AEAAAKEAAA KEAAAKEAAA KALEAEAAK EAAAKEAAK EAAKA (SEQ ID NO: 34)	FLV_P10273_3mutA TLQLEEEYRLFEPSTQKQEMDIWLKNFPQAWAETG GMGTAHCQAPVLIQLKATATPISIRQYPMPEAYQGI KPHIRRLDQGILKPCQSPWNTPLLVPKKPGTEDYRP VQDLREVNKRVEDIHPTVPNPYNLLSTLPPSHPWYTV LDLKDAFFCLRLHSESQLLFAFEWRDPEIGLSGQLTW TRLPQGFKNSPTLTFNEALHSDLADFRVRYPALVLLQY VDDLIIAAATRTECLEGTKALLETLGNGGYRASAKK AQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPK NSRQVREFLGKAGYCRLFIPGFAELAPLYPLTRPGT LFQWGTEQQLAFEDIKALLSSPALGLPDITKPFELFI DENSNGFAKGVLVQKLGWKRVPVAYLSKKLDTVASG WPPCLRMVAAAILVKDAGKLTGQPLTILTSHPVEA LVRQPPNKWLSNARMTHYQAMLLDAERVHFGPTVS LNPATLLPLPSGGNHHDCILAEETHGTRPDLTDQPL PDADLTWYTDGSSFIRNGEREAGAAVTESEVIWAA PLPPGTSAQRAELIALTQALKMAEGKKLTVYTDSRY AFATTHVHGEIYRRRGWLTSEGKEIKNKNEILALLEA LFLPKRLSIHCPGHQKGDSPQAKGNRLADDTAKKA ATETHSSLTVLP
25	AEAAAKEAAA KEAAAKEAAA KALEAEAAK EAAAKEAAK EAAKA (SEQ ID NO: 34)	MLVMS_P03355_3mutA_WS TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETG GMGLAVRQAPLIPLKATSTPVSIKQYPMSQEARLGIK PHIQRLLDQGILVPCQSPWNTPLLVPKKPGTNDYRPV QDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTV LDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTW TRLPQGFKNSPTLTFNEALHRDLADFRIQHPDLILLQY VDDLIIAATSELDCQQGTRALLQTLGNLGYRASAKK AQICQKQVKYLGILLKEGQRWLTEARKETVMGQPT PKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKP

		GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPF ELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDP VAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQ FGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPD LTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTET EVIWAKALPAGTSAQRAELIALTQALKMAEGKKLV YTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEI LALLKALFLPKRLSIHCPGHQKGHSAEARGNRMAD QAARKAAITETPDTSTLL
26	AEAAAKEAAA KEAAAKEAAA KALEAEAAK EAAAKEAAK EAAKA (SEQ ID NO: 34)	SFV3L_P27401_2mutA MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPQAF LEEEVPIKNIWIKTIHGEKEQPYYLTFKIQGRKVEAE VISSPYDYILVSPSDIPWLMKKPLQLTTLVPLQEYEEER LLKQTMLTGSYKEKLQSLFLKYDALWQHWENQVGH RRIKPHHIATGTVNPRPQKQYPINPKAKASIQTVINDL LKQGVLIQQNSIMNTPVYPVPKPDGKWRMVLDYRE VNKTIPLIAAQNHSAAGILSSIFRGKYKTTLDSLNGFW AHSITPESYWLTAFTWLGGQYCWTRLPGQFLNSPAL FNADVVDLLKEVPNVQVYVDDIYISHDDPREHLEQL EKVFSLLLNAGYVVSLLKSEIAQHEVEFLGFNITKEG RGLTETFKQKLLNITPPRDLKQLQSILGKLNFAFNFI NFSELVKPLYNIIATAPGKYITWTTDNSQQQLQNIISML NSAENLEERNPEVRLIMKVNTSPSAGYIRFYNEFAKR PIMYLNYYVYTKAEVKFTNTEKLLTTIHKGLIKALDLG MGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMS YLEDPRIQFHYDKTLPELQQVPTVTDDIIAKIKHPSEF SMVFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFT VINTWSIPLGDHTAQLAEVAAVEFACKKALKIDGPVL IVTDSFYVAESVNKELPYWQSNQFFNNKKKPLKHVS KWKSIADCIQLKPDIIIIHEKGHQPTASTFHTEGNLA DKLATQGSYVVN

Table 3. Full length amino acid sequence for certain drivers including those of Table 2

SEQ ID NO:	Amino acid sequence
27	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSITKADLRLLI YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAK AII SARLSKSRRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKL QLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDK GASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY

	HDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTTRSDKARGKSDN VPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIN NYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMS PQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKEGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAGYKEVKKDLIIKLP KYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKSGPEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVI LADANLDKVL SAYNKH RDKP IREQAENI IHLEF TLNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ LGGDGGAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGT APLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTA LPVRVRQYPITILEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRM VQDLREVNKRVE TIHPTVPNPYTLLSLLPPDRIWYSVLDLKD AFFC I PLAPESQL I FAFEWADAE EGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSL LQ YVDDLIIAADTQAACLSATRDLLMTLAE LGYRVSGKKAQLCQEEV TYLGFKI HKG SRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQFLYAATRPG NDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVL TQAL GPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNL ESLLRSPDPKWL TNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHH CLDTLDSL TSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLD SVIWA EPLPIGTS AQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGM IYRERGWL TAGGKAIKNAPEILALLTAVWLPKRVA VMHCKGHQKDDAPTSTGNRRADEVAREV AIRPLSTQATISAGKRTADGSEFEKRTADGSEFESPKKKAKVE
28	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHS IK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF FH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LI YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENP INASGVDAK AII SARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL QLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFWMTRKSEETITPWNFE EVVDK GASAQSFIERMTNFDKNLPNEKVL PKHSLLEYEFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT Y HDLLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTTRSDKARGKSDN VPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIN NYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMS

	PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDIIKLP KYSLFELENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVI LADANLDKVL SAYNKH RD KPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ LGGDGGAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGT LQLEEEYRLFEP ESTQKQEMDIWLKNFPQAWAETGGMGTAHCQAPVLIQLKATAT PISIRQYPMPEAYQGIKPHIRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPV QDLREVNKRVEDIHPTVNPYNLLSTLPPSHPWYTVL DLKDAFFCLRLHSESQLL FAFEWRDPEIGLSGQLTWTRLPQGFKNSPTLFNEALHSDLADFRVRYPALVLLQY VDDL LLAATRTECLEG TKALLET LGNKGYRASAKKAQICLQEV TYLGYS LKDGQ RWLT KARKEAILSIPVPKNSRQVREFLGAGYCRLFIPGFAELAAPLYPLTRPGT LFQWGTEQQ LAFEDIKKALLSSPALGLPDITKPFELFIDENSGFAKGV LVQKLGP WKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKDAGKLT LGQPLTILTSHPVEA LVRQPPNKWLSNARMTHYQAMLLDAERVHFGPTVSLNPATLLPLPSGGNHHDCLQ ILAE THGTRPDLTDQPLPDADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPL PPGTSAQRAELIALTQALKMAEGK KLT VYTDSRYAFATTHVHGEIYRRRGWLTSE GKEIKNKNEI LALLEALFLPKRLSI IHCPGHQKGDSPQAKGNRLADDTAKKAATE THSSLTVLPAGKRTADGSEFEKRTADGSEFESPKKKAKVE
29	MPAAKRVKLDGGDKKYSIGLDIGTNSVGAVITDEYKVPSKKFKVLGNTRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADRLI YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAK AII SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL QLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK GASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDN VPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREIN NYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDIIKLP KYSLFELENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVI LADANLDKVL SAYNKH RD KPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ LGGDGGAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGT LNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLI IPLKATS TPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRP VQDLREVNKRVEDIHPTVNPYNLLSGLPPSHQWYTVL DLKDAFFCLRLHPTSQP

	LFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQ YVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLKEG QRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIGFAEMAAPLYPLTKPG TLFNWGPDQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVE ALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETETEWAKA LPAGTSAQRAELIALTQALKMAEGKKLNVYTDSTRYAFATAHIGEIYRRRGWLTS EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLAGKRTADGSEFEKRTADGSEFESPKKKAKVE
30	MPAARKVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TKADLR LI YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAK AILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL QLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLKIIKDKDFLDNEENEDILEDIVLTTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQM LIHDDSLTF KEDIQKAQVSGQGDLSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDN VPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIN NYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQF IGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SM PQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQ LGGDGGAEEAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGM DPLQLLQPLEAEIKGTKLKAHWNSGATITCVPQAFLEEEVPIKNIWIKTIHGEKE QPVEYLT FKIQGRKVEAEVISSPYDYILVSPSDIPWLMKKPLQLTTLVPLQEYEE RLLKQTMLTGSYKEKLQSLFLKYDALWQHWHENQVGHRIKPHHIATGTVNPRPQK QYPINPKAKASIQTVINDLLKQGVLIQQNSIMNTPVYPVPKPDGKWRMVL DYREV NKTIPLIAAQNQHSAGILSSI FRGKYKTTLDLSNGFWAHSITPESYWLTAFTWLG QQYCWTRLPQGFLNSPALFNADVVDLLKEVPNVQVYVDDIYISHDDPREHLEQLE KVFSLLL NAGYVVS LKKSEIAQHEVEFLGFNITKEGRGLTETFQKLLNITPPRD LKQLQSILGKLN FARNFIPNFSELVKPLYNIIATAPGKYITWTTDNSQQQLQNIIS MLNSAENLEERNPEVRLIMKVNTSPSAGYIRFYNEFAKRPIMYLNLYVYTKAEVKF TNTEKLLTTIHKGLIKALDLGMGQEILVYSPIVSMTKIQKTPLPERKALPIRWIT WMSYLEDPRIQFHYDKTLPELQQVPTVTDDIIAKIKHPSEFSMVFYTDGSAIKHP NVNKS HNAGMGIAQVQFKPEFTVINTWSIPLGDHTAQLAEVA AVEFACKKALKID

	GPVLIVTDSFYVAESVNKELPYWQSNNGFFNNKKKPLKHVSKWKS IADCIQLKPDII I IHEKGHQPTASTFHTEGNNLADKLATQGSYVVNAGKRTADGSEFEKRTADGSEFESPKKKAKVE
31	MPAAKRVKLDGGSDLVLGLDIGIGSVGVGILNKVTGEI IHKNSRI FPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHGRGISYLLDDASDDGNSSVGDYQIVKENSQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFP TSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYHGPNGEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPD EFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQI INYVKNKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLT LNTEREGIQEAL EHEFADGSFSQKQVDELVQFRKANSSI FGKGWHNFSVKLMME LIPELYETSEEQMTILTRLGKQKTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISI HDLINNSNQFEVDHILPLSITFDDSLANKVLVYATAAQEKQORTPYQALDSMDDAWSFRELKAFVRESKTL SNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVS VVRGQFTSQLRRHWGIEKTRDTYHHHAVDALI IAASSQLNLWKKQKN TLVSYS EDQLLDIETGELISDDEYKESVFKAPYQH FVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQEKYNDIKKKEGVDS DSEFKFTLYKNDLLLVKDTETKEQQ LFRFLSRTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSQCKKGLGKSNISIYKVRTDVLGNQHI IKNEGDKPKLDFGGWQAAESYE VGGTAPLEEEYRLFLEAPIQNVTLL EQWKREIPKVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKD AFFCIPLAPESQLIFAFEWADAE EGESGQLTWTRL PQGFKNSPTLFNEALNRDLQGFRLDHPSVSL LQYVDDLLIAADTQAACLSATRDLLMTLAE LGYRVSGKKAQLCQEEV TYLGFKIHKGSRSLNSRTQAILQIPVPKTKRQVREFLGKIGYCR LFIPGFAELAQPLYAATRP GNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVL TQALGPWKRPVAYLSKR LDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNL ESLLRSPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHC LDTLDSLSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLD SVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTD SRYAFATLHVHGM IYRERGWLTAGGKA IKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFESPKKKAKVE
32	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEI FSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLF IQLVQTYNQ LFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDL DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY

HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDLVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDN VPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIN NYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLMS PQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDEQ KQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENI IHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQ LGGDGGAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGT APLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTA LPVRVRQYPITILEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRM VQDLREVNKRVETIHPTVPNPYTLSSLPPDRIWYSVLDLKD AFFCIPLAPESQL IFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQ YVDDLLIAADTQAACLSATRDLLMTLAE LGYRVSGKKAQLCQEEV TYLGFKIHKG SRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPG NDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVL TQAL GPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNL ESLLRSPDPKWL TNARI TQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHH CLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLD S VIWA EPLPIGTSQAQKALIALTKALEWSKDKSVNIYTDSRYAFATLHVHGM IYRERGWL TAGGKA IKN APEILALLTAVWLPKRVA VMHCKGHQKDDAPTSTGNRRADEVAREV AIRPLSTQATISAGKRTADGSEFEKRTADGSEFESPKKKAKVE
--

Endonuclease domains and DNA binding domains

[0294] In some embodiments, a gene modifying polypeptide possesses the function of

5 DNA target site cleavage via an endonuclease domain. In some embodiments, a gene modifying polypeptide comprises a DNA binding domain, e.g., for binding to a target nucleic acid. In some embodiments, a domain (e.g., a Cas domain) of the gene modifying polypeptide comprises two or more smaller domains, e.g., a DNA binding domain and an endonuclease domain. It is understood that when a DNA binding domain (e.g., a Cas domain) is said to bind

10 to a target nucleic acid sequence, in some embodiments, the binding is mediated by a gRNA.

[0295] In some embodiments, a domain has two functions. For example, in some embodiments, the endonuclease domain is also a DNA-binding domain. In some embodiments, the endonuclease domain is also a template nucleic acid (e.g., template RNA) binding domain. For example, in some embodiments, a polypeptide comprises a CRISPR-associated

15 endonuclease domain that binds a template RNA comprising a gRNA, binds a target DNA

sequence (e.g., with complementarity to a portion of the gRNA), and cuts the target DNA sequence. In some embodiments, an endonuclease domain or endonuclease/DNA-binding domain from a heterologous source can be used or can be modified (e.g., by insertion, deletion, or substitution of one or more residues) in a gene modifying system described herein.

5 **[0296]** In some embodiments, the gene modifying polypeptide comprises a Cas domain, and the Cas domain comprises an endonuclease domain and a DNA-binding domain. In some embodiments, the Cas domain comprises a Cas9 or a mutant or variant thereof (e.g., as described herein). In embodiments, the Cas domain is associated with a guide RNA (gRNA), e.g., as described herein. In embodiments, the Cas domain is directed to a target
10 nucleic acid (e.g., DNA) sequence of interest by the gRNA.

[0297] In some embodiments, the endonuclease domain has nickase activity and cleaves one strand of a target DNA. In some embodiments, nickase activity reduces the formation of double-stranded breaks at the target site relative to an endonuclease domain with double strand break activity. In certain embodiments, the heterologous endonuclease is derived
15 from a CRISPR-associated protein, e.g., Cas9. In certain embodiments, the heterologous endonuclease is engineered to have only nickase activity, e.g., be a Cas9 nickase, e.g., SpCas9 with D10A, H840A, or N863A mutations. In some embodiments, the endonuclease domain has nickase activity and does not form double-stranded breaks. In some embodiments, the endonuclease domain forms single-stranded breaks at a higher frequency than double-stranded
20 breaks, e.g., at least 90%, 95%, 96%, 97%, 98%, or 99% of the breaks are single-stranded breaks, or less than 10%, 5%, 4%, 3%, 2%, or 1% of the breaks are double-stranded breaks. In some embodiments, the endonuclease forms substantially no double-stranded breaks. In some embodiments, the endonuclease does not form detectable levels of double-stranded breaks.

[0298] In some embodiments, the endonuclease domain has nickase activity that nicks
25 the target site DNA of the first strand; e.g., in some embodiments, the endonuclease domain cuts the genomic DNA of the target site near to the site of alteration on the strand that will be extended by the writing domain. In some embodiments, the endonuclease domain has nickase activity that nicks the target site DNA of the first strand and does not nick the target site DNA of the second strand. For example, when a polypeptide comprises a CRISPR-associated
30 endonuclease domain having nickase activity, in some embodiments, said CRISPR-associated endonuclease domain nicks the target site DNA strand containing the PAM site (e.g., and does not nick the target site DNA strand that does not contain the PAM site). As a further example,

when a polypeptide comprises a CRISPR-associated endonuclease domain having nickase activity, in some embodiments, said CRISPR-associated endonuclease domain nicks the target site DNA strand not containing the PAM site (e.g., and does not nick the target site DNA strand that contains the PAM site).

5 [0299] In some embodiments, a gene modifying polypeptide described herein comprises a Cas domain. In some embodiments, the Cas domain can direct the gene modifying polypeptide to a target site specified by a gRNA spacer, thereby modifying a target nucleic acid sequence in “cis.” In some embodiments, a reverse transcriptase domain of a gene modifying
10 polypeptide is fused to a Cas domain. CRISPR systems are adaptive defense systems originally discovered in bacteria and archaea. CRISPR systems use RNA-guided nucleases termed CRISPR-associated or “Cas” endonucleases (e.g., Cas9 or Cpf1) to cleave foreign DNA. For example, in a typical CRISPR-Cas system, an endonuclease is directed to a target nucleotide sequence (e. g., a site in the genome that is to be sequence-edited) by sequence-specific, non-coding “guide RNAs” that target single- or double-stranded DNA sequences. Three classes (I-
15 III) of CRISPR systems have been identified. The class II CRISPR systems use a single Cas endonuclease (rather than multiple Cas proteins). One class II CRISPR system includes a type II Cas endonuclease such as Cas9, a CRISPR RNA (“crRNA”), and a trans-activating crRNA (“tracrRNA”). The crRNA contains a “spacer” sequence, a typically about 20-nucleotide RNA sequence that corresponds to a target DNA sequence (“protospacer”). In the wild-type system,
20 and in some engineered systems, crRNA also contains a region that binds to the tracrRNA to form a partially double-stranded structure that is cleaved by RNase III, resulting in a crRNA/tracrRNA hybrid molecule. A crRNA/tracrRNA hybrid then directs the Cas endonuclease to recognize and cleave a target DNA sequence. A target DNA sequence is generally adjacent to a “protospacer adjacent motif” (“PAM”) that is specific for a given Cas
25 endonuclease and required for cleavage activity at a target site matching the spacer of the crRNA. CRISPR endonucleases identified from various prokaryotic species have characteristic PAM sequence requirements; examples of PAM sequences include 5′-NGG (*Streptococcus pyogenes*), 5′-NNAGAA (*Streptococcus thermophilus* CRISPR1), 5′-NGGNG (*Streptococcus thermophilus* CRISPR3), and 5′-NNNGATT (*Neisseria meningitidis*). Some
30 endonucleases, e.g., Cas9 endonucleases, are associated with G-rich PAM sites, e.g., 5′-NGG), and perform blunt-end cleaving of the target DNA at a location 3 nucleotides upstream from (5′ from) the PAM site. Another class II CRISPR system includes the type V endonuclease Cpf1, which is smaller than Cas9; examples include AsCpf1 (from *Acidaminococcus* sp.) and

LbCpf1 (from *Lachnospiraceae* sp.). Cpf1-associated CRISPR arrays are processed into mature crRNAs without the requirement of a tracrRNA; in other words, a Cpf1 system, in some embodiments, comprises only Cpf1 nuclease and a crRNA to cleave a target DNA sequence. Cpf1 endonucleases, are typically associated with T-rich PAM sites, e. g., 5'-TTN. Cpf1 can also recognize a 5'-CTA PAM motif. Cpf1 typically cleaves a target DNA by introducing an offset or staggered double-strand break with a 4- or 5-nucleotide 5' overhang, for example, cleaving a target DNA with a 5-nucleotide offset or staggered cut located 18 nucleotides downstream from (3' from) from a PAM site on the coding strand and 23 nucleotides downstream from the PAM site on the complementary strand; the 5-nucleotide overhang that results from such offset cleavage allows more precise genome editing by DNA insertion by homologous recombination than by insertion at blunt-end cleaved DNA. See, e.g., Zetsche et al. (2015) *Cell*, 163:759 – 771.

[0300] A variety of CRISPR associated (Cas) genes or proteins can be used in the technologies provided by the present disclosure and the choice of Cas protein will depend upon the particular conditions of the method. Specific examples of Cas proteins include class II systems including Cas1, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cpf1, C2C1, or C2C3. In some embodiments, a Cas protein, e.g., a Cas9 protein, may be from any of a variety of prokaryotic species. In some embodiments a particular Cas protein, e.g., a particular Cas9 protein, is selected to recognize a particular protospacer-adjacent motif (PAM) sequence. In some embodiments, a DNA-binding domain or endonuclease domain includes a sequence targeting polypeptide, such as a Cas protein, e.g., Cas9. In certain embodiments a Cas protein, e.g., a Cas9 protein, may be obtained from a bacteria or archaea or synthesized using known methods. In certain embodiments, a Cas protein may be from a gram-positive bacteria or a gram-negative bacteria. In certain embodiments, a Cas protein may be from a *Streptococcus* (e.g., a *S. pyogenes*, or a *S. thermophilus*), a *Francisella* (e.g., an *F. novicida*), a *Staphylococcus* (e.g., an *S. aureus*), an *Acidaminococcus* (e.g., an *Acidaminococcus* sp. BV3L6), a *Neisseria* (e.g., an *N. meningitidis*), a *Cryptococcus*, a *Corynebacterium*, a *Haemophilus*, a *Eubacterium*, a *Pasteurella*, a *Prevotella*, a *Veillonella*, or a *Marinobacter*.

[0301] In some embodiments, a gene modifying polypeptide may comprise a Cas domain as listed in Table 7 or 8 of International Application WO/2023/039440 (which Tables are incorporated herein by reference in their entireties), or a functional fragment thereof, or a

sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% identity thereto.

[0302] In some embodiments, a gene modifying polypeptide comprises a Cas domain according to SEQ ID NO: 33, or a functional fragment thereof, or a sequence having at least 5 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% identity thereto.

[0303] DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG
ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLV
EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF
RGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
10 ENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLL
AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKA
LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR
EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPL
ARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHK
15 SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED
YFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFE
DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
KSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV
KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE
20 HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
KVLTRSDKARGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
KDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
25 ATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIH
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD
30 (SEQ ID NO: 33)

[0304] In some embodiments, a Cas protein requires a protospacer adjacent motif (PAM) to be present in or adjacent to a target DNA sequence for the Cas protein to bind and/or

function. In some embodiments, the PAM is or comprises, from 5' to 3', NGG, YG, NNGRRT, NNNRRT, NGA, TYCV, TATV, NTTN, or NNGGATT, where N stands for any nucleotide, Y stands for C or T, R stands for A or G, and V stands for A or C or G. In some embodiments, a Cas protein is a protein listed in Table 7 or 8 of International Application WO/2023/039440.

- 5 In some embodiments, a Cas protein comprises one or more mutations altering its PAM. Exemplary advances in the engineering of Cas enzymes to recognize altered PAM sequences are reviewed in Collias et al Nature Communications 12:555 (2021), incorporated herein by reference in its entirety.

- [0305] In some embodiments, the Cas protein is modified to deactivate or partially
10 deactivate the nuclease, e.g., nuclease-deficient Cas9. Whereas wild-type Cas9 generates double-strand breaks (DSBs) at specific DNA sequences targeted by a gRNA, a number of CRISPR endonucleases having modified functionalities are available, for example: a “nickase” version of Cas9 that has been partially deactivated generates only a single-strand break; a catalytically inactive Cas9 (“dCas9”) does not cut target DNA. In some embodiments, a DNA-
15 binding domain comprises a catalytically inactive Cas9, e.g., dCas9. Many catalytically inactive Cas9 proteins are known in the art. In some embodiments, dCas9 comprises mutations in each endonuclease domain of the Cas protein, e.g., D10A and H840A or N863A mutations. In some embodiments, a catalytically inactive or partially inactive CRISPR/Cas domain comprises a Cas protein comprising one or more mutations, e.g., one or more of the mutations
20 listed in Table 7 of International Application WO/2023/039440. In some embodiments, a Cas protein described on a given row of Table 7 of International Application WO/2023/039440 comprises one, two, three, or all of the mutations listed in the same row of Table 7 of International Application WO/2023/039440. In some embodiments, a Cas protein, e.g., not described in Table 7 of International Application WO/2023/039440, comprises one, two, three,
25 or all of the mutations listed in a row of Table 7 of International Application WO/2023/039440 or a corresponding mutation at a corresponding site in that Cas protein.

- [0306] In some embodiments, an endonuclease domain or DNA binding domain comprises a *Streptococcus pyogenes* Cas9 (SpCas9 or SpyCas9) or a functional fragment or variant thereof. In some embodiments, the endonuclease domain or DNA binding domain
30 comprises a modified SpCas9. In embodiments, the modified SpCas9 comprises a modification that alters protospacer-adjacent motif (PAM) specificity. In embodiments, the PAM has specificity for the nucleic acid sequence 5'-NGT-3'. In some embodiments, the

endonuclease domain or DNA binding domain comprises a Cas domain, e.g., a Cas9 domain. In embodiments, the endonuclease domain or DNA binding domain comprises a nuclease-active Cas domain, a Cas nickase (nCas) domain, or a nuclease-inactive Cas (dCas) domain. In embodiments, the endonuclease domain or DNA binding domain comprises a nuclease-active

5 Cas9 domain, a Cas9 nickase (nCas9) domain, or a nuclease-inactive Cas9 (dCas9) domain. In some embodiments, the endonuclease domain or DNA binding domain comprises a Cas9 domain of Cas9 (e.g., dCas9 and nCas9), Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, or Cas12i. In some embodiments, the endonuclease domain or DNA binding domain comprises a Cas9 (e.g., dCas9 and nCas9),

10 Cas12a/Cpf1, Cas12b/C2cl, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, or Cas12i. In some embodiments, the endonuclease domain or DNA binding domain comprises an *S. pyogenes* or an *S. thermophilus* Cas9, or a functional fragment thereof. In some embodiments, the endonuclease domain or DNA binding domain comprises a Cas9 sequence, e.g., as described in Chylinski, Rhun, and Charpentier (2013) RNA Biology 10:5, 726-737;

15 incorporated herein by reference.

[0307] In some embodiments, the endonuclease domain or DNA binding domain comprises spCas9, SpCas9-SpRY, spCas9-VRQR, spCas9-VRER, xCas9 (sp), saCas9, saCas9-KKH, spCas9-MQKSER, spCas9-LRKIQK, or spCas9-LRVSQL.

Additional components of exemplary Gene Modifying Polypeptides

20 [0308] In addition to the writing/RT domain and endonuclease and DNA binding domains, the exemplary gene modifying polypeptides for use in the gene modifying systems may comprise additional components as described herein.

Linkers

25 [0309] In some embodiments, a gene modifying polypeptide may comprise a linker, e.g., a peptide linker, e.g., a linker as described in Table 10 of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety). In some embodiments, a gene modifying polypeptide comprises, in an N-terminal to C-terminal

30 direction, a Cas domain (e.g., a Cas domain of Table 8 of International Application WO/2023/039440), a linker of Table 10 of International Application WO/2023/039440 (or a sequence having at least 70%, 80%, 85%, 90%, 95%, or 99% identity thereto), and an RT domain (e.g., an RT domain of Table 6 of International Application WO/2023/039440). In

some embodiments, a gene modifying polypeptide comprises a flexible linker between the endonuclease and the RT domain, e.g., a linker comprising the amino acid sequence AEAAAKEAAAKEAAAKEAAKALEAEAAAKEAAAKEAAAKEAAKA (SEQ ID NO: 34), or a sequence having at least 70%, 80%, 85%, 90%, 95%, or 99% identity thereto. In some embodiments, an RT domain of a gene modifying polypeptide may be located C-terminal to the endonuclease domain. In some embodiments, an RT domain of a gene modifying polypeptide may be located N-terminal to the endonuclease domain.

[0310] In some embodiments, a gene modifying polypeptide comprises: (i) a Cas domain comprising a Cas sequence as listed in a row of Table 4 of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety), or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto; (ii) a linker comprising a linker sequence as listed in the same row of Table 4, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto; and (iii) an RT domain comprising an RT domain sequence as listed in the same row of Table 4, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto.

Localization sequences

[0311] In certain embodiments, a gene modifying system RNA further comprises an intracellular localization sequence, e.g., a nuclear localization sequence (NLS). In some embodiments, a gene modifying polypeptide or a nucleic acid (e.g., RNA) encoding a gene modifying polypeptide comprises an NLS.

[0312] In some embodiments, a polypeptide described herein comprises one or more (e.g., 2, 3, 4, 5) nuclear targeting sequences, for example a nuclear localization sequence (NLS). In some embodiments, the NLS is a bipartite NLS. In some embodiments, an NLS facilitates the import of a protein comprising an NLS into the cell nucleus. In some embodiments, the NLS is fused to the N-terminus of a gene modifying polypeptide as described herein. In some embodiments, the NLS is fused to the C-terminus of the gene modifying polypeptide. In some embodiments, the NLS is fused to the N-terminus or the C-terminus of a Cas domain. In some embodiments, a linker sequence is disposed between the NLS and the neighboring domain of the gene modifying polypeptide.

[0313] In some embodiments, an NLS comprises an amino acid sequence as disclosed in Table 11 of International Application WO/2023/039440 (which Table is incorporated herein

by reference in its entirety), or a sequence having at least 80%, 85%, 90%, or 95% identity thereto. An NLS may be utilized with one or more copies in a polypeptide in one or more locations in a polypeptide, e.g., 1, 2, 3, or more copies of an NLS in an N-terminal domain, between peptide domains, in a C-terminal domain, or in a combination of locations, in order to
5 improve subcellular localization to the nucleus. Multiple unique sequences may be used within a single polypeptide. Sequences may be naturally monopartite or bipartite, e.g., having one or two stretches of basic amino acids, or may be used as chimeric bipartite sequences.

[0314] In some embodiments, the NLS is a bipartite NLS. A bipartite NLS typically comprises two basic amino acid clusters separated by a spacer sequence (which may be, e.g.,
10 about 10 amino acids in length). A monopartite NLS typically lacks a spacer.

[0315] In some embodiments, the gene modifying polypeptide comprises, in N-terminal to C-terminal order, one or more (e.g., 1, 2, 3, 4, 5, or all 6) of an N-terminal methionine residue, a first nuclear localization signal (NLS), a DNA binding domain, a linker, an RT domain, and/or a second NLS. In some embodiments, a gene modifying polypeptide
15 comprises, in N-terminal to C-terminal order, a first NLS, a DNA binding domain, a linker, an RT domain, and a second NLS.

[0316] In some embodiments, a gene modifying polypeptide comprises: (i) an N-terminal NLS comprising an NLS sequence of PAAKRVKLDGG (SEQ ID NO: 35), or a functional fragment thereof (e.g., an amino acid sequence having at least 70%, 75%, 80%, 85%,
20 90%, 95%, or 99% identity thereto); (ii) a Cas domain comprising a Cas sequence of Table 7 or 8 of International Application WO/2023/039440, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto; (iii) a linker comprising a linker sequence of AEAAAKEAAAKEAAAKEAAKALEAAAKEAAAKEAAAKEAAAKA (SEQ ID NO: 34), or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%,
25 or 99% identity thereto; (iv) an RT domain comprising an RT domain sequence of Table 6 of International Application WO/2023/039440, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto; and (v) a C-terminal NLS of KRTADGSEFEKRTADGSEFESPKKKAKVE (SEQ ID NO: 36), or a functional fragment thereof (e.g., an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99%
30 identity thereto). In some embodiments, the gene modifying polypeptide comprises the amino acid sequence of SEQ ID NO 23, or an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity thereto.

[0317] In some embodiments, the gene modifying polypeptide comprises an N-terminal methionine residue.

[0318] In certain embodiments, the gene modifying polypeptide further comprises a spacer sequence between the first NLS and the DNA binding domain. In certain embodiments, the spacer sequence between the first NLS and the DNA binding domain comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. In certain embodiments, the spacer sequence between the first NLS and the DNA binding domain comprises the amino acid sequence GG.

[0319] In certain embodiments, the gene modifying polypeptide further comprises a spacer sequence between the DNA binding domain and the linker. In certain embodiments, the spacer sequence between the DNA binding domain and the linker comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. In certain embodiments, the spacer sequence between the DNA binding domain and the linker comprises the amino acid sequence GG.

[0320] In certain embodiments, the gene modifying polypeptide further comprises a spacer sequence between the linker and the RT domain. In certain embodiments, the spacer sequence between the linker and the RT domain comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. In certain embodiments, the spacer sequence between the linker and the RT domain comprises the amino acid sequence GG.

[0321] In certain embodiments, the gene modifying polypeptide further comprises a spacer sequence between the RT domain and the second NLS. In certain embodiments, the spacer sequence between the RT domain and the second NLS comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids. In certain embodiments, the spacer sequence between the RT domain and the second NLS comprises the amino acid sequence AG.

Nucleic Acids encoding gene modifying polypeptides

[0322] Provided herein are nucleic acids encoding gene modifying polypeptides for use in the gene modifying systems disclosed herein. In some embodiments, the nucleic acid is mRNA.

[0323] In some embodiments, a nucleic acid (e.g., mRNA) encoding a gene modifying polypeptide is flanked by untranslated regions (UTRs) that modify protein expression levels. Various 5' and 3' UTRs may be used. In some embodiments, a nucleic acid (e.g., an RNA)

encoding a gene modifying polypeptide further comprises a poly(A) tail for enhancing expression of the polypeptide.

Template nucleic acids

[0324] The gene modifying systems described herein can modify a host target DNA site using a template nucleic acid sequence. In some embodiments, the gene modifying systems described herein transcribe an RNA sequence template into host target DNA sites by target-primed reverse transcription (TPRT). By modifying DNA sequence(s) via reverse transcription of the RNA sequence template directly into the host genome, the gene modifying system can insert an object sequence into a target genome without the need for exogenous DNA sequences to be introduced into the host cell (unlike, for example, CRISPR systems), as well as eliminate an exogenous DNA insertion step. The gene modifying system can also delete a sequence from the target genome or introduce a substitution using an object sequence. Therefore, the gene modifying system provides a platform for the use of customized RNA sequence templates containing object sequences, e.g., sequences comprising heterologous gene coding and/or function information.

[0325] In some embodiments, the template nucleic acid comprises one or more sequence (e.g., 2 sequences) that binds the gene modifying polypeptide.

[0326] In some embodiments a system or method described herein comprises a single template nucleic acid (e.g., template RNA). In some embodiments a system or method described herein comprises a plurality of template nucleic acids (e.g., template RNAs).

[0327] In some embodiments, a template RNA can comprise a gRNA sequence, e.g., to direct the gene modifying polypeptide to a target site of interest. In some embodiments, a template RNA comprises (e.g., from 5' to 3') (i) a gRNA spacer that binds a target site (e.g., a second strand of a site in a target genome), (ii) a gRNA scaffold that binds a polypeptide described herein (e.g., a gene modifying polypeptide, e.g., the Cas domain of the gene modifying polypeptide), (iii) a heterologous object sequence comprising a mutation region (optionally the heterologous object sequence comprises, from 5' to 3', a first homology region, a mutation region, and a second homology region), and (iv) a primer binding site (PBS) sequence comprising a 3' target homology domain.

[0328] In some embodiments, a template RNA molecule for use in the system comprises, from 5' to 3' (1) a gRNA spacer; (2) a gRNA scaffold; (3) heterologous object sequence (4) a primer binding site (PBS) sequence. In some embodiments:

- (1) Is a gRNA spacer of ~18-22 nt, e.g., is 20 nt
- 5 (2) Is a gRNA scaffold comprising one or more hairpin loops, e.g., 1, 2, of 3 loops for associating the template with a Cas domain, e.g., a nickase Cas9 domain. In some embodiments, the gRNA scaffold comprises the sequence, from 5' to 3',
GTTTGTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
TTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 37).
- 10 (3) In some embodiments, the heterologous object sequence is, e.g., 7-74, e.g., 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, or 70-80 nt or, 80-90 nt in length.
- (4) In some embodiments, the PBS sequence that binds the target priming sequence after nicking occurs is e.g., 3-20 nt, e.g., 7-15 nt, e.g., 12-14 nt. In some embodiments, the PBS sequence has 40-60% GC content.

15

[0329] The template nucleic acid (e.g., template RNA) component of a genome editing system described herein typically is able to bind the gene modifying polypeptide of the system. In some embodiments, the template nucleic acid (e.g., template RNA) may associate with the DNA binding domain of the polypeptide, e.g., a gRNA scaffold associating with a Cas9-derived DNA binding domain. In some embodiments, the binding region may also provide DNA target recognition, e.g., a gRNA hybridizing to the target DNA sequence and binding the polypeptide, e.g., a Cas9 domain.

20

[0330] In some embodiments the template RNA has a poly-A tail at the 3' end. In some embodiments the template RNA does not have a poly-A tail at the 3' end.

25

[0331] In some embodiments, the template RNA comprises one or more chemically modified nucleotides. In some embodiments, the template RNA comprises a backbone modification, e.g., a modification to a sugar or phosphate group in the backbone. In some embodiments, the template RNA comprises a nucleobase modification.

30

[0332] In some embodiments, a system described herein comprises two nucleic acids which together comprise the sequences of a template RNA described herein. In some embodiments, the two nucleic acids are associated with each other non-covalently, e.g., directly

associated with each other (e.g., via base pairing), or indirectly associated as part of a complex comprising one or more additional molecule.

[0333] Herein, when an RNA sequence (e.g., a template RNA sequence or an RNA sequence encoding a gene modifying polypeptide) is said to comprise a particular sequence that comprises thymine (T), it is of course understood that the RNA sequence may (and frequently does) comprise uracil (U) in place of T. For instance, the RNA sequence may comprise U at every position shown as T in the sequence. More specifically, the present disclosure provides an RNA sequence according to every nucleic acid sequence herein that comprises a T, wherein the RNA sequence has a U in place of each T. Additionally, it is understood that terminal Us and Ts may optionally be added or removed from tracrRNA sequences and may be modified or unmodified when provided as RNA.

[0334] A template RNA described herein may comprise, from 5' to 3': (1) a gRNA spacer; (2) a gRNA scaffold; (3) heterologous object sequence (4) a primer binding site (PBS) sequence. Each of these components is now described in more detail.

15 gRNA spacer and gRNA scaffold

[0335] A template RNA described herein may comprise a gRNA spacer that directs the gene modifying system to a target nucleic acid, and a gRNA scaffold that promotes association of the template RNA with the Cas domain of the gene modifying polypeptide. The systems described herein can also comprise a gRNA that is not part of a template nucleic acid. For example, a gRNA that comprises a gRNA spacer and gRNA scaffold, but not a heterologous object sequence or a PBS sequence, can be used, e.g., to induce second strand nicking, e.g., as described in the section herein entitled "Second Strand Nicking".

[0336] In some embodiments, the gRNA is a short synthetic RNA composed of a scaffold sequence that participates in CRISPR-associated protein binding and a user-defined ~20 nucleotide targeting sequence for a genomic target.

[0337] In some embodiments, the template nucleic acid (e.g., template RNA) has at least 15, 16, 17, 18, 19, 20, 21, 22, 23, or 24 bases of at least 80%, 85%, 90%, 95%, 99%, or 100% homology to the target site, e.g., at the 5' end, e.g., comprising a gRNA spacer sequence of length appropriate to the Cas9 domain of the gene modifying polypeptide.

[0338] Table 12 of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety) provides parameters to define components for designing gRNA and/or Template RNAs to apply Cas variants listed in Table 8 of International Application WO/2023/039440 for gene modifying. The gRNA for a given enzyme can be assembled by concatenating the crRNA, Tetraloop, and tracrRNA sequences, and further adding a 5' spacer of a length within Spacer (min) and Spacer (max) that matches a protospacer at a target site. In some embodiments, a gRNA scaffold described herein comprises a nucleic acid sequence comprising, in the 5' to 3' direction, a crRNA of Table 12 of International Application WO/2023/039440, a tetraloop from the same row of Table 12, and a tracrRNA from the same row of Table 12, or a sequence having at least 70%, 80%, 85%, 90%, 95%, or 99% identity thereto. In some embodiments, the gRNA or template RNA comprising the scaffold further comprises a gRNA spacer having a length within the Spacer (min) and Spacer (max) indicated in the same row of Table 12. In some embodiments, the gRNA or template RNA having a sequence according to Table 12 of International Application WO/2023/039440 is comprised by a system that further comprises a gene modifying polypeptide, wherein the gene modifying polypeptide comprises a Cas domain described in the same row of Table 12.

Heterologous object sequence

[0339] A template RNA described herein may comprise a heterologous object sequence that the gene modifying polypeptide can use as a template for reverse transcription, to write a desired sequence into the target nucleic acid. In some embodiments, the heterologous object sequence comprises, from 5' to 3', a post-edit homology region, the mutation region, and a pre-edit homology region. Without wishing to be bound by theory, an RT performing reverse transcription on the template RNA first reverse transcribes the pre-edit homology region, then the mutation region, and then the post-edit homology region, thereby creating a DNA strand comprising the desired mutation with a homology region on either side.

[0340] In some embodiments, the heterologous object sequence is 10-100, 10-90, 10-80, 10-70, 10-60, 10-50, 10-40, 10-30, or 10-20 nt in length, e.g., 10-80, 10-50, or 10-20 nt in length, e.g., about 10-20 nt in length. In some embodiments, the heterologous object sequence is 8-30, 9-25, 10-20, 11-16, or 12-15 nucleotides in length, e.g., is 11-16 nt in length.

[0341] The heterologous object sequence can be designed to result in insertions, substitutions, or deletions at the target DNA locus.

[0342] In some embodiments, the pre-edit homology domain comprises a nucleic acid sequence having 100% sequence identity with a nucleic acid sequence comprised in a target nucleic acid molecule. In some embodiments, the post-edit homology domain comprises a nucleic acid sequence having 100% sequence identity with a nucleic acid sequence comprised in a target nucleic acid molecule.

PBS sequence

[0343] In some embodiments, a template nucleic acid (e.g., template RNA) comprises a PBS sequence. In some embodiments, a PBS sequence is disposed 3' of the heterologous object sequence and is complementary to a sequence adjacent to a site to be modified by a system described herein, or comprises no more than 1, 2, 3, 4, or 5 mismatches to a sequence complementary to the sequence adjacent to a site to be modified by the system/gene modifying polypeptide. In some embodiments, the PBS sequence binds within 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides of a nick site in the target nucleic acid molecule. In some embodiments, binding of the PBS sequence to the target nucleic acid molecule permits initiation of target-primed reverse transcription (TPRT), e.g., with the 3' homology domain acting as a primer for TPRT. In some embodiments, the PBS sequence is 5-20, 8-16, 8-14, 8-13, 9-13, 9-12, or 10-12 nucleotides in length, e.g., 9-12 nucleotides in length.

Exemplary template RNA sequences – liver cells

[0344] In some embodiments of the systems and methods herein, the template RNA comprises a gRNA spacer comprising a gRNA spacer sequence of **Table 5 or 6**.

[0345] Table 5 provides a gRNA spacer sequence for correcting the pathogenic R408W mutation in the PAH gene, e.g., for treating PKU. Table 6 provides a gRNA spacer sequence for correcting the pathogenic E342K mutation in the SERPINA1 gene, e.g., for treating AATD. List of spacers, PAMs, and Cas variants for generating a nick at an appropriate position are provided herein to enable installation of a desired genomic edit with a gene modifying system. The spacers in these tables are designed to be used with a gene modifying polypeptide comprising a nickase variant of the Cas species indicated in the table. In some embodiments, a payload described herein comprises a template RNA comprising a spacer of Table 5, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto. In some embodiments, a payload described herein comprises a template RNA comprising a spacer of Table 6, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

Table 5: Exemplary gRNA spacer Cas pair for PAH

Spacer name	gRNA spacer	Cas species
hPKU5	UAGCGAACUGAGAAGGGCCA (SEQ ID NO: 38)	SpyCas9

Table 6: Exemplary gRNA spacer Cas pairs for SERPINA1

Spacer name	gRNA spacer	Cas species
ED4	AAGGCUGUGCUGACCAUCGA (SEQ ID NO: 39)	St1Cas9

5

10

15

20

25

[0346] In some embodiments, a template RNA described herein comprises a gRNA scaffold that binds SpyCas9. In some embodiments, the gRNA scaffold that binds SpyCas9 comprises a sequence according to GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 40) or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto, or a fragment thereof.

[0347] In some embodiments, a template RNA described herein comprises a gRNA scaffold that binds St1Cas9. In some embodiments, the gRNA scaffold that binds St1Cas9 comprises a sequence according to GUCUUUGUACUCUGGUACCAGAAGCUACAAAGAUAAAGGCUUCAUGCCGAAAU CAACACCCUGUCAUUUUAUGGCAGGGUGUUUU (SEQ ID NO: 41) or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto, or a fragment thereof.

[0348] In some embodiments of the systems and methods herein, e.g., for modifying PAH, e.g., for correcting a R408W mutation, the heterologous object sequence comprises an RT template sequence from Table 7. The RT sequences and PBS sequences of Table 7 are compatible with, for example, the spacer listed in the first column of Table 7. In some embodiments, the primer binding site (PBS) sequence has a sequence comprising the PBS sequence from the same row of Table 7 as the RT template sequence. In some embodiments, a payload described herein comprises a template RNA comprising one or both of an RT and PBS of Table 7, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

[0349] In some embodiments of the systems and methods herein, e.g., for modifying SERPINA1, e.g., for correcting a E342K mutation, the heterologous object sequence comprises an RT template sequence from Table 8. The RT sequences and PBS sequences of Table 8 are compatible with, for example, the spacer listed in the first column of Table 8. In some
5 embodiments, the primer binding site (PBS) sequence has a sequence comprising the PBS sequence from the same row of Table 8 as the RT template sequence. In some embodiments, a payload described herein comprises a template RNA comprising one or both of an RT and PBS of Table 8, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

10 **Table 7: Exemplary RT sequence (heterologous object sequence) and PBS sequence pairs for PAH**

Exemplary compatible spacer name	RT Template Sequence	PBS Sequence
hPKU5	AAUACCUCGG (SEQ ID NO: 42)	CCCUUCUCAG (SEQ ID NO: 43)

Table 8: Exemplary RT sequence (heterologous object sequence) and PBS sequence pairs for SERPINA1

Exemplary compatible spacer name	RT Template Sequence	PBS Sequence
ED4	UUUCUCGUCG (SEQ ID NO: 44)	AUGGUCAG (SEQ ID NO: 46)
ED4	CCUUUCUCGUCG (SEQ ID NO: 45)	AUGGUCAG (SEQ ID NO: 46)

15 [0350] In some embodiments of the systems and methods herein, the system further comprises a second strand-targeting gRNA that directs a nick to the second strand of the human SERPINA1 gene or the human PAH gene. In some embodiments, the systems and methods provided herein, e.g., for modifying PAH, e.g., for correcting a R408W mutation, may
20 comprise a template sequence listed in **Table 9**. **Table 9** shows the full-length sequence of a template RNA described herein, e.g., comprising a spacers of **Table 5**, e.g., comprising the heterologous object sequence and/or PBS sequence of **Table 7**. In some embodiments, a template RNA comprises a sequence according to **Table 9**, or a sequence having at least 80%,

85%, 90%, 95%, 97%, 98%, or 99% identity thereto. In some embodiments, the template RNA according to **Table 9** herein comprises one or more chemical modifications.

[0351] In some embodiments, the systems and methods provided herein, e.g., for modifying SERPINA1, e.g., for correcting a E342K mutation, may comprise a template sequence listed in **Table 10**. **Table 10** shows the full-length sequences of various template RNAs described herein, e.g., comprising a spacer of **Table 6**, e.g., comprising the heterologous object sequence and/or PBS sequence of **Table 8**. In some embodiments, a template RNA comprises a sequence according to **Table 10**, or a sequence having at least 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity thereto. In some embodiments, the template RNA according to **Table 10** herein comprises one or more chemical modification.

Table 9: Exemplary template RNA sequences for PAH

RN ACS #	Name	Sequence	SEQ ID NO
RN ACS 4134	hPKU 5_R1 0P10	UAGCGAACUGAGAAGGGCCAGUUUUAGAGCUAGAAAUAGCAAGUU AAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCAAUACCUCGGCCCUUCUCAG	47

Table 10: Exemplary template RNA sequences for SERPINA1

Name	Sequence	SEQ ID NO
A1AT_St 1_ED4- _G_5FE_ _PB8	AAGGCUGUGCUGACCAUCGAGUCUUUGUACUCUGGUACCAG AAGCUACAAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUG UCAUUUUAUGGCAGGGUGUUUUUUCUCGUCGAUGGUCAG	48
A1AT_St 1_ED4- _G_7FE_ _PB8	AAGGCUGUGCUGACCAUCGAGUCUUUGUACUCUGGUACCAGAAGCUAC AAAGAUAAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUAUGGCAG GGUGUUUUCUUUCUCGUCGAUGGUCAG	49

[0352] Certain heterologous object sequences and PBS sequences provided herein, e.g., in **Table 7**, may be designed to correct the PAH mutation by replacing a “T” nucleotide with a “C” nucleotide at the mutation site via gene editing, to reverse a R408W mutation in the corresponding protein. Certain heterologous object sequences and PBS sequences provided herein, e.g., in **Table 8**, may be designed to correct the AATD mutation by replacing an “A”

nucleotide with a “G” nucleotide at the mutation site via gene editing, to reverse an E342K mutation in the corresponding protein.

5 [0353] In some embodiments, a template RNA of this section, for modifying PAH, is used together with a gene modifying polypeptide having a sequence of SEQ ID NO: 32, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto. In some embodiments, the gene modifying polypeptide is provided as a nucleic acid (e.g., an mRNA) encoding the amino acid sequence of SEQ ID NO: 32, or a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto.

10 [0354] In some embodiments, a template RNA of this section, for modifying SERPINA1, is used together with a gene modifying polypeptide having a sequence of SEQ ID NO: 31, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto. In some embodiments, the gene modifying polypeptide is provided as a nucleic acid (e.g., an mRNA) encoding the amino acid sequence of SEQ ID NO: 31, or a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity thereto.

15 [0355] The template RNA sequences shown herein may be customized depending on various factors. For example, in some embodiments it is desired to inactivate a PAM sequence upon editing (e.g., using a “PAM-kill” modification) to decrease the potential for further gene editing (e.g., by Cas retargeting) following the initial edit. Consequently, certain template RNAs described herein are designed to write a mutation (e.g., a substitution) into the PAM of
20 the target site, such that upon editing, the PAM site will be mutated to a sequence no longer recognized by the gene modifying polypeptide. Thus, a mutation region within the heterologous object sequence of the template RNA may comprise a PAM-kill sequence. Without wishing to be bound by theory, in some embodiments, a PAM-kill sequence prevents re-engagement of the gene modifying polypeptide upon completion of a genetic modification,
25 or decreases re-engagement relative to a template RNA lacking a PAM-kill sequence. In some embodiments, a PAM-kill sequence does not alter the amino acid sequence encoded by a gene, e.g., the PAM-kill sequence results in a silent mutation. In other embodiments, it is desired to leave the PAM sequence intact (no PAM-kill).

30 [0356] Similarly, in some embodiments, to decrease the potential for further gene editing (e.g., by Cas retargeting) following the initial edit, it may be desirable to alter the first three nucleotides of the RT template sequence via a “seed-kill” motif. Consequently, certain

template RNAs described herein are designed to write a mutation (e.g., a substitution) into the portion of the target site corresponding to the first three nucleotides of the RT template sequence, such that upon editing, the target site will be mutated to a sequence with lower homology to the RT template sequence. Thus, a mutation region within the heterologous object
5 sequence of the template RNA may comprise a seed-kill sequence. Without wishing to be bound by theory, in some embodiments, a seed-kill sequence prevents re-engagement of the gene modifying polypeptide upon completion of genetic modification, or decreases re-engagement relative to an otherwise similar template RNA lacking a seed-kill sequence. In some embodiments, a seed-kill sequence does not alter the amino acid sequence encoded by a
10 gene, e.g., the seed-kill sequence results in a silent mutation. In other embodiments, it is desired to leave the seed region intact, and a seed-kill sequence is not used.

[0357] In further embodiments, to optimize or improve gene editing efficiency, it may be desirable to evade the target cell's mismatch repair or nucleotide repair pathways or to bias the target cell's repair pathways toward preservation of the edited strand. In some
15 embodiments, multiple silent mutations (for example, silent substitutions) may be introduced within the RT template sequence to evade the target cell's mismatch repair or nucleotide repair pathways or to bias the target cell's repair pathways toward preservation of the edited strand.

Exemplary template RNA sequences – HSCs

20 [0358] In some embodiments, a template RNA described herein may comprise a single nucleotide polymorphism (SNP) relative to a reference sequence. In some embodiments, a cell line used for screening may contain one or more additional SNPs in the *HBB* locus relative to a patient or reference sequence, e.g., the hg38 human genome reference sequence, and a landing pad containing the target mutation is optionally designed to carry the one or more non-
25 pathogenic SNPs to match the endogenous cell line *HBB* locus, e.g., designed to carry a mutation that recapitulates a SNP present in the endogenous *HBB* locus in HEK293T cells. Without wishing to be limited by example, it is understood that template RNA sequences found to successfully edit a target mutation at a site containing an additional SNP relative to a reference sequence would differ from a therapeutic template RNA in any region overlapping
30 the additional SNP. For example, a successful template RNA in a HEK293T-based screening assay where a genomic landing pad contains the target mutation (corresponding to the endogenous E6V mutation caused by DNA substitution NC_000011.10:g.5227002T>A) and an additional substitution relative to hg38 (corresponding to the

NC_000011.10:g.5227013T>C mutation at the endogenous *HBB* locus in HEK293T cells) in the protospacer may provide a candidate composition where the corresponding therapeutic template RNA would thus have a substitution (C>T) in the spacer region relative to the corresponding spacer region of the screening template RNA, in order to enable therapeutic

5 correction of the E6V mutation at a target site lacking the additional substitution, e.g., at a target site comprising the pathogenic E6V mutation but otherwise matching the hg38 reference sequence. In this example, a screening cell line containing a target site landing pad comprising the pathogenic mutation with an additional T>C substitution in the protospacer region might be corrected using a screening template RNA comprising the spacer sequence 5'-

10 CATGGTGCACCTGACTCCTG-3' (SEQ ID NO: 50) , whereas the corresponding therapeutic template RNA might comprise the spacer sequence 5'-CATGGTGCATCTGACTCCTG-3'(SEQ ID NO: 51), where the underlined nucleotides indicate the position that is altered to match either the screening cell target sequence or the hg38 target sequence. In some embodiments, the spacer, PBS, and/or RT template regions may

15 need to be adjusted in this manner to account for any discrepancies between screening and reference target sequences. It is further contemplated that a given patient or patient population may possess one or more SNPs relative to hg38 at the target locus in addition to the pathogenic E6V mutation and thus a similar adaptation of candidate template RNA molecules could be used to generate template RNA sequences specific for the patient or patient population.

20 [0359] In some embodiments, a gene modifying system used for correcting the pathogenic E6V mutation in *HBB* is described in WO2023039440, which is hereby incorporated by reference in its entirety.

[0360] Table 11 provides a plurality of gRNAs for correcting the pathogenic EV6 mutation in *HBB*. List of spacers, PAMs, and Cas variants for generating a nick at an

25 appropriate position to enable installation of a desired genomic edit with a gene modifying system. The spacers in this table are designed to be used with a gene modifying polypeptide comprising a nickase variant of the Cas species indicated in the table. In some embodiments, a payload described herein comprises a template RNA comprising a spacer of Table 11, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

Table 11: Exemplary gRNA spacer Cas pairs

Spacer name	gRNA spacer	Cas species
HBB5 with SNP	CATGGTGCACCTGACTCCTG (SEQ ID NO: 50)	SpyCas9
HBB5 no SNP	CATGGTGCATCTGACTCCTG (SEQ ID NO: 51)	SpyCas9
HBB8	GTAACGGCAGACTTCTCCAC (SEQ ID NO: 52)	SpyCas9

[0361] In some embodiments, a template RNA described herein comprises a gRNA scaffold that binds SpyCas9. In some embodiments, the gRNA scaffold that binds SpyCas9 comprises a sequence according to GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA AAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 53) or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto, or a fragment thereof.

[0362] In some embodiments of the systems and methods herein, the heterologous object sequence comprises an RT template sequence from Table 12. The RT sequences of Table 12 are compatible with, for example, the spacer HBB5 with SNP or HBB5 no SNP. In some embodiments, a payload described herein comprises a template RNA comprising an RT of Table 12, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

Table 12: Exemplary RT sequence (heterologous object sequence)

Exemplary compatible spacer name	RT Template Sequence	SEQ ID NO
HBB5	AGTAACGGCAGACTTCTCTTCCG	54
HBB5	GTAACGGCAGACTTCTCTTCCG	55
HBB5	TAACGGCAGACTTCTCTTCCG	56
HBB5	AACGGCAGACTTCTCTTCCG	57
HBB5	ACGGCAGACTTCTCTTCCG	58
HBB5	CGGCAGACTTCTCTTCCG	59
HBB5	GGCAGACTTCTCTTCCG	60
HBB5	GCAGACTTCTCTTCCG	61
HBB5	CAGACTTCTCTTCCG	62
HBB5	AGACTTCTCTTCCG	63
HBB5	GACTTCTCTTCCG	64
HBB5	ACTTCTCTTCCG	65
HBB5	CTTCTCTTCCG	66

HBB5	TTCTCTTCCG	67
HBB5	AGTAACGGCAGACTTCTCTGCCG	68
HBB5	GTAACGGCAGACTTCTCTGCCG	69
HBB5	TAACGGCAGACTTCTCTGCCG	70
HBB5	AACGGCAGACTTCTCTGCCG	71
HBB5	ACGGCAGACTTCTCTGCCG	72
HBB5	CGGCAGACTTCTCTGCCG	73
HBB5	GGCAGACTTCTCTGCCG	74
HBB5	GCAGACTTCTCTGCCG	75
HBB5	CAGACTTCTCTGCCG	76
HBB5	AGACTTCTCTGCCG	77
HBB5	GACTTCTCTGCCG	78
HBB5	ACTTCTCTGCCG	79
HBB5	CTTCTCTGCCG	80
HBB5	TTCTCTGCCG	81

[0363] In some embodiments of the systems and methods herein, the PBS comprises an PBS template sequence from Table 13. The RT sequences of Table 13 are compatible with, for example, the spacer HBB5 with SNP or HBB5 no SNP. In some embodiments, a payload described herein comprises a template RNA comprising an PBS of Table 13, or a sequence with no more than 1, 2, or 3 sequence differences relative thereto.

Table 13: Exemplary PBS sequence

Exemplary compatible spacer name	PBS Sequence	SEQ ID NO
HBB5	GAGTCAGGTGCACCATG	82
HBB5	GAGTCAGGTGCACCAT	83
HBB5	GAGTCAGGTGCACCA	84
HBB5	GAGTCAGGTGCACC	85
HBB5	GAGTCAGGTGCAC	86
HBB5	GAGTCAGGTGCA	87
HBB5	GAGTCAGGTGC	88
HBB5	GAGTCAGGTG	89
HBB5	GAGTCAGGT	90
HBB5	GAGTCAGG	91

[0364] In some embodiments of the systems and methods herein, the heterologous object sequence comprises an RT template sequence from Table 14. The RT sequences and PBS sequences of Table 14 are compatible with the spacers listed in the first column of Table

14. In some embodiments, the primer binding site (PBS) sequence has a sequence comprising the PBS sequence from the same row of Table 14 as the RT template sequence.

[0365] Table 14 provides exemplified PBS sequences and heterologous object sequences (reverse transcription template regions) of a template RNA for correcting the pathogenic EV6 mutation in HBB. gRNA spacers were filtered, e.g., filtered by occurrence within 15 nt of the desired editing location and use of a Tier 1 Cas enzyme. PBS sequences and heterologous object sequences (reverse transcription template regions) were designed relative to the nick site directed by the cognate gRNA, as described in this application. For exemplification, these regions were designed to be 8-17 nt (priming) and 1-50 nt extended beyond the location of the edit (RT). Without wishing to be limited by example, given variability of length, sequences are provided that use the maximum length parameters and comprise all templates of shorter length within the given parameters. Sequences are shown with uppercase letters indicating core sequence and lowercase letters indicating flanking sequence that may be truncated within the described length parameters.

Table 14: Exemplary RT sequence (heterologous object sequence) and PBS sequence pairs

Exemplary compatible spacer name	RT Template Sequence	PBS Sequence
HBB5	AGACTTCTCTTCCG (SEQ ID NO: 63)	GAGTCAGGTG (SEQ ID NO: 89)
HBB8	ctgtgtcactagcaacctcaaacagacaccatggtgcatctgactcctGAG (SEQ ID NO: 92)	GAGAAGTCtgccgttac (SEQ ID NO: 93)
HBB5	<u>AGTAACGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 54)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>GTAACGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 55)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>TAACGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 56)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>AACGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 57)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>ACGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 58)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>CGGCAGACTTCTCTTCCG</u> (SEQ ID NO: 59)	GAGUCAGGUGC (SEQ ID NO: 94)

HBB5	<u>GGCAGACTTCTCTTCCG</u> (SEQ ID NO: 60)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>GCAGACTTCTCTTCCG</u> (SEQ ID NO: 61)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>CAGACTTCTCTTCCG</u> (SEQ ID NO: 62)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>AGACTTCTCTTCCG</u> (SEQ ID NO: 63)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>GACTTCTCTTCCG</u> (SEQ ID NO: 64)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>ACTTCTCTTCCG</u> (SEQ ID NO: 65)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>CTTCTCTTCCG</u> (SEQ ID NO: 66)	GAGUCAGGUGC (SEQ ID NO: 94)
HBB5	<u>TTCTCTTCCG</u> (SEQ ID NO: 67)	GAGUCAGGUGC (SEQ ID NO: 94)

Capital letters indicate “core nucleotides” while lower case letters indicate “flanking nucleotides.”

[0366] In some embodiments of the systems and methods herein, the system further comprises a second strand-targeting gRNA that directs a nick to the second strand of the human HBB gene. In some embodiments, the second strand-targeting gRNA comprises a gRNA spacer sequence, gRNA scaffold sequence, or both, from **Table 15**. **Table 15** provides exemplified second-nick gRNA species for optional use for correcting the pathogenic E6V mutation in HBB.

Table 15: Exemplary second nick gRNAs

Name	RNA sequence	SEQ ID NO
HBB5_216rv	CUUGCCCCACAGGGCAGUAAGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	95
HBB5_24rv	UGCAGGAGUCAGGUGCACCAGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	96
HBB5_34rv	CAGACUUCUCUGCAGGAGUCGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	97

HBB5 _34rv_ hs1	CAGACUUCUCUGCCGGAGUCGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	98
HBB5 _41rv	GUAACGGCAGACUUCUCUGCGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	99
HBB5 _122rv	AAGCAAAUGUAAGCAAUAGAGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	100
HBB5 _92rv	CUGACUUUUUUGCCCCAGCCCGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	101
HBB5 _g27	CCUUGAUACCAACCUGCCCAGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	102
HBB5 _g37	CACGUUCACCUUGCCCCACAGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	103
HBB5 _g38	CCACGUUCACCUUGCCCCACGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	104
HBB5 _g39	ACCUUGAUACCAACCUGCCCCGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	105
HBB5 _g40	UCCACAUGCCCAGUUUCUAUGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	106
HBB8 _gRN A1	CAGGGCUGGGCAUAAAAGUCGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	107
HBB8 _gRN A2	AGGGCUGGGCAUAAAAGUCAGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	108
HBB8 _gRN A3	GCAACCUCAAACAGACACCAGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	109
HBB8 _gRN A4	GGAGGGCAGGAGCCAGGGCUGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	110
HBB8 _231f w	GUCUGCCGUUACUGCCCUGUGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	111
HBB8 _237f w	CGUUACUGCCCUGUGGGGCAGUUUUAGAGCUAGAAAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC ACCGAGUCGGUGCUUUU	112

HBB8 _246f w	CCUGUGGGGCAAGGUGAACGGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	113
HBB8 _256f w	AAGGUGAACGUGGAUGAAGUGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	114
HBB8 _270f w	UGAAGUUGGUGGUGAGGCCCGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	115
HBB8 _279f w	UGGUGAGGCCCGUGGGCAGGUGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	116
HBB8 _299f w	UGGUAUCAAGGUUACAAGACGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	117
HBB8 _306f w	AAGGUUACAAGACAGGUUUAGUUUUAGAGCUAGAAAUAGCA AGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUU	118

[0367] In some embodiments, the systems and methods provided herein, e.g., for modifying HBB, e.g., for correcting an E6V mutation, may comprise a template sequence listed in Table 16. Table 16 shows the full-length sequences of various template RNAs described herein, e.g., comprising spacers of Table 11. In some embodiments, a template RNA comprises a sequence according to Table 16, or a sequence having at least 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity thereto. In some embodiments, the template RNA according to Table 16 herein comprises one or more chemical modification, e.g., chemical modifications as described in Table 14 of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety).

Table 16: Exemplary template RNA sequences

Name	Sequence	SEQ ID NO
tg14_hs1	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUAGCA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGACUUCUCUGCCGGAGU CAGGUG	119
tg14_hs1-SCD-wt	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUAGCA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGACUUCUCUCCGGAGU CAGGUG	120

tgRNA34_H BB8h-SCD-M	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCUGCGGAGA AGUCUGC	121
tgRNA34_H BB8h-SCD-wt	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCUGAGGAGA AGUCUGC	122
tgRNA34_H BB8hs4-MK	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGCAGAGA AGUCUGC	123
tgRNA34_H BB8hs4-WT	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGAAGAGA AGUCUGC	124
tgRNA34_H BB8hs7-MK	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGCUGAGA AGUCUGC	125
tgRNA34_H BB8hs7-WT	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGAUGAGA AGUCUGC	126
tgRNA34_H BB8hs10-MK	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGCCGAGA AGUCUGC	127
tgRNA34_H BB8hs10-WT	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGACGAGA AGUCUGC	128
tgRNA34_H BB8hs13-MK	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGCGGAGA AGUCUGC	129
tgRNA34_H BB8hs13-WT	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACCUGACUCCCGAGGAGA AGUCUGC	130
tg14_PAMT _hs1-SCD-wt	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGACUUCUCAGCCGGAGU CAGGUG	131

tg14_PAMC _hs1-SCD- wt	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGACUUCUCGGCCGGAGU CAGGUG	132
HBB8_RT20 _PBS8	GUAACGGCAGACUUCUCCACGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUGGUGCACCUGACUCCUG AGGAGAAGUC	133
tg43_hs1	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGUAACGGCAGACUUCUC UGCCGGAGUCAGGUGC	134
tg42_hs1	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUAACGGCAGACUUCUCUG CCGGAGUCAGGUGC	135
tg41_hs1	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCACGGCAGACUUCUCUGCC GGAGUCAGGUGC	136
tg19_hs1	CAUGGUGCACCUGACUCCUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCAGACUUCUCUGCCGG AGUCAGGUGC	137

[0368] The heterologous object sequences and PBS sequences may be designed to correct the SCD mutation by replacing a “T” nucleotide with an “A” nucleotide (wild type) or with a “C” (Makassar installation) at the mutation site using a gene modifying system described herein.

[0369] The template RNA sequences shown herein may be customized depending on various factors. For example, in some embodiments it is desired to inactivate a PAM sequence upon editing (e.g., using a “PAM-kill” modification) to decrease the potential for further gene editing (e.g., by Cas retargeting) following the initial edit. Consequently, certain template RNAs described herein are designed to write a mutation (e.g., a substitution) into the PAM of the target site, such that upon editing, the PAM site will be mutated to a sequence no longer recognized by the gene modifying polypeptide. Thus, a mutation region within the heterologous object sequence of the template RNA may comprise a PAM-kill sequence. Without wishing to be bound by theory, in some embodiments, a PAM-kill sequence prevents re-engagement of the gene modifying polypeptide upon completion of a genetic modification, or decreases re-engagement relative to a template RNA lacking a PAM-kill sequence. In some

embodiments, a PAM-kill sequence does not alter the amino acid sequence encoded by a gene, e.g., the PAM-kill sequence results in a silent mutation. In other embodiments, it is desired to leave the PAM sequence intact (no PAM-kill).

[0370] In some embodiments it is desired to inactive a PAM sequence upon editing (“PAM-kill”) and in other embodiments it is preferred to leave the PAM sequence intact (no PAM-kill). The RT template can be designed as a “PAM-kill” or “no PAM-kill” version, for example, as shown below.

HBB5 Spacer (no SNP): CATGGTGCATCTGACTCCTG (SEQ ID NO: 51)
 HBB5 PBS (no SNP): GAGTCAGAtgcaccatg (SEQ ID NO: 138)
 10 HBB5 RT template (no PAM-kill): aacggcagactTCTCGTCAG (SEQ ID NO: 139)
 HBB8 RT template (no SNP): tggatcatctgACTCCTGAG (SEQ ID NO: 140)

[0371] Similarly, in some embodiments, to decrease the potential for further gene editing (e.g., by Cas retargeting) following the initial edit, it may be desirable to alter the first three nucleotides of the RT template sequence via a “seed-kill” motif. Consequently, certain template RNAs described herein are designed to write a mutation (e.g., a substitution) into the portion of the target site corresponding to the first three nucleotides of the RT template sequence, such that upon editing, the target site will be mutated to a sequence with lower homology to the RT template sequence. Thus, a mutation region within the heterologous object sequence of the template RNA may comprise a seed-kill sequence. Without wishing to be bound by theory, in some embodiments, a seed-kill sequence prevents re-engagement of the gene modifying polypeptide upon completion of genetic modification, or decreases re-engagement relative to an otherwise similar template RNA lacking a seed-kill sequence. In some embodiments, a seed-kill sequence does not alter the amino acid sequence encoded by a gene, e.g., the seed-kill sequence results in a silent mutation. In other embodiments, it is desired to leave the seed region intact, and a seed-kill sequence is not used.

[0372] In further embodiments, to optimize or improve gene editing efficiency, it may be desirable to evade the target cell’s mismatch repair or nucleotide repair pathways or to bias the target cell’s repair pathways toward preservation of the edited strand. In some embodiments, multiple silent mutations (for example, silent substitutions) may be introduced within the RT template sequence to evade the target cell’s mismatch repair or nucleotide repair pathways or to bias the target cell’s repair pathways toward preservation of the edited strand.

[0373] Table 7A of International Application WO/2023/039440 (which Table is incorporated herein by reference in its entirety) provides exemplary silent mutations for various positions within the HBB gene. These silent mutations may be incorporated into a template RNA described herein.

5 Second Strand Nicking

[0374] In some embodiments, a gene modifying system described herein comprises a nickase activity (e.g., in the gene modifying polypeptide) that nicks the first strand, and a nickase activity (e.g., in the same gene modifying polypeptide or a polypeptide separate from
10 the gene modifying polypeptide) that nicks the second strand of target DNA. As discussed herein, without wishing to be bound by theory, nicking of the first strand of the target site DNA is thought to provide a 3' OH that can be used by an RT domain to reverse transcribe a sequence of a template RNA, e.g., a heterologous object sequence. Without wishing to be bound by theory, it is thought that introducing an additional nick to the second strand may bias the
15 cellular DNA repair machinery to adopt the heterologous object sequence-based sequence more frequently than the original genomic sequence. In some embodiments, the additional nick to the second strand is made by the same endonuclease domain (e.g., nickase domain) as the nick to the first strand. In some embodiments, the same gene modifying polypeptide performs both the nick to the first strand and the nick to the second strand. In some embodiments, the gene
20 modifying polypeptide comprises a CRISPR/Cas domain and the additional nick to the second strand is directed by an additional nucleic acid, e.g., comprising a second gRNA directing the CRISPR/Cas domain to nick the second strand. In other embodiments, the additional second strand nick is made by a different endonuclease domain (e.g., nickase domain) than the nick to the first strand. In some embodiments, that different endonuclease domain is situated in an
25 additional polypeptide (e.g., a system of the invention further comprises the additional polypeptide), separate from the gene modifying polypeptide. In some embodiments, the additional polypeptide comprises an endonuclease domain (e.g., nickase domain) described herein. In some embodiments, the additional polypeptide comprises a DNA binding domain, e.g., described herein.

30 [0375] It is contemplated herein that the position at which the second strand nick occurs relative to the first strand nick may influence the extent to which one or more of: desired gene modifying DNA modifications are obtained, undesired double-strand breaks (DSBs) occur, undesired insertions occur, or undesired deletions occur. Without wishing to be bound by

theory, second strand nicking may occur in two general orientations: inward nicks and outward nicks. Inward nick and outward nick orientations are described in more detail in International Application WO/2023/039440.

[0376] Exemplary gRNAs suitable for introducing second strand nicks are provided,
5 e.g., in Table 13 herein.

Regulatory Edits

[0377] In some embodiments, the gene modifying system can be used to introduce a regulatory edit. In some embodiments, the regulatory edit is introduced to a regulatory sequence of a gene, for example, a gene promoter, gene enhancer, gene repressor, or a sequence
10 that regulates gene splicing. In some embodiments, the regulatory edit increases or decreases the expression level of a target gene. In some embodiments, the target gene is the same as the gene containing a disease-causing mutation. In some embodiment, the target gene is different from the gene containing a disease-causing mutation. For example, the systems or methods provided herein can be used to upregulate the expression of fetal hemoglobin by introducing a
15 regulatory edit at the promoter of *bcl11a*, thereby treating sickle cell disease.

[0378] In some embodiments, multiple components of a gene modifying system may be prepared as a single LNP formulation, e.g., an LNP formulation comprises mRNA encoding for the gene modifying polypeptide and an RNA template. Ratios of nucleic acid components may be varied in order to maximize the properties of a therapeutic. In some embodiments, the
20 ratio of RNA template to mRNA encoding a gene modifying polypeptide is about 1:1 to 100:1, e.g., about 1:1 to 20:1, about 20:1 to 40:1, about 40:1 to 60:1, about 60:1 to 80:1, or about 80:1 to 100:1, by molar ratio. In other embodiments, a system of multiple nucleic acids may be prepared by separate formulations, e.g., one LNP formulation comprising a template RNA and a second LNP formulation comprising an mRNA encoding a gene modifying polypeptide. In
25 some embodiments, the system may comprise more than two nucleic acid components formulated into LNPs. In some embodiments, the system may comprise a protein, e.g., a gene modifying polypeptide, and a template RNA formulated into at least one LNP formulation.

[0379] In some embodiments, the gene modifying system can be used to introduce a regulatory edit. In some embodiments, the regulatory edit is introduced to a regulatory
30 sequence of a gene, for example, a gene promoter, gene enhancer, gene repressor, or a sequence that regulates gene splicing. In some embodiments, the regulatory edit increases or decreases

the expression level of a target gene. In some embodiments, the target gene is the same as the gene containing a disease-causing mutation. In some embodiment, the target gene is different from the gene containing a disease-causing mutation. For example, the systems or methods provided herein can be used to upregulate the expression of fetal hemoglobin by introducing a regulatory edit at the promoter of bcl11a, thereby treating sickle cell disease.

[0380] Table 17 provides exemplary indications (column 1), genes (column 2), and regulatory edits that can be introduced using the systems or methods described herein (column 3).

Table 17. Indications, genes, and compensatory regulatory edits.

Disease	Gene	Nucleotide Change [#]
homozygous familial hypercholesterolaemia	LDLR	c.81C>T
Porphyrias	ALAS1	c.3G>A
Porphyrias	ALAS1	c.2T>C
Porphyrias	ALAS1	c.46C>T
Porphyrias	ALAS1	c.91C>T
Porphyrias	ALAS1	c.91C>T
Porphyrias	ALAS1	c.226C>T
Porphyrias	ALAS1	c.226C>T
Porphyrias	ALAS1	c.226C>T
Porphyrias	ALAS1	c.229C>T
Porphyrias	ALAS1	c.247C>T
Porphyrias	ALAS1	c.247C>T
Porphyrias	ALAS1	c.250C>T
Porphyrias	ALAS1	c.250C>T
Porphyrias	ALAS1	c.340C>T
Porphyrias	ALAS1	c.340C>T
Porphyrias	ALAS1	c.349C>T
Porphyrias	ALAS1	c.391C>T
Porphyrias	ALAS1	c.391C>T
Porphyrias	ALAS1	c.403C>T
Porphyrias	ALAS1	c.403C>T
Porphyrias	ALAS1	c.199+1G>A
Porphyrias	ALAS1	c.199+1G>A
Porphyrias	ALAS1	c.199+1G>A
Porphyrias	ALAS1	c.199+1G>A
Porphyrias	ALAS1	c.199+2T>C

Disease	Gene	Nucleotide Change [#]
sickle cell disease	HBG1/2	C. -114~ -102 deletion
sickle cell disease	HBG1/2	C. -114~ -102 deletion
sickle cell disease	HBG1/2	c. -90 BCL11A Binding
sickle cell disease	HBG1/2	c. -90 BCL11A Binding
sickle cell disease	HBG1/2	C. -202 C>T, -201 C>T, -198 T>C, -197 C>T, -196 C>T, -195 C>G
sickle cell disease	HBG1/2	C. -197 C>T, -196 C>T, -195 C>G

[#]: See J T den Dunnen and S E Antonarakis, Hum Mutat. 2000;15(1):7-12, herein incorporated by reference in its entirety, for details of the nomenclatures of gene mutations

T cell editing

5

[0381] The conjugates described herein may be used to target and modify immune cells. In some embodiments, the conjugates may be used to modify T cells. In some embodiments, T-cells may include any subpopulation of T-cells, e.g., CD4+, CD8+, gamma-delta, naïve T cells, stem cell memory T cells, central memory T cells, or a mixture of subpopulations. In some embodiments, the conjugates may be used to deliver or modify a T-cell receptor (TCR) in a T cell. In some embodiments, the conjugates may be used to deliver at least one chimeric antigen receptor (CAR) to T-cells. For instance, in specific embodiments, the conjugates can be used to deliver mRNA encoding a CAR to T-cells. In some embodiments, the conjugates may be used to deliver at least one CAR to natural killer (NK) cells. In some embodiments, the conjugates be used to deliver at least one CAR to natural killer T (NKT) cells. In some embodiments, the conjugates may be used to deliver at least one CAR to a progenitor cell, e.g., a progenitor cell of T, NK, or NKT cells. In some embodiments, cells modified with at least one CAR (e.g., CAR-T cells, CAR-NK cells, CAR-NKT cells), or a combination of cells modified with at least one CAR (e.g., a mixture of CAR-NK/T cells) are used to treat a condition as identified in the targetable landscape of CAR therapies in MacKay, et al. Nat Biotechnol 38, 233-244 (2020), incorporated by reference herein in its entirety. In some embodiments, the immune cells comprise a CAR specific to a tumor or a pathogen antigen selected from a group consisting of AChR (fetal acetylcholine receptor), ADGRE2, AFP (alpha fetoprotein), BAFF-R, BCMA, CAIX (carbonic anhydrase IX), CCR1, CCR4, CEA (carcinoembryonic antigen), CD3, CD5, CD8, CD7, CD10, CD13, CD14, CD15, CD19, CD20, CD22, CD30, CD33, CLLI, CD34, CD38, CD41, CD44, CD49f, CD56, CD61, CD64, CD68, CD70, CD74, CD99, CD117, CD123, CD133, CD138, CD44v6, CD267, CD269, CDS,

25

CLEC12A, CS1, EGP-2 (epithelial glycoprotein-2), EGP-40 (epithelial glycoprotein-40), EGFR(HER1), EGFR-VIII, EpCAM (epithelial cell adhesion molecule), EphA2, ERBB2 (HER2, human epidermal growth factor receptor 2), ERBB3, ERBB4, FBP (folate-binding protein), Flt3 receptor, folate receptor-a, GD2 (ganglioside G2), GD3 (ganglioside G3), GPC3
 5 (glypican-3), GPI00, hTERT (human telomerase reverse transcriptase), ICAM-1, integrin B7, interleukin 6 receptor, IL13Ra2 (interleukin-13 receptor 30 subunit alpha-2), kappa-light chain, KDR (kinase insert domain receptor), LeY (Lewis Y), L1CAM (LI cell adhesion molecule), LILRB2 (leukocyte immunoglobulin like receptor B2), MARTI, MAGE-A1 (melanoma associated antigen A1), MAGE- A3, MSLN (mesothelin), MUC16 (mucin 16), MUCI (mucin
 10 I), KG2D ligands, NY-ESO-1 (cancer-testis antigen), PRI (proteinase 3), TRBCI, TRBC2, TFM-3, TACI, tyrosin, hTERT, oncofetal antigen (h5T4), p53, PSCA (prostate stem cell antigen), PSMA (prostate-specific membrane antigen), hROR1, TAG-72 (tumor- associated glycoprotein 72), VEGF-R2 (vascular endothelial growth factor R2), WT-1 (Wilms tumor protein), and antigens of HIV (human immunodeficiency virus), hepatitis B, hepatitis C, CMV
 15 (cytomegalovirus), EBV (Epstein-Barr virus), HPV (human papilloma virus).

[0382] In some embodiments, immune cells, e.g., T-cells, NK cells, NKT cells, or progenitor cells are modified ex vivo and then delivered to a patient. In some embodiments, a nucleic acid (e.g., mRNA) is delivered by one of the methods mentioned herein, and immune cells, e.g., T-cells, NK cells, NKT cells, or progenitor cells are modified in vivo in the patient.

20 [0383] In certain embodiments, the targeting moiety is a T-cell targeting moiety, for example, an antibody, Fab fragment or ScFv that binds to a T-cell antigen selected from the group consisting of CD2, CD3, CD4, CD5, CD7, CD8, CD28, CD137, CD45, T-cell receptor (TCR) β , TCR- α , TCR- α/β , TCR- γ/δ , PD1, CTLA4, TIM3, LAG3, CD18, IL-2 receptor, CD11a, TLR2, TLR4, TLR5, IL-7 receptor, or IL-15 receptor.

25 V. METHODS OF TREATMENT

[0384] In some embodiments, an LNP or conjugate (targeted LNP) described herein is used to treat a disease, disorder, or condition. In some embodiments, an LNP or conjugate described herein, or component or portion thereof, is used to treat a disease, disorder, or condition listed in any of Tables 18-23. In some embodiments, an LNP or conjugate described
 30 herein is used to treat a hematopoietic stem cell (HSC) disease, disorder, or condition, e.g., as listed in Table 18. In some embodiments, an LNP or conjugate described herein is used to

5 treat a kidney disease, disorder, or condition, e.g., as listed in Table 19. In some embodiments, an LNP or conjugate described herein is used to treat a liver disease, disorder, or condition, e.g., as listed in Table 20. In some embodiments, an LNP or conjugate described herein is used to treat a lung disease, disorder, or condition, e.g., as listed in Table 21. In some embodiments, an LMP or conjugate described herein is used to treat a skeletal muscle disease, disorder, or condition, e.g., as listed in Table 22. In some embodiments, an LNP or conjugate described herein is used to treat a skin disease, disorder, or condition, e.g., as listed in Table 23.

Table 18: Exemplary diseases treated with LNPs of the disclosure

Disease	Gene Affected
Adrenoleukodystrophy (CALD)	ABCD1
Alpha-mannosidosis	MAN2B1
Fanconi anemia	FANCA; FANCC; FANCG
Gaucher disease	GBA
Globoid cell leukodystrophy (Krabbe disease)	GALC
Hemophagocytic lymphohistiocytosis	PRF1; STX11; STXBP2; UNC13D
Malignant infantile osteopetrosis- autosomal recessive osteopetrosis	TCIRG1; Many genes implicated
Metachromatic leukodystrophy	ARSA; PSAP
MPS 1S (Scheie syndrome)	IDUA
MPS2	IDS
MPS7	GUSB
Mucopolidosis II	GNPTAB
Niemann-Pick disease A and B	SMPD1
Niemann-Pick disease C	NPC1
Pompe disease	GAA
Sickle cell disease (SCD)	HBB
Tay Sachs	HEXA
Thalassemia	HBB

Table 19: Kidney

Disease	Gene Affected
Congenital nephrotic syndrome	NPHS2
Cystinosis	CTNS

Table 20: Liver

Disease	Gene Affected
Acute intermittent porphyria	HMBS
Alagille syndrome	JAG1
Alpha-1 antitrypsin deficiency	SERPINA1
Carbamoyl phosphate synthetase I deficiency	CPS1
Citrullinemia I	ASS1
Crigler-Najjar	UGT1A1
Fabry	LPL
Familial chylomicronemia syndrome	GLA
Gaucher	GBE1
GSD1a	G6Pase
GSD IV	GBA
Heme A	F8
Heme B	F9
HoFH	LDLRAP1
Methylmalonic acidemia	MMUT
MPS II	IDS
MPS III	Type IIIa: SGSH Type IIIb: NAGLU Type IIIc: HGSNAT Type IIId: GNS S
MPS IV	Type IVA: GALNS Type IVB: GLB1
MPS VI	ARSB
MSUD	Type Ia: BCKDHA Type Ib: BCKDHB Type II: DBT
OTC Deficiency	OTC
Polycystic Liver Disease	PRKCSH
Pompe	GAA
Primary Hyperoxaluria 1	AGXT (HAO1 or LDHA for CRISPR)
Progressive familial intrahepatic cholestasis type 1	ATP8B1
Progressive familial intrahepatic cholestasis type 2	ABCB11
Progressive familial intrahepatic cholestasis type 3	ABCB4
Propionic acidemia	PCCB; PCCA
W'ison's Disease	ATP7B

Table 21: Lung

Disease	Gene Affected
Alpha-1 antitrypsin deficiency	SERPINA1
Cystic fibrosis	CFTR
Primary ciliary dyskinesia	DNAI1
Primary ciliary dyskinesia	DNAH5
Primary pulmonary hypertension I	BMPR2
Surfactant Protein B (SP-B) Deficiency (pulmonary surfactant metabolism dysfunction 1)	SFTPB

5

Table 22: Skeletal muscle

Disease	Gene Affected
Becker muscular dystrophy	DMD
Becker myotonia	CLCN1
Bethlem myopathy	COL6A2
Centronuclear myopathy, X-linked (motubular)	MTM1
Congenital myasthenic syndrome	CHRNE
Duchenne muscular dystrophy	DMD
Emery-Dreifuss muscular dystrophy, AD	LMNA
Limb-girdle muscular dystrophy 2A	CAPN3
Limb-girdle muscular dystrophy, type 2D	SGCA

10 **Table 23:** Skin

Disease	Gene Affected
Epidermolysis Bullosa Dystrophica Recessive (Hallopeau-Siemens)	COL7A1
Epidermolysis Bullosa Junctional	LAMB3
Epidermolytic Ichthyosis	KRT1; KRT10
Hailey-Hailey Disease	ATP2C1
Lamellar Ichthyosis/Nonbullous Congenital Ichthyosiform Erythroderma (ARCI)	TGM1
Netherton Syndrome	SPINK5

[0385] In some embodiments, an LNP or conjugate described herein is delivered to a tissue or cell from the cerebrum, cerebellum, adrenal gland, ovary, pancreas, parathyroid gland,

hypophysis, testis, thyroid gland, breast, spleen, tonsil, thymus, lymph node, bone marrow, lung, cardiac muscle, esophagus, stomach, small intestine, colon, liver, salivary gland, kidney, prostate, blood, or other cell or tissue type. In some embodiments, an LNP or conjugate described herein is used to treat a disease, such as a cancer, inflammatory disease, infectious disease, genetic defect, or other disease. A cancer can be cancer of the cerebrum, cerebellum, adrenal gland, ovary, pancreas, parathyroid gland, hypophysis, testis, thyroid gland, breast, spleen, tonsil, thymus, lymph node, bone marrow, lung, cardiac muscle, esophagus, stomach, small intestine, colon, liver, salivary gland, kidney, prostate, blood, or other cell or tissue type, and can include multiple cancers.

10 **[0386]** In some embodiments, an LNP or conjugate disclosed herein is delivered *ex vivo* to isolated cells (e.g., T cells or HSCs) of a subject (e.g., human patient) in need thereof. In some such embodiments, T cells or HSCs are first collected from the bone marrow of a subject. Following isolation of the desired cell population, the LNPs of the disclosure can be mixed with the cells, thereby resulting in effective transduction of the payload (e.g., gene
15 modifying system) into the T cells or HSCs. The modified T cells or HSCs can be administered to the patient.

[0387] In some embodiments where the cells are T cells, the cells can be activated prior to *ex vivo* delivery of the LNPs or conjugates. In other embodiments, the T cells are not activated prior to the delivery of the LNPs or conjugates *ex vivo* to T cells.

20 **[0388]** In other embodiments, an LNP, e.g., targeted LNP, as disclosed herein can be introduced into cells, tissues and/or multicellular organisms. In other embodiments, the LNPs, e.g., targeted LNPs, are delivered to the cells via mechanical means or physical means.

[0389] Formulation of protein therapeutics is described in Meyer (Ed.), *Therapeutic Protein Drug Products: Practical Approaches to formulation in the Laboratory, Manufacturing, and the Clinic*, Woodhead Publishing Series (2012).
25

[0390] In some embodiments, the LNP or conjugates described herein are administered *in vivo* to a patient in need thereof. The high level of transduction and protein expression following *in vivo* administration allows for sufficient levels of transduction and expression of proteins (e.g., gene modifying proteins), and hence can be used to treat various diseases
30 including but not limited to the diseases listed in **Tables 18-23**.

[0391] In some embodiments, an LNP or conjugate described herein is delivered to a tissue or cell from or in the bone marrow. In some such embodiments, the LNP, e.g., targeted LNP, is administered by parenteral administration (e.g., intravenous, intramuscular, subcutaneous, intradermal, epidural, intracerebral, intracerebroventricular, epicutaneous, nasal, intra-arterial, intra-articular, intracavernous, intraocular, intraosseous infusion, intraperitoneal, intrathecal, intrauterine, intravaginal, intravesical, perivascular, or transmucosal administration).

[0392] In some embodiments, an LNP or conjugate described herein delivers a therapeutic agent, as described above, to T cells, to HSCs (e.g., LT-HSCs), or to liver cells of a patient following in vivo administration. In some embodiments, an LNP, e.g., targeted LNP, of the disclosure delivers a gene modifying polypeptide, or a nucleic acid encoding the gene modifying polypeptide, to T cells, to HSCs (e.g., LT-HSCs), or to liver cells of a patient following in vivo administration. In some embodiments, an LNP, e.g., targeted LNP, of the disclosure delivers a gene modifying system to T cells, to HSCs (e.g., LT-HSCs), or to liver cells of a patient following in vivo administration. In some embodiments, an LNP, e.g., targeted LNP, of the disclosure delivers gene editing components to T cells, to HSCs (e.g., LT-HSCs), or to liver cells of a patient following in vivo administration. Exemplary gene editing systems are described above. Following in vivo administration, the nucleic acids encoding components of a system for modifying or altering a genome can be expressed, resulting in in vivo gene editing of the T cells, HSC (e.g., LT-HSC), or liver cells.

[0393] It has surprisingly been discovered that particular combinations of targeting moieties (e.g., antibodies, Fab fragments or scFvs), and ionizable lipids in particular LNP formulations (e.g., comprising particular amounts of helper lipid)) are capable of delivering therapeutic agents to more HSCs (e.g., LT-HSCs) and/or delivering greater amounts of therapeutic agents to T cells or to HSCs (e.g., LT-HSCs) following in vivo administration to a subject. In some cases, the LNPs described herein result in enhanced levels of modification or alteration of the genomic DNA of the T cells or HSCs, e.g., enhanced levels of gene editing in the T cells or HSCs. In some embodiments, the percentage of T cells or HSCs exhibiting modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 20%. In some cases, the LNPs described herein result in enhanced levels of modification or alteration of the genomic DNA of T cells or HSCs, e.g., enhanced levels of gene editing in the T cells or HSCs. In some embodiments, the percentage

of T cells or HSCs having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 30%. In some embodiments, the percentage of T cells or HSCs having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 40%. In some
5 embodiments, the percentage of T cells or HSCs having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 50%. In some embodiments, the percentage of T cells or HSCs having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 60%. In some embodiments, the percentage of T cells or HSCs having modified or
10 altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 70%. In some embodiments, the percentage of T cells or HSCs (e.g., LT-HSCs) having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is from about 20% to about 50%, from about 20% to about 40%, or from about 20% to about 30%. In some embodiments, 10% more T
15 cells or HSCs have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 15% more T cells or HSCs have
20 modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 20% more T cells or HSCs have modified or
25 altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 25% more T cells or HSCs have modified or altered (e.g., edited)
30 genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 30% more T cells or HSCs have modified or altered (e.g., edited) genomic DNA

in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 35% more T cells or HSCs have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 40% more T cells or HSCs have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 45% more T cells or HSCs have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 50% more T cells or HSCs have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, the number of T cells or HSCs comprising modified or altered (e.g., edited) genomic DNA can be determined by DNA sequencing.

[0394] It has surprisingly been discovered that particular combinations of targeting moieties (e.g., antibodies, Fab fragments or scFvs), and ionizable lipids in particular LNP formulations (e.g., comprising particular amounts of helper lipid)) are capable of delivering therapeutic agents to more liver cells and/or delivering greater amounts of therapeutic agents to liver cells following in vivo administration to a subject. In some cases, the LNPs described herein result in enhanced levels of modification or alteration of the genomic DNA of the liver cells, e.g., enhanced levels of gene editing in the liver cells. In some embodiments, the percentage of liver cells exhibiting modified or altered (e.g., edited) genomic DNA in the

subject after administration of the LNPs, e.g., targeted LNPs, is at least 20%. In some cases, the LNPs described herein result in enhanced levels of modification or alteration of the genomic DNA of liver cells, e.g., enhanced levels of gene editing in the liver cells. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 30%. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 40%. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 50%. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 60%. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is at least 70%. In some embodiments, the percentage of liver cells having modified or altered (e.g., edited) genomic DNA in the subject after administration of the LNPs, e.g., targeted LNPs, is from about 20% to about 50%, from about 20% to about 40%, or from about 20% to about 30%. In some embodiments, 10% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 15% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 20% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 25% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene

modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 30% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 35% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 40% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 45% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, 50% more liver cells have modified or altered (e.g., edited) genomic DNA in the subject after administration of an LNP as described herein, e.g., a targeted LNP, comprising a therapeutic agent (e.g., a nucleic acid encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) relative to a baseline LNP (e.g., a LNP comprising V003) comprising the same therapeutic agent. In some embodiments, the number of liver cells comprising modified or altered (e.g., edited) genomic DNA can be determined by DNA sequencing.

[0395] In some embodiments, the LNPs described herein can deliver greater amounts of a payload, e.g., a therapeutic agent or a reporter, to a T cell or an HSC (e.g., an LT-HSC) when administered in vivo to a subject compared to a baseline LNP (e.g., V003). In some embodiments, the targeted LNPs described herein can deliver greater amounts of a payload, e.g., a therapeutic agent or a reporter, to a T cell or an HSC (e.g., an LT-HSC) when administered in vivo to a subject compared to a baseline LNP (e.g., an LNP comprising V003)

comprising the same targeting moieties as the targeted LNP. In some embodiments, the LNPs described herein, e.g., targeted LNPs, can deliver greater amounts of a nucleic acid, e.g., a nucleic acid (e.g., mRNA) encoding a gene modifying polypeptide or one or more nucleic acids comprising a gene modifying system), to a T cell or an HSC (e.g., an LT-HSC) when administered in vivo to a subject compared to a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., a therapeutic agent such as a nucleic acid (e.g., mRNA) encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) that is delivered to T cells or HSCs (e.g., LT-HSCs) in vivo using an LNP as described herein (e.g., a targeted LNP) is expressed at higher levels relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 10% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 15% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 20% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 25% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 30% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 35% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 40% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 45% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a

nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 50% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some embodiments, a nucleic acid payload (e.g., mRNA) that is delivered to T cells or HSCs in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits between 10% to 50%, between 20% to 40%, between 10% to 30%, or between 30% to 50% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, the level of expression of the payload can be determined by measuring protein levels, e.g., by flow cytometry. In some cases, for example, when the payload comprises a gene modifying polypeptide or gene modifying system, the expression level of the payload can be measured by determining the level of modification or alteration of the genomic DNA of the T cells or HSCs, e.g., by DNA sequencing.

[0396] In some embodiments, the LNPs described herein can deliver greater amounts of a payload, e.g., a therapeutic agent or a reporter, to a liver cell when administered in vivo to a subject compared to a baseline LNP (e.g., V003). In some embodiments, the targeted LNPs described herein can deliver greater amounts of a payload, e.g., a therapeutic agent or a reporter, to a liver cell when administered in vivo to a subject compared to a baseline LNP (e.g., an LNP comprising V003) comprising the same targeting moieties as the targeted LNP. In some embodiments, the LNPs described herein, e.g., targeted LNPs, can deliver greater amounts of a nucleic acid, e.g., a nucleic acid (e.g., mRNA) encoding a gene modifying polypeptide or one or more nucleic acids comprising a gene modifying system), to a liver cell when administered in vivo to a subject compared to a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., a therapeutic agent such as a nucleic acid (e.g., mRNA) encoding a gene modifying polypeptide or one or more nucleic acids encoding a gene modifying system) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) is expressed at higher levels relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 10% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 15% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is

delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 20% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 25%
5 higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 30% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to
10 liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 35% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 40% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to
15 liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 45% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits 50% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some embodiments, a nucleic acid payload (e.g., mRNA) that is delivered to liver cells in vivo using an LNP as described herein (e.g., a targeted LNP) exhibits between 10% to 50%, between 20% to 40%, between 10% to 30%, or between 30% to 50% higher expression relative to the same payload delivered using a baseline LNP (e.g., an LNP comprising V003). In some cases, the level of expression of the payload can be determined by
20 measuring protein levels, e.g., by flow cytometry. In some cases, for example, when the payload comprises a gene modifying polypeptide or gene modifying system, the expression level of the payload can be measured by determining the level of modification or alteration of the genomic DNA of the liver cells, e.g., by DNA sequencing.

30 [0397] The high level of in vivo gene editing following administration of the LNPs, e.g., targeted LNPs, of the disclosure enables effective treatment of particular diseases such as cancer, β -hemoglobinopathies, primary immunodeficiencies, and metabolic disorders. In some embodiments, the LNPs, e.g., targeted LNPs, of the disclosure can be used to treat sickle cell

diseases. In other embodiments, the LNP, e.g., targeted LNPs, of the disclosure can be used to treat beta thalassemia.

EXAMPLES

5 **[0398]** The invention is further illustrated by the following examples. The examples are provided for illustrative purposes only and are not to be construed as limiting the scope or content of the invention in any way.

Example 1: Synthesis of ionizable lipids

10 **[0399]** A number of ionizable lipids were synthesized for formulation in LNPs that were tested for delivery to HSPCs.

[0400] General: All reactions were run using anhydrous grade solvents under an atmosphere of nitrogen in flasks or vials with magnetic stirring, unless otherwise noted. Anhydrous solvents were from Wuxi EHS and used as received. Column silica gel chromatography using silica gel 100-200 plates, and compounds were visualized using
15 phosphomolybdic acid. Nuclear magnetic resonance (NMR) spectroscopy was performed either using 400 MHz spectrometer; chemical shifts are reported in δ parts per million (ppm) referenced to tetramethylsilane at $\delta = 0.00$ ppm for CDCl_3 samples, and residual solvent peak ($\delta = 7.27$ ppm) for CDCl_3 samples. Liquid Chromatography-Mass Spectrum (LC-MS) was performed using a detector (ESI*) using one of the following general methods.

20 **[0401] Method A** (20 min run): 60-100AB-20MIN_Phenyl-Hexyl-ELSD-3 : LC/MS (The gradient was 60% B in 0-0.4 min, 60-100% B in 0.4-17 min, 100% B in 17-20 min, 60% B in 0.01, the flow rate was 0.8 ml/min. Mobile phase A was 0.05% Trifluoroacetic Acid in water, mobile phase B was 0.05% Trifluoroacetic Acid in acetonitrile. The column used for chromatography was a Waters XSelect CSH Phenyl-Hexyl 150*4.6mm, 3.5 μm . Detection
25 methods are diode array (DAD) and evaporative light scattering (ELSD) detection as well as positive electrospray ionization. MS range was 100-2000.

[0402] Method B (22 min run): 50-100AB-22MIN-CSH_Phenyl-Hexyl-ELSD-3M : LC/MS (The gradient was 50% B in 0-0.4 min, 50-100% B in 0.4-17 min, 100% B in 17-20 min, 50% B in 0.01, the flow rate was 0.8 ml/min. Mobile phase A was 0.05% Trifluoroacetic
30 Acid in water, mobile phase B was 0.05% Trifluoroacetic Acid in acetonitrile. The column

used for chromatography was a Waters XSelect CSH Phenyl-Hexyl 150*4.6mm, 3.5 μ m. Detection methods are diode array (DAD) and evaporative light scattering (ELSD) detection as well as positive electrospray ionization. MS range was 100-2000.

5 **[0403] Method C** (10 min run): WUXIAB60-Phenyl-Hexyl-10MIN-1: LCMS (The gradient was 60% B in 0.4 min and 60-100% B at 0.4-8.0 min, hold at 100% B for 1 min, 100-0% B in 0.01 min, and then held at 60% B for 2min. Mobile phase A was 0.05% Trifluoroacetic Acid in water, mobile phase B was 0.05% Trifluoroacetic Acid in acetonitrile. The column used for chromatography was a Waters XSelect CSH Phenyl-Hexyl 50*2.1mm, 3.5 μ m. Detection methods are ELSD. The ESI MS range was 100-2000.

10 **[0404] Method D** (10 min run): WUXIAB40-Phenyl -Hexyl_10min : LC/MS (The gradient was 40% B in 0-0.4 min, 40-100% B in 0.4-8 min, 100% B in 8-9 min, 40% B in 9.01-10 min, the flow rate was 0.8 ml/min. Mobile phase A was 0.05% Trifluoroacetic Acid in water, mobile phase B was 0.05% Trifluoroacetic Acid in acetonitrile. The column used for chromatography was a Waters XSelect CSH Phenyl-Hexyl 50*2.1mm, 3.5 μ m. Detection
15 methods are diode array (DAD) and evaporative light scattering (ELSD) detection as well as positive electrospray ionization. MS range was 100-2000.

List of abbreviations

DCM: dichloromethane
DIEA: N,N-diisopropylethylamine.
20 DMAP: 4-(dimethylamino)pyridine
EDCI: 3-(ethyliminomethylideneamino)propyl-dimethylazanium;chloride
Eq or eqv: equivalents
Py: pyridine

25 **Lipid 154**

General procedure for preparation of compound 5

[0405] Compound 5 was prepared according to the synthetic scheme shown in **FIG. 9A**. KOH (25.85 g, 460.8 mmol, 3.0 eq) was dissolved in H₂O (150 mL), the resulting solution was added to a solution of 4,4-dioctoxybutanenitrile (50 g, 153.6 mmol, 1.0 eq) in EtOH (150
30 mL). The mixture was stirred at 110°C for 20 hr. TLC (Petroleum ether: Ethyl acetate = 10:1, the starting material R_f = 0.7, product R_f = 0.5) showed the reaction was completed. The reaction mixture was concentrated under reduced pressure to remove EtOH. The residue was

diluted with water (10 mL). The aqueous layer was acidified to pH=5 with (HCl, 1 M, 50 mL), and extracted with EtOAc (50 mL * 3). The combined organic layers were washed with brine (50 mL), dried over anhydrous Na₂SO₄, filtered and concentrated under reduced pressure to give compound 5 (50 g, crude) as colorless oil. There was no further purification, the product
5 was used directly in next step. ¹H NMR (400 MHz, CDCl₃-d) δ 5.28 (br s, 2H), 4.63 (br s, 3H), 4.45 (br s, 1H), 3.57 - 3.37 (m, 4H), 2.13 (br t, J = 7.6 Hz, 2H), 1.87 - 1.65 (m, 2H), 1.55 (br d, J = 6.0 Hz, 4H), 1.36 - 1.25 (m, 19H), 0.96 - 0.82 (m, 6H).

General procedure for preparation of compound 29

[0406] Compound 29 was prepared according to the synthetic scheme shown in FIG.
10 **9B.** To a solution of octanedioic acid (10 g, 57.41 mmol, 1.0 eq) in DCM (100 mL) was added DIEA (11.13 g, 86.11 mmol, 15.0 mL, 1.5 eq), DMAP (1.40 g, 11.48 mmol, 0.2 eq) and EDCI (16.51 g, 86.11 mmol, 1.5 eq) in order. Then tridecan-7-ol (11.50 g, 57.41 mmol, 1.0 eq) was added by one portion at 20°C. The mixture was degassed and purged with N₂ for 3 times and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 5:1,
15 product R_f = 0.5) showed the reaction was completed. The residue was extracted with DCM (50 mL * 3) and H₂O (50 mL). The combined organic layers were washed with brine (50 mL), dried over anhydrous Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, Petroleum ether: Ethyl acetate = 30:1 to 1:1) to give 8-(1-hexylheptoxy)-8-oxo-octanoic acid (compound 29) (10 g,
20 28.05 mmol, 48.9% yield) as a colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, J = 6.2 Hz, 1H), 2.39 - 2.25 (m, 4H), 1.70 - 1.57 (m, 4H), 1.56 - 1.51 (m, 2H), 1.41 - 1.22 (m, 22H), 0.88 (t, J = 6.8 Hz, 6H).

General procedure for preparation of compound 30

[0407] Compound 30 was prepared according to the synthetic scheme shown in FIG.
25 **9C.** To a mixture of compound 29 (8 g, 22.44 mmol, 1.0 eq) in DCM (80 mL) was added DIEA (4.35 g, 33.66 mmol, 5.86 mL, 1.5 eq), DMAP (548.2 mg, 4.49 mmol, 0.2 eq) and EDCI (6.45 g, 33.66 mmol, 1.5 eq). Then (2,2-dimethyl-1,3-dioxan-5-yl)methanol (3.94 g, 26.93 mmol, 1.2 eq) was added. The mixture was degassed and purged with N₂ for 3 times and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 5:1, the starting
30 material R_f = 0.3, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with H₂O (50 mL) and extracted with DCM (100 mL * 2). The combined organic layers were washed with brine (50 mL * 2), dried over Na₂SO₄, filtered and concentrated under

reduced pressure. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 100:1 to 3:1) to give O1-[(2,2-dimethyl-1,3-dioxan-5-yl)methyl] O8-(1-hexylheptyl) octanedioate (compound 30) (9.7 g, 20.01 mmol, 89.2% yield) as colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, *J* = 6.2 Hz, 1H), 4.17 (d, *J* = 7.0 Hz, 2H), 4.00 (dd, *J* = 4.0, 12.0 Hz, 2H), 3.74 (dd, *J* = 5.8, 12.0 Hz, 2H), 2.30 (td, *J* = 7.6, 12.6 Hz, 4H), 2.01 (br s, 1H), 1.69 - 1.64 (m, 1H), 1.62 - 1.48 (m, 6H), 1.43 (d, *J* = 13.2 Hz, 6H), 1.37 - 1.22 (m, 20H), 0.93 - 0.84 (m, 6H).

General procedure for preparation of compound 31

[0408] Compound 31 was prepared according to the synthetic scheme shown in FIG. 9E. To a solution of O1-[(2,2-dimethyl-1,3-dioxan-5-yl)methyl] O8-(1-hexylheptyl) octanedioate (7.7 g, 15.89 mmol, 1.0 eq) in isopropanol (39 mL) was added HCl (1 M, 39 mL). The mixture was stirred at 15°C for 1 hr under N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 1:1, the starting material R_f = 0.7, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with saturated NaHCO₃ solution (60 mL) and extracted with DCM (50 mL * 2). The combined organic layers were washed with brine (50 mL * 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = from 50:1 to 10:1 to remove impurities, from 5:1 to 1:1 to get product) to give O8-(1-hexylheptyl) O1-[3-hydroxy-2-(hydroxymethyl)propyl] octanedioate (compound 31) (9.7 g, crude) as colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, *J* = 6.4 Hz, 1H), 4.27 (d, *J* = 6.4 Hz, 2H), 3.78 (dq, *J* = 5.2, 11.1 Hz, 4H), 2.40 - 2.24 (m, 6H), 2.09 - 2.00 (m, 1H), 1.69 - 1.64 (m, 2H), 1.62 - 1.57 (m, 1H), 1.51 (br d, *J* = 6.2 Hz, 4H), 1.38 - 1.22 (m, 21H), 0.93 - 0.84 (m, 6H).

General procedure for preparation of compound 32

[0409] Compound 32 was prepared according to the synthetic scheme shown in FIG. 9F. To a solution of 4,4-diethoxybutanoic acid (7.25 g, 21.05 mmol, 1.2 eq) in DCM (80 mL) was added DIEA (4.53 g, 35.08 mmol, 6.11 mL, 2.0 eq), DMAP (428.6 mg, 3.51 mmol, 0.2 eq) and EDCI (6.73 g, 35.08 mmol, 2.0 eq). Then compound 31 (7.8 g, 17.54 mmol, 1.0 eq) was added. The resulting mixture was degassed and purged with N₂ for 3 times and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 3:1, the starting material R_f = 0.3, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with H₂O 50 mL and extracted with DCM (50 mL * 2). The combined organic layers were washed with brine (50 mL * 2), dried over Na₂SO₄, filtered and concentrated under

reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 100:1 to 3:1) to give compound 32 (8.1 g, 10.50 mmol, 59.9% yield) as colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, *J* = 6.4 Hz, 1H), 4.50 (t, *J* = 5.6 Hz, 1H), 4.23 - 4.14 (m, 4H), 3.65 - 3.54 (m, 4H), 3.41 (td, *J* = 6.8, 9.3 Hz, 2H), 2.42 (t, *J* = 7.6 Hz, 2H), 2.35 - 2.17 (m, 6H), 1.98 - 1.91 (m, 2H), 1.69 - 1.65 (m, 1H), 1.62 - 1.47 (m, 10H), 1.38 - 1.22 (m, 42H), 0.93 - 0.85 (m, 12H).

General procedure for preparation of Lipid 154

[0410] Lipid 154 was prepared according to the synthetic scheme shown in **FIG. 11F**. To a mixture of 4-(diethylamino)butanoic acid (2.48 g, 15.56 mmol, 3.0 eq) in DCM (40 mL) was added DIEA (2.01 g, 15.56 mmol, 2.71 mL, 3.0 eq), DMAP (316.8 mg, 2.59 mmol, 0.5 eq) and EDCI (2.98 g, 15.56 mmol, 3.0 eq). Then compound 32 (4 g, 5.19 mmol, 1.0 eq) was added to the mixture. The resulting mixture was degassed and purged with N₂ for 3 times, and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Dichloromethane: Methanol = 10:1, the starting material R_f = 0.7, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with H₂O (50 mL) and extracted with DCM (50 mL * 2). The combined organic layers were washed with brine (50 mL * 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = from 80:1 to 10:1 to remove impurity, from 10:1 to 3:1 to get product) for 2 times to give Lipid 154 (3.52 g, 3.86 mmol, 74.4% yield) as colorless oil. LCMS (Method C): Rt 3.47 min, m/z calculated [M+H]⁺: 912.7, found: 912.8. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, *J* = 6.4 Hz, 1H), 4.49 (t, *J* = 5.6 Hz, 1H), 4.13 (d, *J* = 5.8 Hz, 6H), 3.57 (td, *J* = 6.8, 9.2 Hz, 2H), 3.41 (td, *J* = 6.8, 9.2 Hz, 2H), 2.54 - 2.26 (m, 15H), 1.97 - 1.90 (m, 2H), 1.81 - 1.70 (m, 2H), 1.68 - 1.47 (m, 14H), 1.38 - 1.22 (m, 42H), 1.01 (t, *J* = 7.2 Hz, 6H), 0.92 - 0.85 (m, 12H).

Lipid 155

General procedure for preparation of compound 34

[0411] Compound 34 was prepared according to the synthetic scheme shown in **FIG. 10A**. In a 100 mL round-bottom flask equipped with a stir bar, octanedioic acid (5.0 g, 28.70 mmol, 1.0 eq) was dissolved in DCM (50 mL). DIEA (5.56 g, 43.06 mmol, 7.50 mL, 1.5 eq), DMAP (701.3 mg, 5.74 mmol, 0.2 eq) and EDCI (8.25 g, 43.06 mmol, 1.5 eq) was added in order. Then undecan-3-ol (4.95 g, 28.70 mmol, 1.0 eq) was added to the mixture, the mixture was degassed and purged with N₂ for 3 times. The mixture was stirred at 20°C for 16 hr under

N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 5:1, the starting material R_f = 0.8, product R_f = 0.5) showed undecan-3-ol was consumed completely and one major new spot was detected. The reaction mixture was diluted with H₂O (50 mL) and extracted with DCM (50 mL * 2). The combined organic layers were washed with brine (50 mL * 2), dried over Na₂SO₄,
5 filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 30:1 to 1:1) to give 8-(1-ethylnonyl)-8-oxo-octanoic acid (compound 34) (3.0 g, 9.13 mmol, 31.8% yield) as colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.82 (t, *J* = 6.2 Hz, 1H), 2.31 (td, *J* = 7.6, 12.1 Hz, 4H), 1.69 - 1.65 (m, 1H), 1.73 - 1.46 (m, 7H), 1.43 - 1.23 (m, 17H), 0.93 - 0.83 (m, 6H).

10 **General procedure for preparation of compound 35**

[0412] Compound 35 was prepared according to the synthetic scheme shown in FIG. 10B. In a 100 mL round-bottom flask equipped with a stir bar, compound 34 (2.5 g, 7.61 mmol, 1.0 eq) was dissolved in DCM (25 mL). DIEA (1.48 g, 11.42 mmol, 1.99 mL, 1.5 eq), DMAP (186.0 mg, 1.52 mmol, 0.2 eq) and EDCI (2.19 g, 11.42 mmol, 1.5 eq) was added in order.
15 Then (2,2-dimethyl-1,3-dioxan-5-yl)methanol (1.34 g, 9.13 mmol, 1.2 eq) was added to the mixture by one portion. The mixture was degassed and purged with N₂ for 3 times and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Petroleum ether: Ethyl acetate = 5:1, the starting material R_f = 0.3, product R_f = 0.5) showed starting material was consumed completely and one major new spot was detected. The reaction mixture was diluted with H₂O (25 mL) and
20 extracted with DCM (25 mL * 2). The combined organic layers were washed with brine (25 mL * 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 100:1 to 14:1) to give O1-[(2,2-dimethyl-1,3-dioxan-5-yl)methyl] O8-(1-ethylnonyl) octanedioate (compound 35) (3.1 g, 6.31 mmol, 83.0% yield) as colorless oil. ¹H NMR (400
25 MHz, CDCl₃-d) δ 4.82 (quin, *J* = 6.2 Hz, 1H), 4.17 (d, *J* = 7.0 Hz, 2H), 4.00 (dd, *J* = 4.0, 12.0 Hz, 2H), 3.74 (dd, *J* = 5.8, 12.0 Hz, 2H), 2.35 - 2.26 (m, 4H), 2.05 - 1.96 (m, 1H), 1.68 - 1.47 (m, 9H), 1.43 (d, *J* = 13.4 Hz, 6H), 1.37 - 1.21 (m, 16H), 0.93 - 0.83 (m, 6H).

General procedure for preparation of compound 36

[0413] Compound 36 was prepared according to the synthetic scheme shown in FIG. 10C. In a 50 mL round-bottom flask equipped with a stir bar, compound 35 (1.9 g, 4.16 mmol, 1.0 eq) was dissolved in IPA (10 mL) and HCl (1 M, 10 mL). The mixture was degassed and purged with N₂ for 3 times, and then the mixture was stirred at 20°C for 2 hr under N₂
30

atmosphere. TLC (Petroleum ether: Ethyl acetate = 1:1, the starting material R_f = 0.8, product R_f = 0.5) showed starting material was consumed completely and one major new spot was detected. Adjust the pH to basic (pH=8~9) with NaHCO_3 (10 mL). The reaction mixture was diluted with H_2O (20mL) and extracted with DCM (20 mL * 2). The combined organic layers were washed with brine (20 mL * 2), dried over Na_2SO_4 , filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO_2 , petroleum ether: ethyl acetate = 100:1 to 13:1) to give O8-(1-ethylnonyl) O1-[3-hydroxy-2-(hydroxymethyl)propyl] octanedioate (compound 36) (1.6 g, 3.84 mmol, 92.3% yield, 100% purity) as colorless oil. ^1H NMR (400 MHz, CDCl_3 - d) δ 4.82 (quin, J = 6.2 Hz, 1H), 4.27 (d, J = 6.4 Hz, 2H), 3.78 (dq, J = 5.2, 11.0 Hz, 4H), 2.32 (td, J = 7.4, 17.8 Hz, 6H), 2.05 (td, J = 5.6, 10.9 Hz, 1H), 1.69 - 1.47 (m, 9H), 1.39 - 1.22 (m, 17H), 0.92 - 0.84 (m, 6H).

General procedure for preparation of compound 37

[0414] Compound 37 was prepared according to the synthetic scheme shown in FIG. 10D. In a 50 mL round-bottom flask equipped with a stir bar, compound 36 (900 mg, 2.16 mmol, 1.0 eq) was dissolved in DCM (10 mL). DIEA (418.8 mg, 3.24 mmol, 564.5 μL , 1.5 eq), DMAP (52.79 mg, 432.1 μmol , 0.2 eq) and EDCI (621.2 mg, 3.24 mmol, 1.5 eq) was added in order. Then 4,4-diethoxybutanoic acid (893.2 mg, 2.59 mmol, 1.2 eq) was added to the mixture. The mixture was degassed and purged with N_2 for 3 times, and stirred at 20°C for 12 hr under N_2 atmosphere. TLC (Petroleum ether: Ethyl acetate = 5:1, the starting material R_f = 0.5, product R_f = 0.4) showed the reaction was completed. The reaction mixture was diluted with H_2O (10 mL) and extracted with DCM (10 mL * 2). The combined organic layers were washed with brine (10 mL * 2), dried over Na_2SO_4 , filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO_2 , petroleum ether: ethyl acetate = 100: 1 to 13: 1) to give compound 37 (1.4 g, 1.86 mmol, 86.3% yield, 98.9% purity) as colorless oil. ^1H NMR (400 MHz, CDCl_3 - d) δ 4.82 (quin, J = 6.2 Hz, 1H), 4.50 (t, J = 5.4 Hz, 1H), 4.23 - 4.13 (m, 4H), 3.66 - 3.53 (m, 4H), 3.41 (td, J = 6.8, 9.2 Hz, 2H), 2.42 (t, J = 7.4 Hz, 2H), 2.36 - 2.16 (m, 6H), 1.98 - 1.90 (m, 2H), 1.68 - 1.48 (m, 14H), 1.40 - 1.21 (m, 37H), 0.94 - 0.85 (m, 12H).

General procedure for preparation of Lipid 155

[0415] Lipid 155 was prepared according to the synthetic scheme shown in FIG. 10E. In a 50 mL round-bottom flask equipped with a stir bar, compound 37 (600 mg, 807.4 μmol , 1.0 eq) was dissolved in DCM (10 mL). DIEA (156.5 mg, 1.21 mmol, 211.0 μL , 1.5 eq), DMAP

(19.73 mg, 161.5 μ mol, 0.2 eq) and EDCI (232.2 mg, 1.21 mmol, 1.5 eq) was added in order. Then 4-(diethylamino)butanoic acid (385.7 mg, 2.42 mmol, 3.0 eq) was added to the mixture. The mixture was degassed and purged with N₂ for 3 times and stirred at 20°C for 12 hr under N₂ atmosphere. TLC (Dichloromethane: Methanol = 10:1, the starting material R_f = 0.8, product R_f = 0.5) showed the reaction was completed. The mixture was washed by H₂O (10 mL * 3) and extracted with DCM (10 mL * 3). The organic phase was separated, washed with Na₂CO₃ solution (10 mL * 5), brine (10 mL * 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = from 100:1 to 13:1 to remove impurities, from 13:1 to 1:1 to get product) for 3 times to give Lipid 155 (217 mg, 245.4 μ mol, 30.4% yield, 100% purity) as colorless oil. LCMS (Method A): Rt 11.04 min, m/z calculated [M+H]⁺: 884.7, found: 884.7. ¹H NMR (400 MHz, CDCl₃-d) δ 4.82 (quin, *J* = 6.2 Hz, 1H), 4.49 (t, *J* = 5.4 Hz, 1H), 4.13 (d, *J* = 5.8 Hz, 6H), 3.57 (td, *J* = 6.8, 9.2 Hz, 2H), 3.41 (td, *J* = 6.8, 9.2 Hz, 2H), 2.56 - 2.26 (m, 15H), 1.97 - 1.89 (m, 2H), 1.81 - 1.72 (m, 2H), 1.67 - 1.47 (m, 13H), 1.39 - 1.21 (m, 37H), 1.01 (br t, *J* = 6.8 Hz, 6H), 0.93 - 0.84 (m, 12H).

Lipid 232

General procedure for preparation of compound 29

[0416] Compound **29** was prepared according to the synthetic scheme shown in FIG. 24A. To a solution of acid **27** (50 g, 287.0 mmol, 1 eq) in CH₂Cl₂ (500 mL) was added DIEA (55.65 g, 430.6 mmol, 74.99 mL, 1.5 eq), DMAP (7.01 g, 57.41 mmol, 0.2 eq) and EDCI (82.54 g, 430.6 mmol, 1.5 eq) at 20°C. Then alcohol **28** (57.51 g, 287.0 mmol, 1 eq) was added to the reaction mixture. The mixture was stirred at 20°C for 12 hr. LCMS and TLC (Petroleum ether: Ethyl acetate= 5: 1, R_f = 0.5) showed reactant **27** was consumed completely and desired MS of compound **2** was detected. The residue was extracted with DCM (150 mL x 3) and H₂O (250 mL). The combined organic layers were washed with brine (200 mL), dried over anhydrous Na₂SO₄, filtered and concentrated under reduced pressure to give compound **2** (100 g, crude) as light-yellow oil. The crude material was used directly in next step without purification. LCMS: m/z calculated [M+Na]⁺: 379.3, found: 379.3.

General procedure for preparation of compound 30

[0417] Compound **30** was prepared according to the synthetic scheme shown in FIG. 24B. To a solution of compound **29** (80 g, 224.4 mmol, 1 eq) in CH₂Cl₂ (800 mL) was added DIPEA (30.45 g, 235.6 mmol, 41.04 mL, 1.05 eq), DMAP (2.74 g, 22.44 mmol, 0.1 eq) and

EDCI (45.16 g, 235.6 mmol, 1.05 eq) at 20°C. Then alcohol **1** (26.24 g, 179.50 mmol, 0.8 eq) was added to the reaction mixture. The mixture was stirred at 20°C for 1 hr. LCMS and TLC (Petroleum ether: Ethyl acetate = 5: 1, the starting material R_f = 0.3, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with H₂O (500 mL) and extracted with CH₂Cl₂ (150 mL x 2). The combined organic layers were washed with brine (200 mL x 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give compound **30** (100 g, crude) as light-yellow oil. The crude material was used directly in next step without purification. LCMS: m/z calculated [M+Na]⁺: 507.4, found: 507.4.

General procedure for preparation of compound 31

[0418] Compound **31** was prepared according to the synthetic scheme shown in FIG. 24C. To a solution of compound **30** (100 g, 206.3 mmol, 1 eq) in IPA (500 mL) was added HCl (1M, 500 mL). The mixture was stirred at 20°C for 0.5 hr under N₂ atmosphere. LCMS and TLC (Petroleum ether: Ethyl acetate = 1: 1, the starting material R_f = 0.7, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with NaHCO₃ in aqueous solution 550 mL and extracted with CH₂Cl₂ (100mL x 2). The combined organic layers were washed with brine (100 mL x 2), dried over Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 50:1, 30:1, 10:1, product 5:1 to 1:1) to give compound **31** (31.5 g, 70.77 mmol, 34.30% yield, 99.9% purity) as colorless oil. ¹H NMR (400 MHz, CDCl₃-d) δ 4.87 (t, J = 6.2 Hz, 1H), 4.27 (d, J = 6.4 Hz, 2H), 3.78 (dq, J = 5.2, 11.0 Hz, 4H), 2.40 - 2.24 (m, 6H), 2.10 - 2.00 (m, 1H), 1.70 - 1.57 (m, 5H), 1.51 (br d, J = 6.0 Hz, 4H), 1.40 - 1.30 (m, 6H), 1.27 (br s, 14H), 0.88 (t, J = 6.8 Hz, 6H). LCMS: m/z calculated [M+Na]⁺: 467.3, found: 467.1.

General procedure for preparation of compound 32

[0419] Compound **32** was prepared according to the synthetic scheme shown in FIG. 24D. To a solution of acid **5** (32.54 g, 94.46 mmol, 1.2 eq) in THF (400 mL) was added DIPEA (20.35 g, 157.4 mmol, 27.42 mL, 2 eq), DMAP (1.92 g, 15.74 mmol, 0.2 eq) and EDCI (30.18 g, 157.4 mmol, 2 eq) at 20°C. Then compound **31** (35 g, 78.71 mmol, 1 eq) was added to the reaction mixture. The mixture was stirred at 20°C for 1 hr. LCMS and TLC (Petroleum ether: Ethyl acetate = 3:1, the starting material R_f = 0.3, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted with H₂O (100 mL) and extracted with CH₂Cl₂ (100mL x 2). The combined organic layers were washed with brine (100 mL x 2), dried over

Na₂SO₄, filtered and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 20:1 to 1:1) to give compound **32** (22 g, 28.53 mmol, 36.24% yield, 100% purity) as colorless oil. ¹H NMR (400 MHz, CDCl₃-*d*) δ 4.87 (t, *J* = 6.2 Hz, 1H), 4.50 (t, *J* = 5.4 Hz, 1H), 4.23 - 4.13 (m, 4H), 3.65 - 3.54 (m, 4H), 3.41 (td, *J* = 6.8, 9.2 Hz, 2H), 2.42 (t, *J* = 7.6 Hz, 2H), 2.36 - 2.17 (m, 6H), 1.99 - 1.90 (m, 2H), 1.69 - 1.58 (m, 5H), 1.55 - 1.46 (m, 6H), 1.38 - 1.23 (m, 40H), 0.94 - 0.84 (m, 12H). LCMS: *m/z* calculated [M+Na]⁺: 793.6, found: 793.8.

General procedure for preparation of Lipid 232

[0420] Lipid 232 was prepared according to the synthetic scheme shown in **FIG. 24E**.
10 To a solution of compound **32** (20 g, 25.94 mmol, 1 eq) in MeCN (200 mL) was added pyridine (3.28 g, 41.50 mmol, 3.35 mL, 1.6 eq), DMAP (950.5 mg, 7.78 mmol, 0.3 eq) and chloroformate **5A** (8.36 g, 41.50 mmol, 1.6 eq) at 20°C. The mixture was stirred at 20°C for 0.5 hr. LCMS and TLC (Petroleum ether: Ethyl acetate = 3: 1, R_f = 0.7) showed reactant **32** was consumed completely and desired MS of **42** was detected. The reaction mixture was used
15 in situ to next step without any workup. LCMS: *m/z* calculated [M+Na]⁺: 958.6, found: 958.8.

[0421] To the resulting mixture of intermediate **42** in MeCN (total ~250 mL) was added alcohol **6A** (10.09 g, 76.90 mmol, 11.47 mL, 3 eq) at 20°C. Then the mixture was stirred at 20°C for 16 hr. LCMS and TLC (Dichloromethane: Methanol = 10:1, the starting material R_f = 0.7, product R_f = 0.5) showed the reaction was completed. The reaction mixture was diluted
20 with H₂O (200 mL) and extracted with CH₂Cl₂ (100 mL * 2). The combined organic layers were washed with Na₂CO₃(aq) (200 mL x 2) and brine (200 mL x 2), dried over Na₂SO₄, filtrated, and concentrated under reduced pressure to give a residue. The residue was purified by column chromatography (SiO₂, petroleum ether: ethyl acetate = 20:1, 5:1 to 3:1) for 2 times to give **Lipid 232** (14.48 g, 15.60 mmol, 60.85% yield, 100% purity ELSD, 96.0% purity CAD)
25 as colorless oil. ¹H NMR (400 MHz, CDCl₃-*d*) δ 4.87 (t, *J* = 6.2 Hz, 1H), 4.49 (t, *J* = 5.6 Hz, 1H), 4.22 - 4.13 (m, 8H), 3.60 - 3.53 (m, 2H), 3.44 - 3.37 (m, 2H), 2.51 (q, *J* = 6.8 Hz, 6H), 2.45 - 2.37 (m, 3H), 2.30 (td, *J* = 7.6, 11.4 Hz, 4H), 1.97 - 1.88 (m, 2H), 1.85 - 1.77 (m, 2H), 1.65 - 1.49 (m, 12H), 1.36 - 1.23 (m, 40H), 1.01 (t, *J* = 7.2 Hz, 6H), 0.92 - 0.85 (m, 12H). LCMS: *m/z* calculated [M+H]⁺: 928.7, found: 928.6.

30 Example 2: Screening of Base LNPs comprising novel ionizable lipids for HSPC delivery in C57BL/6 Mice

[0422] A general screening paradigm was developed to facilitate the identification of ionizable lipids that can be formulated in LNPs, including targeted LNPs (tLNPs), to enhance delivery of a payload to HSPCs in vivo in humans. The paradigm is as follows: a) screen “base LNPs” that each have an identical formulation except for the ionizable lipid (selected from the ionizable lipids in **Table 1**) incorporated in the formulation in C57BL/6 mice to identify those that exhibit enhanced delivery of a reporter GFP mRNA to mouse LSK cells (HSCs) and LT-HSCs relative to a base LNP comprising a baseline control ionizable lipid (the V003 lipid); b) conjugate a targeting moiety (as described herein) that can bind to a protein on the surface of human HSPCs/LT-HSCs to those base LNPs exhibiting higher HSC delivery in C57BL/6 mice relative to the baseline control LNP to create and screen for targeted LNPs (tLNPs comprising the base LNPs plus targeting moiety) exhibiting higher delivery of a reporter GFP mRNA to human HSPCs/LT-HSCs engrafted in NBSGW mice; c) optionally optimize the LNP formulations (e.g., the relative proportions of LNP lipid components) of those tLNPs exhibiting higher delivery to human HSPCs/LT-HSCs in NBSGW mice to further enhance delivery of the reporter to human HSPCs/LT-HSCs engrafted in NBSGW mice; and d) then test those optimized tLNPs exhibiting the highest HSPC/LT-HSC delivery in the mouse models in nonhuman primates (NHPs) to evaluate their ability to deliver GFP mRNA to primate HSPCs/LT-HSCs in vivo.

[0423] **Ionizable Lipid Synthesis:** Ionizable lipids were produced according to or by methods similar to those described in **Example 1**.

[0424] **Base LNP Production:** Base LNPs were formulated with an ionizable lipid from **Table 1**, or with the V003 ionizable lipid. LNPs were prepared using molar ratios of either 47% ionizable lipid: 42.5% cholesterol: 8.0% DSPC: 2.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO or 47% ionizable lipid: 43.5% cholesterol: 8.0% DSPC: 1.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO using a N:P of 6. LNPs were formulated using microfluidic mixing using a Precision Nanosystems NanoAssemblr Ignite according to the manufacturer’s protocol. A ratio of aqueous: organic solvent of 3:1 was maintained during mixing by using differential flow rates. After mixing, the LNPs were collected, diluted in citrate buffered solution pH = 6.0 (approximately 3:1 v/v). After 15 minutes, the diluted LNP solution was neutralized using 1 vol% Tris HCl pH = 8.0. After neutralization each LNP formulated was characterized by dynamic light scattering (DLS) to measure the size (e.g., diameter) and polydispersity (PDI). Following neutralization, buffer exchange was conducted on the solution

using dialysis in sucrose buffer (pH = 7.6, 9.0 wt% sucrose, 50 mM Tris, 50 mM salt) at 4°C. The buffer was replaced at 2 hours and dialysis continued overnight against a 10kDa filter. Following removal from the dialysis cassettes, each LNP formulated was characterized by DLS to measure the size and PDI. Each LNP formulation was concentrated using 100kDa Amicon
5 Filters at 4°C and 1500 RCF. The concentrated solution was then sterile filtered using a 0.2 µm syringe filter and measured using DLS. The filtered LNP samples were evaluated for mRNA content using the Ribogreen assay. Additionally, the pKa of LNPs was measured using the 2-(p-toluidinyl)naphthalene-6-sulfonic acid (TNS) assay. LNP solutions were diluted to the desired concentration and stored at -80°C for future use.

10 **[0425] Characterization by DLS:** LNP hydrodynamic diameter and polydispersity index (PDI) were measured using single cell dynamic light scattering (DLS) (Malvern Zetasizer Pro Red Label), high throughput DLS (Wyatt DynaPro III), or high resolution DLS (Malvern NanoSight NS300). Samples were prepared by mixing LNP mixture with 1X PBS (GIBCO 10010-023) at a 1:5 ratio before LNP concentration or a 1:20 ratio after LNP
15 concentration and then analyzed.

[0426] Characterization by RiboGreen: Concentration and encapsulation efficiency of the LNPs were measured using the Quant-iT RiboGreen RNA Kit. In brief, the LNP sample, mRNA standard, RiboGreen reagent, and Triton X-100 were all diluted in TE buffer to the appropriate concentration. Each sample well was comprised of 15 µL of LNP sample, 100 µL
20 of RiboGreen reagent, 50 µL of Triton X-100, and the remaining volume being TE buffer up to 200 µL total. Samples were incubated in the dark for 30 minutes and the fluorescence was measured using an excitation wavelength of 485 nm and an emission wavelength of 525 nm. Nucleic acid concentration for unlysed and lysed LNPs were found as per manufacturer's instruction with the encapsulated mRNA used as the known standard. The concentration of the
25 mRNA in the lysed LNP was used as the concentration of the LNP, and the encapsulation was given as the ratio of the difference in lysed and unlysed LNPs to the lysed LNPs.

[0427] Characterization by TNS: To determine the pKa of the LNPs, stock buffers of 25 mM PBS and 125 mM sodium chloride were prepared. The pH of the various buffers was adjusted to roughly pH 4 – 9.5 using hydrogen chloride and sodium hydroxide. Using 3
30 replicates of each pH value for each LNP sample, 386 µL of pH buffer, 10 µL of LNP sample (0.04 mg/mL mRNA), and 4 µL of TNS sample (0.6 mM) were added to each well and incubated in the dark for 30 minutes. Fluorescence of each sample was taken (Molecular

Devices SpectraMax i3x) using an excitation wavelength of 325 nm and an emission wavelength of 435 nm.

5 **[0428] Base Lipid (LNP) Screening in C57BL/6 mice:** A series of studies was conducted to compare base LNPs comprising different ionizable lipids selected from **Table 1** (above) to a base LNP comprising the V003 ionizable lipid (baseline LNP). The V003 ionizable lipid was used as a standard control to set a baseline delivery level to C57BL/6 LSK cells and LT-HSCs by which to compare the delivery levels of the test ionizable lipids in each study. Those base LNPs comprising an ionizable lipid from **Table 1** that demonstrated an increase in the percent delivery to and/or GFP expression in LSK cells and/or LT-HSCs relative to the baseline LNP comprising the V003 ionizable lipid were selected for further testing in humanized NBSGW mice.

15 **[0429]** Female C57BL/6 mice 6-8 weeks of age were purchased from Jackson Labs (Bar Harbor, ME; Catalog #000664). Following acclimation to the vivarium for 3-5 days, animals were randomized into groups of 4-5 and baseline weights were obtained. Animals were then treated via intravenous tail vein injection with 200 μ l of candidate base lipids formulated with GFP mRNA at a dose of 2 mg/kg. Following a 16-hour incubation, animals were euthanized via CO₂ asphyxiation, weighed and their femurs were collected and then used to isolate bone marrow cells. In some cases, serum samples were collected to measure ALT and AST levels as biomarkers for liver damage.

20 **[0430] Flow cytometric analysis of C57BL/6 mouse bone marrow:** Bone marrow cells (1×10^6 in 0.1 ml PBS) from each test mouse were subsequently aliquoted to individual wells in 96 well plates and antibody fluorochrome conjugate mixtures (Table 24) were added to the cells in a total volume of 0.1 ml. Following a 30-minute incubation at room temperature, the plates were centrifuged for 5 minutes at 500G and supernatants were removed. The cells were then resuspended in 0.2 ml PBS and evaluated on an Aligent NovoCyte Quanteon (2010098) or Cytex Northern Lights (N7-00008) flow cytometer using FlowJo software (Ashland, OR). Cells stained with relevant isotype control fluorochrome conjugated antibodies were used to set the negative control gates. HSCs and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-.

25

30

Table 24. Flow cytometry Antibodies

Marker	Fluorophore	Clone	Vendor	Catalog
Live/Dead	Yellow		ThermoFisher	L34959
Lineage	PE	145-2C11 ; RB6-8C5 ; RA3-6B2 ; Ter-119 ; M1/70	Biolegend	133303
Sca1	APC	D7	Biolegend	108112
CD150	BV421	TC15-12F12.2	Biolegend	115926
CD48	PerCP-Cy5.5	HM48-1	Biolegend	103422
CD45	PEcy7	30-F11	Biolegend	103114
CD117	BV605	2B8	Biolegend	105847
CD11b	BV650	M1/70	Biolegend	101259

[0431] **Summary of Base Lipid Screening Studies:** Table 25 provides the results of the screening studies. Base LNPs containing Lipid 154, Lipid 155, and Lipid 163 on average exhibited increased delivery to and/or increased GFP expression in LSK cells and/or LT-HSCs in C57BL/6 mice relative to a base LNP containing the V003 ionizable lipid (baseline LNP). **FIGs. 11A and 11B** show a subset of the results for Base LNPs containing high performing ionizable lipids. **FIGs. 11A to 11D** show enhanced transduction of GFP and/or enhanced GFP expression in LSK cells and in LT-HSCs using LNPs containing ionizable Lipid 154 or ionizable Lipid 155 relative to an LNP containing ionizable Lipid V003 (Study 2 in Table 25). These results indicated that LNPs formulated with these ionizable lipids could exhibit enhanced delivery of a payload gene to HSCs in vivo for therapeutic applications. A subset of these ionizable lipids were selected for testing in a humanized NBSGW mouse model harboring human HSPCs to determine whether they are capable of enhancing delivery to human cells.

15

Table 25: Results of Base LNP screening in C57BL/6 Mice.

Study #	Number	iLipid Tested	LNP Formulation (iLipid:chol:DSPC:PEG:PEG-TCO)	Average % GFP+ in LSK cells	Average GFP MFI in LSK cells	Average % GFP+ in LT-HSCs	Average GFP MFI in LT-HSCs
Study 1	2	Control	-	0	0	0	0
	5	V003	47:42.5:8:2:0.5	13.38%	9804	19.14%	10133
	5	Lipid 163	47:42.5:8:2:0.5	8.81%	7230	13.02%	10424
Study 2	2	Control	-	0	0	0	1755
	4	V003	47:42.5:8:2:0.5	19.9%	8807	24.11%	2106
	4	Lipid 154	47:42.5:8:2:0.5	39.17%	12209	47.75%	4796
	4	Lipid 155	47:42.5:8:2:0.5	27.72%	8926	37.8%	3546

Example 3: Screening of targeted LNPs (tLNPs) comprising novel ionizable lipids for HSPC delivery in a humanized mouse model

5

[0432] A subset of the base LNPs were modified to try to increase their ability to target human HSPCs in vivo, and then evaluated for their delivery GFP mRNA to HSPCs/LT-HSCs in a humanized mouse model in which human HSPCs were engrafted.

[0433] **Targeted LNP (tLNP) production:** Human HSCs express CD117 on their surface. Therefore, a targeting moiety comprising an antibody Fab that binds specifically to CD117 was conjugated to the surface of the best performing base LNPs of Example 2 to further enhance their delivery potential for human HSPCs in vivo. Specifically, the anti-CD117 targeting moiety was conjugated to the surface of the base LNPs of Example 2 formulated with the Lipid 154 and Lipid 155 ionizable lipids to produce targeted LNPs (tLNPs) to be tested in an NBSGW mouse model engrafted with human HSPCs.

[0434] Base LNPs were prepared using molar ratios of 47% ionizable lipid: 42.5% cholesterol: 8.0% DSPC: 2.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO using a N:P of 6. LNPs were formulated using microfluidic mixing using a Precision Nanosystems NanoAssemblr Ignite according to the manufacturer's protocol. A ratio of aqueous: organic solvent of 3:1 was maintained during mixing by using differential flow rates. After mixing, the LNPs were collected, diluted in citrate buffered solution pH = 6.0 (approximately 3:1 v/v). After 15 minutes, the diluted LNP solution was neutralized using 1 vol% Tris HCl pH = 8.0. After neutralization each LNP formulated was characterized by dynamic light scattering (DLS) to measure the size (e.g., diameter) and polydispersity (PDI). Following neutralization, buffer exchange was conducted on the solution using dialysis in sucrose buffer (pH = 7.6, 9.0 wt% sucrose, 50 mM Tris, 50 mM salt) at 4°C. The buffer was replaced at 2 hours and dialysis continued overnight against a 10kDa filter. Following removal from the dialysis cassettes, each LNP formulated was characterized by DLS to measure the size and PDI. Each LNP formulation was concentrated using 100kDa Amicon Filters at 4C and 1500 RCF. The concentrated solution was then sterile filtered using a 0.2 µm syringe filter and measured using DLS. The filtered LNP samples were evaluated for mRNA content using the Ribogreen assay. Additionally, the pKa of LNPs was measured using the 2-(p-toluidinyl)naphthalene-6-sulfonic acid (TNS) assay. LNP solutions were diluted to the desired concentration and stored at -80°C for future use.

[0435] An anti-CD117 Fab was produced with an additional sortase tag (LPETG) on the C terminus of the heavy chain. 10uM CD117 Fab-LPETG and 1mM Trisglycine methyltetrazine (GGG-meTz) were prepared in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.4). 2uM Sortase A5 was then added to initiate the reaction. The above solution was incubated at 30°C for 3 hours with gentle mixing. Excess reagents were removed, and buffer was exchanged to PBS containing 10mM EDTA by a 10K MWCO Amicon centrifugal UF device. To confirm a DOL (degree of labeling) of close to 1 meTz per antibody, sortase modified anti-CD117 Fab was allowed to react with 5eq AF594 containing TCO group in PBS for 2 hours at room temperature. After the reaction, excess dye was removed via Zeba desalting columns. DOL was determined by first measuring absorbances at A280 and A590 using a Nanodrop, and then calculating the DOL using extinction coefficients and correction factors. The resulting meTz-modified anti-CD117 Fab was added to the TCO modified-LNP solution described above for antibody-LNP conjugation, with a 1:1 mass ratio between antibody and mRNA (eGFP). The solution was incubated at room temperature for 2 hours and then at 4°C overnight. Any unreacted anti-CD117 Fab was removed by a 300K mPES TFF membrane. The resulting antibody-LNP product was further concentrated using a 100k MWCO Amicon centrifugal UF device. A ribogreen assay was then performed to quantify mRNA content and encapsulation efficiency.

[0436] **Characterization:** LNPs were characterized by DLS, RiboGreen, and TNS as described in Example 2.

[0437] **Targeted Lipid (tLNP) screening in “humanized” NBSGW mice:** A series of studies was conducted to compare anti-CD117 tLNPs formulated with the Lipid 154 or Lipid 155 to an ionizable lipid to a targeted LNP (tLNP) formulated with the V003 ionizable lipid. The V003 ionizable lipid was used as a standard control to set a baseline delivery level to human HSPCs/LT-HSCs engrafted in NBSGW mice by which to compare the delivery levels of the test ionizable lipids in each study. Those tLNPs demonstrating an increase in the percent delivery to and/or GFP expression in human HSPCs and/or LT-HSCs relative to the tLNP comprising the V003 ionizable lipid (baseline tLNP) were selected for further testing in nonhuman primates (NHPs) (see Example 5).

[0438] Briefly, female NBSGW mice 6-8 weeks of age (Catalog #026622) were purchased from Jackson Labs (Bar Harbor, ME). Following acclimation to the vivarium for 3-5 days, animals received a tail vein injection of 2×10^6 human CD34⁺ HSPCs (HemaCare

Division of Charles River Labs, Wilmington, MA) and the human cells were allowed to engraft for 12 weeks before the animals were used in screening studies. For each tLNP screening study, “humanized” NBSGW mice were randomized into groups of 4-5 animals and baseline weights were obtained. Animals were then treated via intravenous tail vein injection with 100
 5 ul of candidate tLNPs formulated with GFP mRNA at a dose of 2 mg/kg. Following a 16-hour incubation, animals were euthanized via CO₂ asphyxiation, weighed and their femurs and tibias were collected and the bone marrow was isolated. In some cases, serum samples were collected to measure ALT and AST levels as biomarkers for liver damage.

[0439] Flow cytometric analysis of “humanized” NBSGW mouse bone marrow:

10 Bone marrow cells (1×10^6 in 0.1 ml PBS) from each test mouse were subsequently aliquoted to individual wells in 96 well plates and antibody fluorochrome conjugate mixtures (Table 26) were added to the cells in a total of 0.1 ml. Following a 30-minute incubation at room temperature, the plates were centrifuged for 5 minutes at 500G and supernatants were removed. The cells were then resuspended in 0.2 ml PBS and evaluated on an Aligent NovoCyte
 15 Quanteon (2010098) or Cytex Northern Lights (N7-00008) flow cytometer using FlowJo software (Ashland OR). Cells stained with relevant isotype control fluorochrome conjugated antibodies were used to set the negative control gates. HSC and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-.

20 **Table 26. Flow cytometry antibodies**

Marker	Fluorophore	Clone	Vender	Catalog
mCD45	PE-Cy5	30-F11	Biolegends	103110
CD3	PerCp-Cy5.5	UCHT1	BD	560835
CD19	R718	H1B19	BD	567343
CD11b	PE/Dazzle 594	ICRF44	Biolegends	301348
hCD45	PE	HI30	Biolegends	304008
CD16	PE-Cy7	3G8	BD	557744
Live/Dead	Yellow		ThermoFisher	L34959
CD45RA	BV650	HI100	Biolegends	304136
CD34	BV421	561	Biolegends	343610
CD38	APC-Cy7	HIT2	Biolegends	303534
CD90	APC	5 E10	BD	561971
CD117	BV605	104D2	Biolegends	313218
CD71	BV786	L01.1	BD	745555

[0440] **Summary of Targeted LNP screening studies:** Table 27 provides the results of the screening studies. Anti-CD117 targeted LNPs containing Lipid 154 or Lipid 155 on average exhibited increased delivery to and/or increased GFP expression in human HSPCs and/or LT-HSCs in the humanized NBSGW mouse model relative to the targeted LNP containing the V003 ionizable lipid (baseline tLNP).

[0441] The results from Study 1 in Table 27 are depicted graphically in FIGs. 12A-D. These results show that tLNPs formulated with Lipid 154 or Lipid 155 exhibit higher GFP transduction and GFP expression levels in HSPCs and in LT-HSCs in the humanized mouse model relative to the baseline tLNP formulated with the V003 lipid. Lipid 155 tLNPs exhibited higher GFP transduction levels relative to the V003 baseline tLNP.

[0442] These results indicate that targeted LNPs formulated with a subset of ionizable lipids could exhibit enhanced delivery of a payload gene to HSPCs/LT-HSCs in vivo for therapeutic applications. Some of these ionizable lipids were selected for testing in nonhuman primates to determine whether they are capable of enhancing delivery in a larger animal model (see Example 5).

Table 27. Screening results of targeted LNP (tLNP) screening in humanized NBSGW mice.

Study #	Number	iLipid Tested	LNP Formulation (iLipid:chol:DSPC: PEG:PEG-TCO)	Average % GFP+ in HSPCs	Average GFP MFI in HSPCs	Average % GFP+ in LT-HSPCs	Average GFP MFI in LT-HSPCs
Study 1	2	Control	-	0.05%	1250	0	1340
	4	V003	47:42.5:8:2:0.5	23.65%	2750	41.8%	5405
	4	Lipid 154	47:42.5:8:2:0.5	53.48%	7577	81.38%	18279
	4	Lipid 155	47:42.5:8:2:0.5	34.03%	3322	56.47%	6906

Example 4: Screening of targeted LNPs (tLNPs) comprising novel ionizable lipids for HSPC delivery in a humanized mouse model

[0443] Another subset of the base LNPs (comprising Lipid 154) exhibiting enhanced delivery to HSPCs relative to the baseline control LNP in Example 2 were modified to try to increase their ability to target human HSPCs in vivo, and then evaluated for their delivery of GFP mRNA to HSPCs/LT-HSCs in a humanized mouse model in which human HSPCs were engrafted. Experiments were conducted as described in Example 3, but with increased quantities of helper lipid (22% DSPC or sphingomyelin). Results are shown in FIGs. 13A-D.

The results show that LNPs comprising 22% helper lipid deliver payload mRNA to a high percentage of HSPCs and LT-HSPCs, leading to GFP expression.

Example 5: Targeted LNPs formulated with Lipid 154 exhibit enhanced delivery of RNA payloads comprising exemplary gene modifying systems to HSPCs *in vivo*

5 **[0444]** Targeted LNPs formulated with the Lipid 154 ionizable lipid and conjugated to an anti-CD117 Fab fragment was tested for its ability to deliver exemplary gene modifying system payloads to HSCs *in vivo*. As set forth in the prior examples, these ionizable lipids were shown to enhance delivery of an mRNA payload encoding GFP reporter to HSCs *in vivo* previously. The gene modifying system comprised either an mRNA encoding a wild-type
10 Cas9-RT fusion protein and a gRNA specific for the first exon of the human Beta 2 Microglobulin (B2M) gene or a Cas9 nickase fused to an RT and a template RNA suitable to produce an insertion of five nucleotides (TAGTG) into the first exon of B2M. The introduction of the wtCas9-RT and gRNA leads to the introduction and repair of a DNA double-stranded break induced at this site which introduces the formation of small insertions and deletions
15 (indels), largely resulting in frameshifts and thus loss of B2M expression from the cell surface. Introduction of the nCas9-RT and B2M template RNA leads to a single-stranded DNA break and the insertion of the 5 nucleotides thereby generating a frameshift mutation which results in loss of expression of B2M from the cell surface. The tLNP was administered to humanized mice and evaluated for its effectiveness in delivering the payloads *in vivo* as measured by the
20 percentage of either B2M indels in HSPCs or the percentage of reads including the desired insertion of 5 nucleotides (TAGTG) in the first exon of B2M (assayed by Amp-Seq).

[0445] As described above, tLNP containing Cas9-RT and gRNA was formulated to encapsulate an mRNA encoding a wild-type Cas9 molecule fused to a reverse transcriptase domain (“wtCas9-RT”, SEQ ID NO: 141) and the gRNA (“B2M gRNA1”, SEQ ID NO: 143)
25 at a mass ratio of 1:1. The tLNP containing nCas9-RT and template RNA was formulated to encapsulate an mRNA encoding Cas9 nickase fused to an RT (“nCas9-RT”, SEQ ID NO: 142) and the template RNA (“B2M template RNA” SEQ ID NO: 144) at a mass ratio of 1:1. The gene modifying polypeptides comprising wtCas9-RT and nCas9-RT comprise the amino acid sequences indicated in Table 28.

30 **Table 28: Gene modifying polypeptide sequences**

Name	AA sequence	SEQ ID NO
wtCas9 RT	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRL SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSDTKADLRILIYALAHMI KFRGHFLIEGDLNPNDSVDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLN LAQIGDQYADFLFAAKNLSDAILSDILRVNTEITKAPLSASMIKREYDEHHQDLTLK LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGP LARGNSRFWMTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVL HSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLKKIKDKDFLDNEENEIDILEIVLTLT DREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKKLINGIRDKQSGKTILDFLK SDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIAKKGILQTV KVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEIGIKELGDSQILKE HPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDQVHIVPQSFLLKDDSDIN KVLTRSDKNRGSNDVPSSEVVKMKMKNYWRQLLNKLTQRKFDNLTKAERGGSL EDLKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSFLVSDFRK DFQFYKVRINNYHHAHDAYLNAVVGTALEIKKYPKLESEFVYGDYKVDVVRKMIAS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSPMPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKPYGGFDSPTV AYSVLVAVKEKGSKKLKSVEKELLGITIMERSSFEKNPIDLEAKGYKEVKKDIJIK PKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFYLASHYKELKSGSPEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVLADANLKDVL SAYNKHDKPIREQAENIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDGGAE AAAKEAAAEEAAKEAAKALEAAEAAKEAAEAAEAAEAAKAGGTAPLEEEYRL FLEAPIQNVTLLEQWKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITL EAKRSLRETIRKFRAGIIRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVTETI HPTVPNPYTLTLLSPDRIWYSVLDDKDAFFCPLAPESQLIFAEFWADAEESGQQL TWTRLPQGGFKNSPTLFNEALNRDLQGGFRDLHPSVSLQYVDDLLIAADTQAACLSAT RDLLMTLAEGLYRVSGKKAQLCQEEVYTLGFKIHKGSRSLSNSRTQAILQIPVPKTK RQVREFLGKIGYCRFLFIPGFAELAQPLAATRPNGNDPLVWGEKEEEAFQSLKLALTO PPALALPSLDKPPFLVEETSGAAGKGLTQALGPWKRPVAYLSKRLDPVAGWPR CLRAIAAAALLTREASKITFGQDIEITSSHNESSLRSPDQKWLNTARITQYQVLLDP PRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLSTRPDLTDQPLAQAEATLFTD GSSYIPDGKRYAGAAVVTLDVSIWAEPLPGTSAQAELIALTKALEWSKDKSVNIYT DSRYAFATLHVHGMIRERGWLTAGGKAIKNAPEILALLTAVWLPRKRVAVMHCKGH QKDDAPTSTGNRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFES PKKKAKVE	141
nCas9 -RT	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRL SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSDTKADLRILIYALAHMI KFRGHFLIEGDLNPNDSVDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLN LAQIGDQYADFLFAAKNLSDAILSDILRVNTEITKAPLSASMIKREYDEHHQDLTLK LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGP LARGNSRFWMTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVL HSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLKKIKDKDFLDNEENEIDILEIVLTLT DREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKKLINGIRDKQSGKTILDFLK SDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIAKKGILQTV KVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEIGIKELGDSQILKE HPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDQVHIVPQSFLLKDDSDIN KVLTRSDKARGKSDNVPSSEVVKMKMKNYWRQLLNKLTQRKFDNLTKAERGGSL LDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSFLVSDFRK DFQFYKVRINNYHHAHDAYLNAVVGTALEIKKYPKLESEFVYGDYKVDVVRKMIAS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSPMPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKPYGGFDSPTV AYSVLVAVKEKGSKKLKSVEKELLGITIMERSSFEKNPIDLEAKGYKEVKKDIJIK PKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFYLASHYKELKSGSPEDNEQ KQLFVEQHKHYLDEIEQISEFSKRVLADANLKDVL SAYNKHDKPIREQAENIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDGGAE AAAKEAAAEEAAKEAAKALEAAEAAKEAAEAAEAAEAAKAGGTAPLEEEYRL FLEAPIQNVTLLEQWKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITL EAKRSLRETIRKFRAGIIRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVTETI HPTVPNPYTLTLLSPDRIWYSVLDDKDAFFCPLAPESQLIFAEFWADAEESGQQL TWTRLPQGGFKNSPTLFNEALNRDLQGGFRDLHPSVSLQYVDDLLIAADTQAACLSAT RDLLMTLAEGLYRVSGKKAQLCQEEVYTLGFKIHKGSRSLSNSRTQAILQIPVPKTK RQVREFLGKIGYCRFLFIPGFAELAQPLAATRPNGNDPLVWGEKEEEAFQSLKLALTO PPALALPSLDKPPFLVEETSGAAGKGLTQALGPWKRPVAYLSKRLDPVAGWPR CLRAIAAAALLTREASKITFGQDIEITSSHNESSLRSPDQKWLNTARITQYQVLLDP PRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLSTRPDLTDQPLAQAEATLFTD GSSYIPDGKRYAGAAVVTLDVSIWAEPLPGTSAQAELIALTKALEWSKDKSVNIYT DSRYAFATLHVHGMIRERGWLTAGGKAIKNAPEILALLTAVWLPRKRVAVMHCKGH QKDDAPTSTGNRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFES PKKKAKVE	142

	PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDNEQ KQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDGGAE AAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAGGTAPLEEEYRL FLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITL EAKRSLRETIRKFRAAGILRPVHSPWNTPLLVRKSGTSEYRMVQDLREVNKR VETI HPTVPNPYTLLSLLPPDRIWYSVLDLKD AFFCIPLAPESQLIFAFEWADAE EGESGQL TWTRL PQGFKNSPTLFNEALNRDLQGFRLDHPSVSL LQYVDDLLIAADTQAACLSAT RDLLMTLAE LGYRVSGKKAQLCQEEV TYLGFKIHKGSRSLSNSRTQAILQIPVPKTK RQVREFLOKIGYCR LFI PGFAELAQPLYAATRP GNDPLVWGEKEEEAFQSLKLALTQ PPALALPSLDKPFQLFVEETSGAAKG VLTQALGPWKRPVAYLSKRLDPVAA GWPR CLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLDP PRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTD GSSYIRDGKRYAGAAVVTLD SVIWA EPLPIG TSAQKAELIALT KALEWSKDKSVNIYT DSRYAFATLHVHGM IYRERGWLTAGGKA IKNAP EILALLTAVWLPKRVAVMMHCKGH QKDDAPTSTGNRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFES PKKKAKVE	
--	---	--

[0446] The B2M gRNA and template RNA sequences are provided in Table 29.

[0447] Nucleotide modifications are noted as follows: phosphorothioate linkages denoted by an asterisk, 2'-O-methyl groups denoted by an 'm' preceding a nucleotide.

5 Table 29 – B2M gRNA and template sequences

Name	Sequence	SEQ ID NO
B2M gRNA1	mA*mC*mU*rCrArCrGrCrUrGrGrArUrArGrCrCrUrCrCrGrUr UrUrUrArGrAmGmCmUmAmGmAmAmAmUmAmGmCrArA rGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrC rAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAm CmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	143
B2M template RNA (with modifications)	mA*mC*mU*rCrArCrGrCrUrGrGrArUrArGrCrCrUrCrCrGrUr UrUrUrArGrAmGmCmUmAmGmAmAmAmUmAmGmCrArA rGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrC rAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAm CmCmGmAmGmUmCmGmGmUmGmCrUrUrUrCrUrGrGrCr CrUrGrGrArUrGrArUrGrGrGrCrUrArUrCrC*mA*mG*mC	144
B2M template RNA (without modifications)	ACUCACGCUGGAUAGCCUCCGUUUUAGAGCUAGAAAU AGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUG AAAAAGUGGCACCGAGUCGGUGCUUUCUGGCCUGGA UGAUGGGCUAUCCAGC	145

[0448] The B2M gRNA1 contained a spacer (ACUCACGCUGGAUAGCCUCC) (SEQ ID NO: 146) and a gRNA scaffold (GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU) (SEQ ID NO: 147).

[0449] The B2M template RNA contained:

1. a gRNA spacer;
2. a gRNA scaffold;
3. a heterologous object sequence; and
- 5 4. a primer binding site (PBS) sequence.

[0450] The heterologous object sequences and PBS sequences were designed to introduce an insertion of 5 nucleotides (TAGTG) in the first exon of B2M, which generates a frameshift mutation.

Targeted LNP (tLNP) production

10 [0451] Briefly, base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO at an N/P of 6 in the ethanol phase. The aqueous phase was composed of either mRNA encoding wtCas9-RT and the B2M gRNA or nCas9-RT and the B2M template RNA at a mass ratio of 1:1 in 25 mM acetate buffer. The two phases were mixed using a microfluidic device to
15 formulate base LNP at a total flow rate of 18 mL/min (aqueous phase: ethanol phase=3:1). 3x volume of citrate-buffered saline was added to LNP solution and resulted solution was neutralized by 1 vol% of 1M Tris-HCl, pH 8 after 15 minutes.

[0452] The anti-CD117 targeting moiety (Fab) was produced with a sortase tag (LPETG) on the C terminus of the heavy chain. 10uM Fab-LPETG and 1mM Triglycine
20 methyltetrazine (GGG-meTz) were prepared in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.4). 2uM Sortase A5 was then added to initiate the reaction. The above solution was incubated at 30°C for 3 hours with gentle mixing. Excess reagents were removed, and buffer was exchanged to PBS containing 10mM EDTA by a 10K MWCO Amicon centrifugal UF device. To confirm a DOL (degree of labeling) of close to 1 meTz per targeting
25 moiety, sortase modified Fab was allowed to react with 5eq AF594 containing TCO group in PBS for 2 hours at room temperature. After the reaction, excess dye was removed via Zeba desalting columns. DOL was determined by first measuring absorbances at A280 and A590 using a Nanodrop, and then calculating the DOL using extinction coefficients and correction factors. The resulting meTz-modified Fabs were added to the TCO modified-LNP solution
30 described above for Fab-LNP conjugation, with a 1:1 mass ratio between Fab and total RNA

payload. The solution was incubated at room temperature for 2 hours and then at 4°C overnight. Any unreacted Fabs were removed by a 300K mPES TFF membrane. The resulting tLNP products were further concentrated using a 100k MWCO Amicon centrifugal UF device. The tLNPs were characterized using a ribogreen assay to quantify RNA content and encapsulation efficiency and by DLS. *Characterization by DLS:* LNP hydrodynamic diameter and polydispersity index (PDI) were measured using single cell dynamic light scattering (DLS) (Malvern Zetasizer Pro Red Label), high throughput DLS (Wyatt DynaPro III), or high resolution DLS (Malvern NanoSight NS300). Samples were prepared by mixing LNP mixture with 1X PBS (GIBCO 10010-023) at a 1:5 ratio before LNP concentration or a 1:20 ratio after LNP concentration and then analyzed.

[0453] *Characterization by RiboGreen:* Concentration and encapsulation efficiency of the LNPs were measured using the Quant-iT RiboGreen RNA Kit. In brief, the LNP sample, mRNA standard, RiboGreen reagent, and Triton X-100 were all diluted in TE buffer to the appropriate concentration. Each sample well was comprised of 15 uL of LNP sample, 100 uL of RiboGreen reagent, 50 uL of Triton X-100, and the remaining volume being TE buffer up to 200 uL total. Samples were incubated in the dark for 30 minutes and the fluorescence was measured using an excitation wavelength of 485 nm and an emission wavelength of 525 nm. Nucleic acid concentration for unlysed and lysed LNPs were found as per manufacturer's instruction with the encapsulated mRNA used as the known standard. The concentration of the mRNA in the lysed LNP was used as the concentration of the LNP, and the encapsulation was given as the ratio of the difference in lysed and unlysed LNPs to the lysed LNPs.

[0454] *Characterization by TNS:* To determine the pKa of the LNPs, stock buffers of 25 mM PBS and 125 mM sodium chloride were prepared. The pH of the various buffers was adjusted to roughly pH 4 – 9.5 using hydrogen chloride and sodium hydroxide. Using 3 replicates of each pH value for each LNP sample, 386 uL of pH buffer, 10 uL of LNP sample (0.04 mg/mL mRNA), and 4 uL of TNS sample (0.6 mM) were added to each well and incubated in the dark for 30 minutes. Fluorescence of each sample was taken (Molecular Devices SpectraMax i3x) using an excitation wavelength of 325 nm and an emission wavelength of 435 nm.

30 *Administration of tLNPs in "humanized" NBSGW mice*

[0455] Briefly, female NBSGW mice 6-8 weeks of age (Catalog #026622) were purchased from Jackson Labs (Bar Harbor, ME). Following acclimation to the vivarium for 3-

5 days, animals received a tail vein injection of 5×10^5 human CD34⁺ HSPCs (HemaCare Division of Charles River Labs, Wilmington, MA) and the human cells were allowed to engraft for 12 weeks before the animals were used in screening studies. Animals were then treated via intravenous tail vein injection with 100 μ l of candidate tLNPs formulated with the payload at a dose of 4 mg/kg. 7 days later, animals were euthanized via CO₂ asphyxiation, weighed and their femurs, tibias and hips were collected and the bone marrow was isolated. HSCs and early progenitor cells (HSPC) were defined as Lin-/CD34⁺/CD38⁻, while long-term HSCs (LT-HSCs) were defined as Lin-/CD34⁺/CD38⁻/CD90⁺. LT-HSC were sorted as a pool from the mice belonging to the same experimental group. Editing of the B2M locus (indels) was quantified by amplicon-sequencing (Ampseq) of the locus near the binding site of the protospacer. Gene modification of the B2M locus by the nCas9-RT and B2M template RNA gene modifying system was quantified by amplicon-sequencing (Ampseq) of the locus near the spacer by quantifying the reads containing the desired insertion of 5 nucleotides (TAGTG) in the first exon of B2M.

Results

[0456] FIGS. 14A and 14B show that the administered anti-CD117 tLNPs can deliver payloads suitable for gene editing to human HSPCs in an animal model, leading to *in vivo* gene editing in HSPCs. FIG. 14A shows that administration of tLNPs formulated with Lipid 154 and the wtCas9-RT mRNA and B2M gRNA payload resulted in an average of 15.7% indels in B2M in HSPCs. FIG. 14B shows that administration of the tLNPs formulated with Lipid 154 and the nCas9-RT and B2M template RNA payload resulted in an average of 5.2% of the human HSPCs comprising the intended B2M edit, resulting in B2M functional knockout. However, clinical observations indicated that Lipid 154 tLNPs administered at 4mpk showed some potential signs of dose-limiting toxicity.

Example 6: Assaying delivery of an RNA payload to liver cells by targeted LNPs formulated with Lipid 154

[0457] Targeted LNPs that exhibited enhanced delivery of an mRNA payload to human HSCs *in vivo* were assayed to determine to what extent they delivered the RNA payload to the liver. The LNP was formulated using Lipid 154 as the ionizable lipid with a GFP mRNA payload. The LNP was then conjugated to an anti-CD117 Fab fragment, as described in Example 5, to create tLNPs. The tLNPs were then administered to humanized NBSGW mice at 2mpk using the methods described in Example 5 and evaluated to determine how much GFP

protein was expressed in liver tissue at 16 hours following LNP administration. For liver delivery measurement, liver punches were collected from administered mice during tissue takedown, homogenized and lysed; then GFP ELISA assays were performed to quantify the expression of GFP protein in liver tissue.

- 5 **[0458] FIG. 15** shows that the Lipid 154 tLNPs deliver GFP mRNA to liver tissue. Consequently, there may be a correlation between the amount of liver delivery of Lipid 154 tLNPs *in vivo* in the humanized mouse model and the toxicity of the tLNPs observed in these mice (see Example 5).

Example 7: Liver toxicology assay in a rat model

- 10 **[0459]** To further understand the liver delivery levels and potential toxicity of the LNPs in Example 5, Lipid 154 base LNPs was dosed at both 2mpk and 4mpk in a rats to evaluate its propensity to cause liver damage, as measured by ALT serum levels.

Base LNP (LNP) production

- 15 **[0460]** Briefly, base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO at an N/P of 6 in the ethanol phase. The aqueous phase was composed of GFP mRNA in 25 mM acetate buffer. The two phases were mixed using a T-mixer backed by HPLC pumps to formulate base LNP at a total flow rate of 120 mL/min (aqueous phase: ethanol phase=3:1). 3x volume of citrate-buffered saline was added to LNP solution, via an in-line dilution post- T-
20 mixer, at 360 mL/min. The resulted solution was neutralized by 1 vol% of 1M Tris-HCl, pH 8 after 15 minutes. LNPs were buffer exchanged using 300kDa mPES hollow fiber TFFs (Repligen), at a shear rate of 6000 s⁻¹ and TMP of ~5 psig. The LNPs were then further concentrated using 100kDa Amicon Filters (Millipore) at 4C and 1500 RCF. The LNPs were characterized by DLS and Ribogreen as described in Example 5.

25 *Administration of LNPs in rats*

- [0461]** LNPs were administered to sprague-dawley rats through tail vein injection at 2mpk and 4mpk. Whole blood was collected at 6 hours and 24 hours, and liver sections were collected at 24 hours after dosing. Then serum was processed for liver function tests including alanine transaminase (ALT), aspartate transaminase (AST), and alkaline phosphatase (ALP).
30 Liver samples were also processed for histological analysis.

Results

[0462] FIG. 16 shows that Lipid 154 LNPs dosed at 4mpk causes an increase in ALT levels compared to the ALT levels in rats dosed at 2mpk or in the nontreated rat. This elevation in ALT levels in rats correlates with the clinically observed symptoms of toxicity associated with Lipid 154 tLNPs in humanized mice (see Example 5) and the high GFP liver delivery in humanized mice (see Example 6). Thus, while Lipid 154 LNPs exhibits potent delivery of gene modifying systems to HSCs *in vivo* (see Example 5), there is a possibility that the amount of liver delivery could result in some toxicity, possibly reducing the tolerable dose level of this potent ionizable lipid for *in vivo* HSC delivery.

10 **Example 8: Analysis of PEG-lipid anchor length on LNP delivery of an RNA payload to human CD34+ cells in-vitro**

[0463] It was hypothesized that different PEG-lipids and mol % will influence the delivery of LNPs to HSCs *in vivo*. Different length PEG lipid anchors and mol % could affect pharmacokinetics, biodistribution, and endosomal release. To try to understand these relationships, targeted LNPs formulated with a number of different PEG-lipids and mol % conditions were screened in human CD34+ cells *in vitro* to assay their potential functionality for HSC delivery. More specifically, tLNPs comprising an anti-CD117 targeting moiety were formulated with DMG-PEG2k (C14), DPG-PEG2k (C16), DPPE-PEG2k (C16), DSG-PEG2k (C18), or DSPE-PEG2k (C18) at either 1.5, 2.5, or 3.5 mol % of the total lipid.

Targeted LNP (tLNP) production

[0464] LNPs were generally formulated as described above except for the alterations in PEG-lipids and mol %s shown in Table 30. Briefly, base LNPs were prepared at a molar ratio of 47% ionizable lipid 28.5% cholesterol: 22% DSPC: 2.0% Lipid-PEG2000: 0.5% DSPE-PEG2000-TCO at an N/P of 6 in the ethanol phase. Any alterations from a mol % of 2% Lipid-PEG, either higher or lower, as described in Table 30, were accounted for by increasing or decreasing the percent of cholesterol in the formulation proportionally to maintain a total of 100%. A subset of the LNPs were then conjugated to an anti-CD117 targeting moiety (see Table 14) and characterized by DLS and RiboGreen as described in Example 5.

Table 30: LNP Formulations

Formulation No.	Lipid-PEG2k	Lipid-PEG content (mol %)	Targeting Antibody	DLS Z-Avg. Size (nm)	mRNA Encapsulation Efficiency (%)
1	DMG	2	Anti-huCD117 Fab	77	97.3
2	DMG	1.5	Anti-huCD117 Fab	80	96.5
3	DPG	1.5	Anti-huCD117 Fab	85	97.2
4	DSG	1.5	Anti-huCD117 Fab	90	97.2
5	DPPE	1.5	Anti-huCD117 Fab	82	96.3
6	DSPE	1.5	Anti-huCD117 Fab	85	97.0
7	DMG	2.5	Anti-huCD117 Fab	71	97.3
8	DPG	2.5	Anti-huCD117 Fab	82	94.0
9	DSG	2.5	Anti-huCD117 Fab	87	96.7
10	DMG	2.5	None	72	96.8
11	DPG	2.5	None	81	97.4
12	DSG	2.5	None	86	97.3
13	DMG	3.5	Anti-huCD117 Fab	64	97.4
14	DPG	3.5	Anti-huCD117 Fab	78	97.1
15	DSG	3.5	Anti-huCD117 Fab	80	97.7

Administration of LNPs on human CD34+ cells in vitro

[0465] Human primary CD34+ cells isolated from mobilized peripheral blood were cultured in Xvivo-10 media (Lonza, Basel) supplemented with 100ng/mL SCF, TPO, Flt3L (Peprotech, Rocky Hill, NJ) for 16 hours. 1.5×10^5 cells were subsequently transferred to individual wells of a 96 well plate in 80 uL of culture media. Each of the tLNPs were diluted to 20 ug/mL in Xvivo-10 media supplemented with 10% fetal bovine serum. From this stock LNP suspension, four additional sequential 1:2 dilutions were prepared (10, 5, 2.5, 1.25 ug/mL). 20 uL of each tLNP dilution were added to each well of the 96 well plate containing the CD34+ cell culture to achieve a total volume of 100 uL/well and incubated for 24 hours. GFP expression was measured on an Agilent NovoCyte flow cytometer (Agilent, Santa Clara, CA). Analysis and quantification of the analyzed cell samples was performed on FlowJo software (Ashland OR). The flow cytometry results for the cells administered tLNPs were normalized against the flow cytometry results for the negative control cells (untreated).

Results

[0466] FIG. 17 A-B show that the lipid-PEG type as well as mol % influences the number of human CD34+ cells transfected by the tLNPs (%GFP+ cells) and amount of GFP expression (MFI) in those cells. Increasing the percentage of Lipid-PEG in tLNPs results in a decrease of %GFP+ cells, as well as in MFI. Additionally, increasing the lipid anchor length (from C14 to C16 to C18) decreases the %GFP+ cells and MFI levels, even at the same mol %, both in the presence and absence of serum. LNPs formulated with DMG-PEG were observed to be around 10nm smaller than LNPs formulated with C16 or C18 PEG, at the same mol %. In addition, LNP sizes trended smaller as mol % increased, decreasing ~10nm in size from 1.5 mol % to 3.5 mol %, generally. Administration of the non-targeted base LNPs lead to minimal levels of GFP expression. Some of the lead formulations and conditions from this *in-vitro* experiment were selected for *in vivo* screening.

Example 9: Effects of PEG-lipid anchor length on LNP delivery of an RNA payload to liver cells and payload expression levels

[0467] Targeted LNPs formulated with PEG-lipid of varying anchor length and different mol % of PEG-lipid were screened *in vivo*. Anti-CD117 targeted tLNP were prepared with either DMG-PEG2k (C14), DPG-PEG2k (C16), or DSG-PEG2k (C18) at either 1.5, 2 or 2.5 mol % of the total lipid. The tLNPs were formulated with a GFP mRNA payload and administered to NBSGW mice engrafted with human HSCs and evaluated for their

effectiveness in delivering RNA in vivo as measured by GFP expression in HSPCs (measured by FACS). A GFP ELISA assay was performed to quantify the expression of GFP protein in the liver tissue of mice.

[0468] The tLNPs were formulated as described in Example 17 and administered to humanized mice at 2mpk as described in Example 5. Bone marrow cells (1×10^6 in 0.1 ml PBS) from each test mouse were subsequently aliquoted to individual wells in 96 well plates and antibody fluorochrome conjugate mixtures (Table 31) were added to the cells in a total of 0.1 ml. Following a 30-minute incubation at room temperature, the plates were centrifuged for 5 minutes at 500G and supernatants were removed. The cells were then resuspended in 0.2 ml PBS and evaluated on an Aligent NovoCyte Quanteon (2010098) or Cytex Northern Lights (N7-00008) flow cytometer using FlowJo software (Ashland OR). Cells stained with relevant isotype control fluorochrome conjugated antibodies were used to set the negative control gates. HSC and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-. A GFP ELISA assay was also performed to quantify the expression of GFP protein in the liver tissue of mice.

Table 31. Flow cytometry antibodies

Marker	Fluorophore	Clone	Vender	Catalog
mCD45	PE-Cy5	30-F11	Biolegends	103110
CD3	PerCp-Cy5.5	UCHT1	BD	560835
CD19	R718	HIB19	BD	567343
CD11b	PE/Dazzle 594	ICRF44	Biolegends	301348
hCD45	PE	HI30	Biolegends	304008
CD16	PE-Cy7	3G8	BD	557744
Live/Dead	Yellow		ThermoFisher	L34959
CD45RA	BV650	HI100	Biolegends	304136
CD34	BV421	561	Biolegends	343610
CD38	APC-Cy7	HIT2	Biolegends	303534
CD90	APC	5 E10	BD	561971
CD117	BV605	104D2	Biolegends	313218
CD71	BV786	L01.1	BD	745555

Results

[0469] FIG. 18A shows that administration of tLNPs formulated with smaller anchor length PEG (C14) lead to higher GFP+ MFI levels in human LT-HSCs *in vivo*. Administration of tLNPs formulated with C18 anchor length PEG resulted in significantly lower GFP MFI. The administration of tLNPs formulated with C16 anchor length PEG, however, resulted in only a minimal decrease in MFI levels relative to the MFI levels generated using C14 anchor length PEG in LT-HSCs. By contrast, FIG. 18B shows that the tLNPs formulated with C14-PEG exhibited a higher propensity for liver delivery compared to the tLNPs formulated with either the C16- or C18-PEG, regardless of the mol % tested in this study. In summary, tLNPs formulated with C16-PEG are capable of delivering an mRNA payload to human LT-HSCs *in vivo*, leading to expression of a payload mRNA at levels similar to that observed using C14-PEG tLNPs, while also exhibiting a lower propensity to target liver tissues.

15 Example 10: Analysis of PEG-lipid anchor length on liver toxicology assay in rat model

[0470] Base LNPs formulated with Lipid 154 and DPG-PEG2k (C16) were evaluated for their propensity to cause liver damage, as measured by ALT serum levels in rats. LNPs comprising DPG-PEG2k (C16) were shown to exhibit lower liver delivery relative to LNPs comprising a C14-PEG (see Example 9). These rat studies were conducted to determine whether there is also a corresponding reduction in liver damage (ALT levels) when a C16-PEG is used. The base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DPG-PEG2000: 0.5% DSPE-PEG2000-TCO, and generally using the formulation process described in Example 7 (except for the PEG substitution). The LNPs were then administered to rats at 4 mpk and ALT levels were measured, as previously described in Example 7.

Results

[0471] FIG. 19 shows that Lipid 154 LNP, formulated with 2% DPG-PEG2000 (C16), does not lead to elevated serum ALT levels compared to baseline levels. This contrasts with data showing that administration of Lipid 154 LNPs formulated with DMG-PEG2000 (C14) led to elevated serum ALT levels in the humanized mouse model (Example 7). Substitution of C16 PEG for C14 PEG in the LNPs reduced the serum ALT levels, presumably a result of decreased liver delivery (as seen in Example 9).

Example 11: Effects of longer PEG-lipid anchor length on LNP delivery of a payload RNA to HSPCs in nonhuman primates

[0472] Targeted LNPs formulated with Lipid 154 and DPG-PEG2k (C16) were
5 evaluated for their efficiency in delivering an mRNA payload encoding GFP to HSPCs and LT-HSCs in nonhuman primates (NHPs).

Targeted LNP (tLNP) production

[0473] Base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154):
28.5% cholesterol: 22% DSPC: 2.0% DPG-PEG2k (C16): 0.5% DSPE-PEG2000-TCO at an
10 N/P of 6 in the ethanol phase. The aqueous phase was composed of GFP mRNA in 25 mM acetate buffer. The two phases were mixed using a T-mixer backed by HPLC pumps to formulate base LNP at a total flow rate of 120 mL/min (aqueous phase: ethanol phase=3:1). 3x volume of citrate-buffered saline was added to LNP solution, via an in-line dilution post- T-mixer, at 360 mL/min. The resulting solution was neutralized by 1 vol% of 1M Tris-HCl, pH
15 8 after 15 minutes. LNPs were buffer exchanged using 300kDa mPES hollow fiber TFFs (Repligen), at a shear rate of 6000 s⁻¹ and TMP of ~5 psig. The LNPs were then further concentrated using 100kDa Amicon Filters (Millipore) at 4C and 1500 RCF.

[0474] The anti-CD117 targeting moiety (Fab) was produced with a sortase tag (LPETG) on the C terminus of the heavy chain. 10uM Fab-LPETG and 1mM Trglycine
20 methyltetrazine (GGG-meTz) were prepared in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.4). 2uM Sortase A5 was then added to initiate the reaction. The above solution was incubated at 30°C for 3 hours with gentle mixing. Excess reagents were removed, and buffer was exchanged to PBS containing 10mM EDTA by a 10K MWCO Amicon centrifugal UF device. To confirm a DOL (degree of labeling) of close to 1 meTz per targeting
25 moiety, sortase modified Fab was allowed to react with 5eq AF594 containing TCO group in PBS for 2 hours at room temperature. After the reaction, excess dye was removed via Zeba desalting columns. DOL was determined by first measuring absorbances at A280 and A590 using a Nanodrop, and then calculating the DOL using extinction coefficients and correction factors. The resulting meTz-modified Fabs were added to the TCO modified-LNP solution
30 described above for Fab-LNP conjugation, with a 1:1 mass ratio between Fab and total RNA payload. The solution was incubated at room temperature for 2 hours and then at 4°C overnight. Any unreacted Fabs were removed by a 300K mPES TFF membrane. The resulting tLNP products were further concentrated using a 100k MWCO Amicon centrifugal UF device. The

tLNPs were characterized using a ribogreen assay to quantify RNA content and encapsulation efficiency and by DLS, as described in Example 5.

Administration of LNPs to cynomolgus macaques

[0475] Sexually mature male or female cynomolgus macaques (Biomere, Worcester, MA) were enrolled in the study and baseline weights and blood samples were collected. Animals were pre-treated with famotidine (0.5 mg/kg, IM) and diphenhydramine (5 mg/kg, IM) on Day -1, and on Day 1 at 30-60 minutes prior to start of treatment. For each animal, tLNPs were administered intravenously through a saphenous vein, cephalic vein, or tail vein with a temporary IV catheter on Day 1 at a dose of 2 mg/kg via an approximate 60-minute infusion using an infusion pump, followed by about 0.5 mL of saline to flush the dose from the IV catheter. Additional blood samples were collected at 1 hour and 15 to 20 hours post infusion. Lymph node and bone marrow aspirate were also collected at 15-20 hours after the end of infusion. Separately, bone marrow aspirates from non-treated cynomolgus macaques were collected to establish a negative control for baseline %GFP+ and MFI levels. HSCs and early progenitor cells (HSPC) were defined as Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as Lin-/CD34+/CD38-/CD90+.

Flow cytometric analysis of NHP bone marrow

[0476] Bone marrow cells (3×10^6 in 0.1 ml PBS) from each nonhuman primate in the studies were subsequently aliquoted to individual wells in 96 well plates and centrifuged at 500 G for 5 min at RT. A master mix composed of antibody fluorochrome conjugate mixtures (Table 16) was added to the cells in a total of 0.1 ml. Following a 30-minute incubation at 4°C, the plates were centrifuged for 5 min at 500G and supernatants were removed. Cells were resuspended in 200 uL of fixation buffer (BD Cytofix/Cytoperm) and incubated for 20 min at 4°C. After fixation, cells were centrifuged for 5 min at 500G and washed once in 200 uL of PBS. After centrifugation, the supernatants were discarded and the cells were resuspended in 0.2 ml PBS and evaluated on an Agilent NovoCyte Quanteon (2010098) or Cytex Northern Lights (N7-00008) flow cytometer. Analysis and quantification of the analyzed BM samples was performed on FlowJo software (Ashland OR). FMO (fluorescence-minus-one) controls for all channels were used to set the negative control gates. HSCs and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-.

Table 32. Flow cytometry antibodies

Marker	Clone	Manufacturer	Fluorophore	Catalog No.
Live/Dead	-	Thermo Fisher	Yellow	L34959
CD45	MBD-6D6	Miltenyi Biotec	PE	130-123-703
CD34	563	BD	BV421	740081
CD38	AT-1	Novus	APC Cy7	740608
CD45RA	5H9	BD	BV650	740608
CD90	5E10	BD	APC	559869
CD117	104D2	Biolegend	BV605	313218
CD123	9F5	BD	BV786	740989
LDLR	C7	BD	BV711	745480
CD3	SP34-2	BD	BB700	566517
CD20	2H7	BD	R718	566988
CD11b	ICRF44	Biolegend	PE-Dazzle 594	301348
CD16	3G8	BD	PE-Cy7	557744

Results

[0477] FIG. 20A shows that administration of anti-CD117 tLNP formulated with Lipid 154 and 2 mol % C16-PEG in the macaques (n=2) resulted in high percentages of GFP+ HSPCs and LT-HSCs (96% GFP+ and 91% GFP+ cells, respectively). FIG. 20B shows that the expression of GFP mRNA dosed with the Lipid 154 tLNPs. The tLNPs were well tolerated by the animals, as none of the animals dosed with the Lipid 154 tLNPs comprising C16-PEG exhibited any clinical signs of dose-limiting toxicity.

10 **Example 12: Dual targeted LNPs formulated with Lipid 154 and different PEG-lipid anchor lengths exhibit enhanced delivery of an RNA payload comprising a gene editing system to HSPCs *in vivo***

[0478] LNPs formulated with Lipid 154, and either DMG-PEG2k (C14) or DPG-PEG2k (C16), were next conjugated to two targeting moieties that each bind to a different protein on the surfaces of HSPCs to create dual targeted LNPs (dual tLNPs) to determine whether multiple targeting moieties could improve delivery of a mRNA payload to HSCs in vivo. The targeting moieties conjugated to LNPs were an anti-CD34 Fab and an anti-CD117 Fab. The tLNPs were formulated with a gene editing system payload comprising an mRNA encoding a wild-type Cas9-RT fusion protein and a gRNA specific for the first exon of the human Beta 2 Microglobulin (B2M) gene at a mass ratio of 1:1, as described in Example 9. The tLNPs were administered to humanized mice and evaluated for their effectiveness in

delivering the payload *in vivo* as measured by the percentage of B2M indels in HSPCs (assayed by Amp-Seq).

Targeted LNP (tLNP) production

[0479] Briefly, base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DMG--PEG2k (C14) or DPG-PEG2k (C16): 0.5% DSPE-PEG2000-TCO at an N/P of 6 in the ethanol phase, using the method described in Example 5.

[0480] Each targeting moiety (Fab) was produced with an additional sortase tag (LPETG) on the C terminus of the heavy chain. 10uM Fab-LPETG and 1mM Triglycine methyltetrazine (GGG-meTz) were prepared in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.4). 2uM Sortase A5 was then added to initiate the reaction. The above solution was incubated at 30°C for 3 hours with gentle mixing. Excess reagents were removed, and buffer was exchanged to PBS containing 10mM EDTA by a 10K MWCO Amicon centrifugal UF device. To confirm a DOL (degree of labeling) of close to 1 meTz per targeting moiety, sortase modified Fab was allowed to react with 5eq AF594 containing TCO group in PBS for 2 hours at room temperature. After the reaction, excess dye was removed via Zeba desalting columns. DOL was determined by first measuring absorbances at A280 and A590 using a Nanodrop, and then calculating the DOL using extinction coefficients and correction factors. The resulting meTz-modified Fabs were added to the TCO modified-LNP solution described above for Fab-LNP conjugation, with a 1:1:1 mass ratio between Fab1 (anti-CD117 Fab), Fab2 (anti-CD34 Fab), and total RNA payload for dual tLNPs. The solution was incubated at room temperature for 2 hours and then at 4°C overnight. Any unreacted Fabs were removed by a 300K mPES TFF membrane. The resulting dual tLNP products were further concentrated using a 100k MWCO Amicon centrifugal UF device. A ribogreen assay was then performed to quantify RNA content and encapsulation efficiency. The LNPs were characterized by DLS and Ribogreen as described in Example 5.

Administration of tLNPs in “humanized” NBSGW mice

[0481] LNPs were administered to “humanized” NBSGW mice, at 4mpk, and HSPCs from bone marrow were analyzed as described in Example 1. HSCs and early progenitor cells (HSPC) were defined as Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as Lin-/CD34+/CD38-/CD90+. LT-HSC were sorted as a pool from the mice belonging to the

same experimental group. Editing of the B2M locus (indels) was quantified by amplicon-sequencing (Ampseq) of the locus near the binding site of the protospacer.

Results

[0482] FIG. 21 shows that administration of the anti-CD117/anti-CD34 dual tLNPs formulated with Lipid 154 dual tLNPs formulated with C16-PEG generated a significant number of human HSPCs *in vivo* comprising B2M indels (greater than 45% indels). Humanized mice dosed with 4mpk of the Lipid 154 dual tLNPs showed no clinical signs of dose-limiting toxicity in contrast to the humanized mice dosed with 4mpk of Lipid 154 tLNPs comprising C14-PEG (see Example 5). Thus, Lipid 154 tLNPs can deliver a gene modifying system to human HSPCs *in vivo* to more effectively generate engineered cells. Optimization of Lipid 154 tLNPs by substitution of C16-PEG in place of C14-PEG eliminated clinical signs of dose-limiting toxicity associated with these LNPs. The improvement in tolerability offsets any potential decrease in HSC potency associated with a C16-PEG vs. a C14-PEG, suggesting that C16 PEG may be used to improve the tolerability of potent ionizable lipids by de-targeting the liver, enabling increases in dose levels that allow for enhanced *in vivo* delivery.

Example 13: Dual targeted LNPs formulated with Lipid 154 exhibit enhanced delivery of an RNA payload comprising an exemplary gene modifying system to HSPCs *in vivo*

[0483] The antiCD117/anti-CD34 dual tLNPs (comprising Lipid 154 and DPG-PEG2k (C16)) were formulated as described in Example 12, except they encapsulated a gene modifying system comprising the mRNA encoding a nCas9-RT and the template RNA a template RNA suitable to produce an insertion of five nucleotides (TAGTG) into the first exon of B2M, as described in Example 5. The nCas9-RT and template RNA molecule comprised the sequences indicated in Example 5. The five-nucleotide insertion generates a frameshift mutation which results in loss of expression of B2M from the cell surface. The tLNPs were administered to humanized mice engrafted with human HSPCs and evaluated for their effectiveness in delivering the payloads *in vivo* as measured by the percentage of reads including the desired insertion of 5 nucleotides (TAGTG) in the first exon of B2M (assayed by Ampseq).

Targeted LNP (tLNP) production

[0484] Dual tLNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DPG-PEG2000 (C16): 0.5% DSPE-PEG2000-TCO at an

N/P of 6, as described in Example 12, except for the difference in gene modifying system payload.

Administration of tLNPs in “humanized” NBSGW mice

[0485] LNPs were administered to “humanized” NBSGW mice, at 4mpk, and HSPCs from bone marrow were analyzed as described in Example 1. HSCs and early progenitor cells (HSPC) were defined as Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as Lin-/CD34+/CD38-/CD90+. LT-HSC were sorted as a pool from the mice belonging to the same experimental group. Gene modification of the B2M locus was quantified by amplicon-sequencing (Amp-Seq) of the locus near the spacer by quantifying the reads containing the desired insertion of 5 nucleotides (TAGTG) in the first exon of B2M.

Results

[0486] FIG. 22 shows that anti-CD117/anti-CD34 dual-targeted tLNPs, formulated with Lipid 154 and C16 PEG led to a mean 11.8% perfect gene modification in HSPCs and mean 14.3 perfect gene modification in LT-HSCs *in vivo*. This is approximately a three-fold improvement over the B2M rewriting data presented in Example 5 (FIG. 14) when Lipid 154 tLNPs were conjugated to the anti-CD117 targeting moiety and formulated with a C14-PEG. The Lipid 154 dual tLNPs showed no clinical signs of dose-limiting toxicity.

Example 14: Dual targeted LNPs formulated with Lipid 154 exhibit enhanced delivery of an mRNA payload to HSPCs in nonhuman primates

[0487] LNPs were formulated with Lipid 154 and conjugated to two targeting moieties that each bind to a different protein on the surfaces of HSPCs to create dual targeted LNPs (dual tLNPs). Dual tLNPs were conjugated to anti-CD117 and anti-CD34 Fab fragments or to anti-CD117 and anti-CD71 Fab fragments. They were then evaluated for their efficiency in delivering an mRNA payload encoding GFP to HSPCs and LT-HSCs in nonhuman primates (NHPs).

Targeted LNP (tLNP) production

[0488] Base LNPs were prepared at a molar ratio of 47% ionizable lipid (Lipid 154): 28.5% cholesterol: 22% DSPC: 2.0% DPG-PEG2k (C16): 0.5% DSPE-PEG2000-TCO at an N/P of 6 in the ethanol phase. The aqueous phase was composed of GFP mRNA in 25 mM acetate buffer. The two phases were mixed using a T-mixer backed by HPLC pumps to

formulate base LNP at a total flow rate of 120 mL/min (aqueous phase: ethanol phase=3:1). 3x volume of citrate-buffered saline was added to LNP solution, via an in-line dilution post-T-mixer, at 360 mL/min. The resulted solution was neutralized by 1 vol% of 1M Tris-HCl, pH 8 after 15 minutes. LNPs were buffer exchanged using 300kDa mPES hollow fiber TFFs (Repligen), at a shear rate of 6000 s⁻¹ and TMP of ~5 psig. The LNPs were then further concentrated using 100kDa Amicon Filters (Millipore) at 4C and 1500 RCF.

[0489] Each targeting moiety (Fab) was produced with an additional sortase tag (LPETG) on the C terminus of the heavy chain. 10uM Fab-LPETG and 1mM Triglycine methyltetrazine (GGG-meTz) were prepared in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM CaCl₂, pH 7.4). 2uM Sortase A5 was then added to initiate the reaction. The above solution was incubated at 30°C for 3 hours with gentle mixing. Excess reagents were removed, and buffer was exchanged to PBS containing 10mM EDTA by a 10K MWCO Amicon centrifugal UF device. To confirm a DOL (degree of labeling) of close to 1 meTz per targeting moiety, sortase modified Fab was allowed to react with 5eq AF594 containing TCO group in PBS for 2 hours at room temperature. After the reaction, excess dye was removed via Zeba desalting columns. DOL was determined by first measuring absorbances at A280 and A590 using a Nanodrop, and then calculating the DOL using extinction coefficients and correction factors. The resulting meTz-modified Fabs were added to the TCO modified-LNP solution described above for Fab-LNP conjugation, with a 1:1:1 mass ratio between Fab1, Fab2, and total RNA payload for dual tLNPs. The solution was incubated at room temperature for 2 hours and then at 4°C overnight. Any unreacted Fabs were removed by a 300K mPES TFF membrane. The resulting dual tLNP products were further concentrated using a 100k MWCO Amicon centrifugal UF device. A ribogreen assay was then performed to quantify RNA content and encapsulation efficiency.

25 *Administration of LNPs to cynomolgus macaques*

[0490] Sexually mature male or female cynomolgus macaques (Biomere, Worcester, MA) were enrolled in the study and baseline weights and blood samples were collected. Animals were pre-treated with famotidine (0.5 mg/kg, IM) and diphenhydramine (5 mg/kg, IM) on Day -1, and on Day 1 at 30-60 minutes prior to start of treatment. For each animal, tLNPs were administered intravenously through a saphenous vein, cephalic vein, or tail vein with a temporary IV catheter on Day 1 at a dose of 1 mg/kg via an approximate 60-minute infusion using an infusion pump, followed by about 0.5 mL of saline to flush the dose from the

IV catheter. Additional blood samples were collected at 1 hour and 15 to 20 hours post infusion. Lymph node and bone marrow aspirate were also collected at 15-20 hours after the end of infusion. Separately, bone marrow aspirates from non-treated cynomolgus macaques were collected to establish a negative control for baseline %GFP+ and MFI levels. HSCs and early progenitor cells (HSPC) were defined as Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as Lin-/CD34+/CD38-/CD90+.

Flow cytometric analysis of NHP bone marrow

[0491] Bone marrow cells (3×10^6 in 0.1 ml PBS) from each nonhuman primate in the studies were subsequently aliquoted to individual wells in 96 well plates and centrifuged at 500 G for 5 min at RT. A master mix composed of antibody fluorochrome conjugate mixtures (Table 33) was added to the cells in a total of 0.1 ml. Following a 30-minute incubation at 4°C, the plates were centrifuged for 5 min at 500G and supernatants were removed. Cells were resuspended in 200 uL of fixation buffer (BD Cytofix/Cytoperm) and incubated for 20 min at 4°C. After fixation, cells were centrifuged for 5 min at 500G and washed once in 200 uL of PBS. After centrifugation, the supernatants were discarded and the cells were resuspended in 0.2 ml PBS and evaluated on an Agilent NovoCyte Quanteon (2010098) or Cytex Northern Lights (N7-00008) flow cytometer. Analysis and quantification of the analyzed BM samples was performed on FlowJo software (Ashland OR). FMO (fluorescence-minus-one) controls for all channels were used to set the negative control gates. HSCs and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-.

Table 33. Flow cytometry antibodies

Marker	Clone	Manufacturer	Fluorophore	Catalog No.
Live/Dead	-	Thermo Fisher	Yellow	L34959
CD45	MBD-6D6	Miltenyi Biotec	PE	130-123-703
CD34	563	BD	BV421	740081
CD38	AT-1	Novus	APC Cy7	740608
CD45RA	5H9	BD	BV650	740608
CD90	5E10	BD	APC	559869
CD117	104D2	Biolegend	BV605	313218
CD123	9F5	BD	BV786	740989
LDLR	C7	BD	BV711	745480
CD3	SP34-2	BD	BB700	566517
CD20	2H7	BD	R718	566988

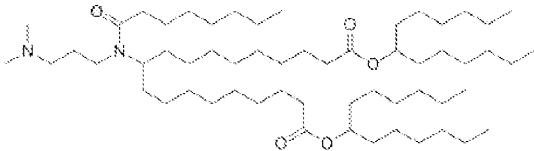
CD11b	ICRF44	Biolegend	PE-Dazzle 594	301348
CD16	3G8	BD	PE-Cy7	557744

Results

[0492] FIG. 23A shows that administration of dual tLNPs formulated with Lipid 154 to cynomolgus macaques (n=2) resulted in an average of over 90% GFP+ HSPCs and LT-HSCs for the anti-CD117/anti-CD34 dual tLNPs, and an average of over 80% GFP+ HSPCs and over 70% GFP+ LT-HSCs for the anti-CD117/anti-CD71 dual tLNPs. FIG. 23B shows that HSPCs and LT-HSCs in the NHPs expressed GFP following administration of each of the dual tLNPs.

Example 15: Evaluating Lipid 232 for LNP delivery of an exemplary all-RNA gene modifying system to correct the human SERPINA1 PiZ mutation in the liver cells of transgenic mice

[0493] LNPs formulated with the Lipid 232 ionizable lipid were tested to determine whether they could effectively deliver an all-RNA gene modifying system to hepatocytes *in vivo*. Lipid 232 was produced using the methods described in Example 1. LNPs were then formulated with the ionizable lipid and an exemplary gene modifying system that was designed to introduce a precise edit to correct the PiZ mutation (corresponding to a A>G base change) in a human *SERPINA1* (*hSERPINA1*) gene *in vivo* in mice modified to carry *hSERPINA1**E342K (*PiZ mutation*) encoding alpha-1-antitrypsin (A1AT) protein. The LNPs were administered to the transgenic mice and then delivery efficiency was measured *in vivo* by assaying the percentage of reads that included the desired modification of the *hSERPINA1* gene (assayed by Amp-Seq). Benchmark LNPs were also produced using either the V003 or V005 baseline ionizable lipid and the same gene modifying system, and then administered to the transgenic mice. The ionizable lipid V005 has the following structure:



V005

[0494] As described above, the LNPs were formulated to encapsulate an exemplary gene modifying system comprising an mRNA encoding the gene modifying polypeptide (an

nCas9-RT polypeptide, RNAIVT6898, SEQ ID NO: 148) and a template RNA (“SERPINA1 template RNA,” RNACS20686, SEQ ID NO: 150) at a mass ratio of 1:1.

[0495] The gene modifying polypeptide comprised the amino acid sequence provided in Table 34.

5 Table 34: Gene modifying polypeptide sequence

Name	AA sequence	SEQ ID NO
RNAIV T6898	MPAAKRVKLDGGSDDLVLGLDIGISVGVGILNKVTGEIIHKNSRIFP AAQAENNLVRRRTNRQGRRLARRKKHRRVRLNRLFEEESGLITDFTKI SINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLLDDASDDG NSSVGDYAQIVKENSQKLETKTPGQIQRLERYQTYGQLRGDFTVEK DGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTG KRKYHGGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYYPDEFR AAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEK AMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKT LETLDIEQMDRETLDKLAYVLTNTEREGIQEALHEFADGGSFSQK QVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTIL TRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAI KEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAA NQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGTISIHD LINNSNQFEVDHILPLSITFDDSLANKVLVYATAAQEKGQRTPYQA LDSMDDAWSFRELKAFVRESKTL SNKKKEYLLTEEDISKFDVRKKF IERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRH WGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLL DIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSK FNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFM KIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNP FLKYKEEHGYIRKYSKKGNGPEIKSLKYYSKLGHNHIDITPKDSNN KVVQLQSVSPWRADVFNKTTGKYEILGLKYADLQFEKGTGTYKIS QEKYNDIKKKEGVDSSEFKFTLYKNDLLL VKDTETKEQQLFRFLS RTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLG KSNISIIYKVRTDVLGNQHIIKNEGDKPKLDFGGWQAAESYEVGGTA PLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNT PLLVRKSGTSEYRMVQDLREVNKRVTIHPTVPNPYTLLSLLPPDR IWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRL PQ GFKNSPTL FNEALNRDLQGFRLDHP SVSLLQYVDDLLIAADTQAAC LSATRDLLMTL AELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSN SRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRP GNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSG AAKGVLTQALGPWKRPVAYLSKRLDPV AAGWPRCLRAIAAAALL TREASKLTFGQDIEITSSHNL ELLRSPDPKWLTNARITQYQVLLLD PPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSL TSTRPDLTDQ	148

	PLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSEVIWAEPLPIGTSA QKAEALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIRERGRW LTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTG NRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFESP KKKAKVE***	
--	--	--

[0496] Table 35 provides the evaluated exemplary template RNA sequence. Nucleotide modifications are noted as follows: phosphorothioate linkages denoted by an asterisk, 2'-O-methyl groups denoted by an 'm' preceding a nucleotide, 2-Fluororibose denoted by an 'i2F' preceding a nucleotide.

Table 35: Exemplary template RNA and sequence

RNACS	Name	Sequence	SEQ ID NO
RNACS20686	St1_t-lock_ design8_OMe004&F038_R10P8_edit_7	mA*mA*mG*rGrCrUrGrUrGrCrUrGrArCrCrArUrCrGrArGrUrCrUrUrGrUrArCrUmCmUmGmGmGmAmCmUmUmCmGmGmUmCmCmCmAmGrArArGrCrUrArCrArArArGrAmUmAmA/i2FG//i2FG//i2FC//i2FU//i2FU//i2FC/rArU/i2FG//i2FC//i2FC/rGrArArA/i2FU//i2FC//i2FA/rUrUrUrCrUrCrGrUrCrGrArUrGrGrU*mC*mA*mG	149

[0497] Table 36 shows the sequence of Table 19 without chemical modifications. In some embodiments, the sequences of Table 36 may be used without chemical modifications, or with one or more chemical modifications.

Table 36: Table 35 sequences without chemical modifications

RNACS	Name	Sequence	SEQ ID NO
RNACS20686	St1_t-lock_ design8_OMe004&F038_R10P8_edit_7	AAGGCUGUGCUGACCAUCGAGUCUUUGUACUCUGGGA CUUCGGUCCCGAGAAGCUACAAAGAUAAAGGCUUCAUGCC GAAAUCAUUUCUCGUCGAUGGUCAG	150

LNP production

[0498] Briefly, LNPs were prepared at a molar ratio of 50% ionizable lipid: 38% cholesterol: 9% DSPC: 3% DMG-PEG2000 at an N/P of 6 in the ethanol phase. The aqueous phase was composed of mRNA encoding the gene modifying polypeptide and the SERPINA1 template RNA at a mass ratio of 1:1 (w/w) in 25 mM citrate buffer with 100 mM NaCl, pH 5. The two phases were mixed using a microfluidic device to formulate LNPs at a total flow rate

of 7 mL/min (aqueous phase: ethanol phase=2:1). 0.33x volume of water was added to the LNP solution and resulting solution was kept for an hour and then dialyzed against 50 mM Tris buffer, pH=7.2 (50 mM sodium chloride and 9% sucrose). After overnight dialysis, LNPs were concentrated using centrifugation and sterile-filtered using 0.22 µm filters.

5 Administration of LNPs to transgenic mice

[0499] LNPs formulated with the gene modifying system comprising mRNA encoding RNAiVT6898 gene modifying polypeptide and template RNAs described above were delivered to mice. Specifically, 0.1 mg/kg of total RNA equivalent formulated in LNPs were dosed intravenously in about 8-week-old, gender balanced *hSERPINA1 E342K* mice (0.05 mg/kg of template RNA and 0.05 mg/kg of mRNA) in a 10 ml/kg bolus. Mice were administered a dose at time 0 (t = 0). At 8 days post-dosing (as used herein post-dosing refers to time since the first dose), animals were sacrificed, and their liver and serum were collected for analyses.

[0500] Liver samples were analyzed by using Amp-Seq to determine % rewriting and % INDELs in target liver cells. To analyze gene editing activity, primers flanking the target mutation site locus were used to amplify across the locus in the genomic DNA of liver samples. Amplicons were analyzed via short read sequencing using an Illumina MiSeq. Conversion of an A nucleotide to a G nucleotide at position c.1096 in exon 5 in the human *SERPINA1* gene indicated successful editing.

20 Results

[0501] FIG. 36A shows that administration of the Lipid 232 LNPs to the transgenic mice results in an average of about 15.9% of the SERPINA1 alleles in the hepatocytes carrying the precise corrective edit at 8 days post-dosing, respectively, as assayed by Amp-Seq. By contrast, administration of the LNPs formulated with the baseline V003 ionizable lipid resulted in significantly lower rewriting levels (an average of 6.6%). Administration of the LNPs comprising the V005 ionizable lipid resulted in an average of 23% rewriting. FIG. 36B shows that administration of each of the LNPs comprising the gene modifying system resulted in very low indel percentages. The results show that the exemplary Lipid 232 LNPs can deliver a gene modifying system to hepatocytes *in vivo* to effectively introduce precise edits into the human SERPINA1 gene locus at low dose amounts (0.1 mg/kg RNA), while introducing very small numbers of indels.

Example 16: LNPs formulated with novel ionizable lipids can effectively deliver RNA payloads comprising a gene modifying system designed to introduce a precise mutation into the SERPINA1 gene to the liver cells of nonhuman primates

[0502] LNPs formulated with the Lipid 232 ionizable lipid were tested for their ability to deliver an all-RNA gene modifying system payload to the hepatocytes of nonhuman primates (NHPs) to introduce a precise edit into the SERPINA1 locus *in vivo*. The gene modifying system comprising an mRNA encoding a Cas9 nickase fused to an RT and a template RNA was designed to install a silent mutation into the SERPINA1 locus of cynomolgus primates. Delivery efficiency was measured *in vivo* by assaying the percentage of reads that included the desired modification of the SERPINA1 gene (assayed by Amp-seq).

[0503] As described above, the LNPs containing nCas9-RT and template RNA were formulated to encapsulate an mRNA encoding Cas9 nickase fused to an RT (“nCas9-RT”, SEQ ID NO: 151) and the template RNA (“SERPINA1 template RNA” SEQ ID NO: 152) at a mass ratio of 1:1.

[0504] The exemplary gene modifying polypeptide (RNAIVT9315) comprising the amino acid sequence of SEQ ID NO: 151 is provided in Table 37.

Table 37: Gene modifying polypeptide sequence

Name	AA sequence	SEQ ID NO
RNAIVT9315	MPAAKRVKLDGGSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRRTNRQGRRLARRKKHRRVRLNRLFEESGLITDFTKI SINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSQLETKTPGQIQLERYQTYGQLRGDFTVEKD GKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKR KYYHGPNGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAK ASYTAQEFNLLNDLNNLTVPETKKLSKEQKNQIINYVKNEKAMGPAKL FKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMD RETLDKLAYVLTNTEREGIQEAEHEFADGSFSQKQVDELVQFRKANSSIF GKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYI DEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNED DEKKAIQKIQKANKDEKDAAMLKAA NQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGTISIHD LINNSNQFEVDHILPLSITFDDSLANKVLVYATAAQEKGQRTPYQAL DSMDDAWSFRELKAFVRESKTL SNKKKEYLLTEEDISKFDVRKKFI ERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRH	151

	WGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLL DIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSK FNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFM KIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNP FLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGHNHIDITPKDSNN KVVLSQSVSPWRADVFNKTTGKYEILGLKYADLQFEKGTGTYKIS QEKYNDIKKKEGVDSSEFKFTLYKNDLLL VKDTETKEQQLFRFLS RTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSQGCKKGLG KSNISYKVRTDVLGNQHIIKNEGDKPKLDFGGWQAAESYEVGGTA PLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNT PLLPVRKSGTSEYRMVQDLREVNKR VETIHPTVPNPYTLLSLLPPDR IWYSVLDLKD AFFCIPLAPESQLIF AFEWADAEEGESGQLTWTRL PQ GFKNSPTL FNEALNRDLQGFRLDHPVSLLQYVDDLLIAADTQAAC LSATRDLLMTLAE LGYRVSGKKAQLCQEEVTYLGFKIHKGSRLSN SRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRP GNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSG AAKGVL TQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALL TREASKLTFGQDIEITSSHNLESLLRSPDKWL TNARITQYQVLLD PPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSL TSTRPDLTDQ PLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIG TSA QKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIRERGW LTAGGKAIKNAPEILALLTAVWLPKRVA VMHCKGHQKDDAPTSTG NRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFESP KKKAKVE	
--	--	--

[0505] The template RNA sequences is provided in Table 38. Nucleotide modifications are noted as follows: phosphorothioate linkages denoted by an asterisk, 2'-O-methyl groups denoted by an 'm' preceding a nucleotide, 2-Fluororibose denoted by an 'i2F' preceding a nucleotide.

Table 38: Exemplary template RNA

ID	Sequence (with modifications)	SED ID NO	Sequence (no modifications)	SEQ ID NO
RNACS22390	mU*mG*mA*rGrArCrArUrGrCrUrUr CrCrArGrUrArCrArCrGrUrCrGrUrUrG rUrArCrUmCmUmGmGmGmAmCmU mUmCmGmGmUmCmCmCmAmGrAr	152	UGAGACAUGCUUC CAGUACACGUCGU UGUACUCUGGGAC UUCGGUCCCAGAA	153
cA1AT18_S2 l_TLock_RA R_U4G_OMe 004&F038_R1 4P11	ArGrCrUrArCrArArCrGrAmUmAmA/i 2FG//i2FG//i2FC//i2FU//i2FU//i2FC/rA rU/i2FG//i2FC//i2FC/rGrArArArUrCrA rUrGrArGrArUrUrCrCrArGrGrUrGrUr ArCrUrGrGrArA*mG*mC*mA		GCUACAACGAUAA GGCUUCAUGCCGA AAUCAUGAGAUUC CAGGUGUACUGGA AGCA	

LNP production

[0506] Briefly, LNPs were prepared using a molar ratio of 47.5% ionizable lipid: 40% cholesterol: 10% DSPC: 2.5% DMG-PEG2000 at a concentration in the ethanol phase corresponding to an N/P ratio of 6. The aqueous phase was composed of mRNA encoding the exemplary gene modifying polypeptide and the SERPINA1 template RNA at a mass ratio of either 1:1 or 2:1 in 25 mM acetate buffer. The two phases were combined using a mixing tee at a ratio of 3:1 (aqueous phase: ethanol phase). Mixed LNPs were diluted with 1-volume of citrate-buffered saline and the resulting solution was neutralized by 1 vol% of 1M Tris-HCl, pH 8 after 15 minutes. LNPs were buffer exchanged into a suitable storage buffer and to remove ethanol using tangential flow filtration. The final LNPs were sterile filtered and were stored at -80C until use.

Administration of LNPs to cynomolgus macaques

[0507] Sexually mature male cynomolgus macaques (Biomere, Worcester, MA) were enrolled in the study and baseline weights and blood samples were collected. Between 1.5 to 2 hours prior to each Test Article dose administration, and at approximately 24 and 48-hours post-dose, animals will be administered dexamethasone at a dose level of 2 mg/kg (1 mg/kg will be administered by intramuscular (IM) injection and 1 mg/kg will be administered by oral dose). For each animal, LNPs were administered intravenously through a saphenous vein, cephalic vein, or tail vein with a temporary IV catheter on Day 1 at a dose of 2 mg/kg (dose volume of 5 mL/kg) via an approximate 60-minute infusion using an infusion pump, followed by about 0.5 mL of saline to flush the dose from the IV catheter. At 7 days after dosing, liver biopsies were collected and the samples were analyzed by using Amp-Seq to determine % rewriting at the target locus in the genomic DNA and the expressed mRNA. To analyze gene editing activity, primers flanking the target locus were used to amplify across the locus in the genomic DNA or the cDNA of liver samples. Amplicons were analyzed via short read sequencing using an Illumina MiSeq.

Results

[0508] FIG. 37A shows that an average of 22.7% of the SERPINA1 alleles in the hepatocytes of the dosed NHPs carried the precise edit following administration of the LNPs formulated with Lipid 232 comprising the exemplary gene modifying system, as assayed by Amp-Seq. An average of approximately 33.3 percent of the liver cDNA carried the intended precise edit. The rewriting results are summarized in Table 39. The results show the exemplary

Lipid 232 LNPs can effectively deliver a gene modifying system to NHP hepatocytes *in vivo* to introduce precise edits into the SERPINA1 gene locus.

Table 39: SERPINA 1 Rewriting Results in NHPs

Dose (mg/kg)	Animal	gDNA Ampseq (%)	Mean ± SD	cDNA Ampseq (%)	Mean ± SD	Indel (%)	Mean ± SD
2	1	42.2	22.7 ± 17.4	62.7	33.3 ± 26.9	1.5	1.3 ± 0.2
	2	17.1		27.4		1.2	
	3	8.9		9.9		1.2	
3	1	47.9	42.7 ± 7.4	85.5	76.3 ± 13.0	0.9	0.8 ± 0.2
	2*						
	3	37.5		67.1		0.6	

5 [0509] Next, the LNPs were formulated with Lipid 232 as described above and administered to cynomolgus macaques (n=3) at a higher dose (3 mg/kg) to determine whether the higher dose would result in an increase in SERPINA1 editing by the exemplary gene modifying system. FIG. 37B shows that increasing the LNP dose to 3 mg/kg leads to a corresponding increase in perfect rewriting of the SERPINA1 locus in NHP hepatocytes, resulting in an average of 42.7% of SERPINA1 alleles having the intended precise edit. An average of approximately 76.3 percent of the liver cDNA carried the intended precise edit at SERPINA1.

15 **Example 17: GalNAc-conjugated LNPs formulated with a novel ionizable lipid can effectively deliver RNA payloads comprising a gene modifying system designed to introduce a precise mutation into the SERPINA1 gene to the liver cells of nonhuman primates**

20 [0510] LNPs formulated with the Lipid 232 ionizable lipid were tested for their ability to deliver an all-RNA gene modifying system payload to the hepatocytes of nonhuman primates (NHPs) to introduce a precise edit into the NHP SERPINA1 locus *in vivo*. The exemplary gene modifying system comprising an mRNA encoding a gene modifying polypeptide (Cas9 nickase fused to an RT) and a template RNA was designed to install a silent mutation into the SERPINA1 locus of cynomolgus primates. The LNPs comprising the gene modifying system were conjugated to GalNAc molecules and administered to the NHPs. Delivery efficiency was measured *in vivo* by assaying the percentage of reads that included the desired modification of the SERPINA1 gene (assayed by Amp-Seq).

[0511] As described above, the LNPs were formulated to encapsulate an mRNA encoding gene modifying polypeptide (SEQ ID NO: 151) and the template RNA (“SERPINA1 template RNA” SEQ ID NO: 152) at a mass ratio of 1:1.

[0512] The exemplary gene modifying polypeptide (RNAIVT9315) comprising the amino acid sequence of SEQ ID NO: 151 is provided in Table 40.

Table 40: Gene modifying polypeptide sequence

Name	AA sequence	SEQ ID NO
RNAIV T9315	MPAAKRVKLDGGSDDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFP AAQAENNLVRRITNRQGRRLARRKKHRRVRLNRLFEEESGLITDFTKI SINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDASDDG NSSVGDYAQIVKENSQKLETKTPGQIQRLRYQTYGQLRGDFTVEK DGKKHRLINVFPTSAIRSEALRLQTQQEFNPQITDEFINRYLEILT KRKYHGPNEKSRDYGRYRTSGETLDNIFGILIGKCTFYPPDEF AAKASYTAQEFNLLNDLNNLTVPETKKLSKEQKNQIINYVKNEK AMGPAKLKFIYAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMK LETLDIEQMDRETLDKLAIVLTNTEREGIQEAEHEFADGSFSQK QVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTIL TRLGKQKTTSSSNKTKYIDEKLLTEIYNPVVAKSVRQAIKIVNAI KEYGDFDNIVIAMARETNEDDEKKAQKIQKANKDEKDAAMLKAA NQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGTISIHD LINNSNQFEVDHILPLSITFDDSLANKVLVYATAAQEKGQRTPYQA LDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKF IERNLVDRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRH WGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLL DIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSK FNRKISDATIYATRQAKVGDKADETYVLGKIKDIYTQDGYDAFM KIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNP FLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGHNHIDITPKDSNN KVVQLQSVSPWRADVFNKTTGKYELGLKYADLQFEKGTGTYKIS QEKYNDIKKKEGVDSSEFKFTLYKNDLLLVKDTETKEQQLFRFLS RTMPKQKHVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLG KSNISYKVRTDVLGNQHIIKNEGDKPKLDFGGWQAAESYEVGGTA PLEEEYRLFLEAPIQNVTLLEQWKREIPKVAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNT PLLVRKSGTSEYRMVQDLREVNKRVTIHPTVPNPYTLLSLLPPDR IWYSVLDLKDFAFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQ GFKNSPTLNEALNRDLQGFRLDHPSPVSLQYVDDLLIAADTQAAC LSATRDLLMTLAEGLYRVSGKKAQLCQEEVTYLGFKIHKGSRLSN SRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRP GNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSG AAKGVLTAQALGPWKRPVAYLSKRLDPVAAAGWPRCLRAIAAAALL	151

	TREASKLTFGQDIEITSSHNLESLLRSPDKWLTNARITQYQVLLLD PPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLSTSRPDLTDQ PLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSDVIWAEPLPIGTSA QKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIRERGW LTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTG NRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTADGSEFESP KKKAKVE	
--	---	--

[0513] The template RNA sequences are provided in Table 41. Nucleotide modifications are noted as follows: phosphorothioate linkages denoted by an asterisk, 2'-O-methyl groups denoted by an 'm' preceding a nucleotide, 2-Fluororibose denoted by an 'i2F' preceding a nucleotide.

Table 41: Exemplary template RNA

ID	Sequence (with modifications)	SED ID NO	Sequence (no modifications)	SEQ ID NO
RNACS22390	mU*mG*mA*rGrArCrArUrGrCrUrUr CrCrArGrUrArCrArCrGrUrCrGrUrUrG rUrArCrUmCmUmGmGmGmAmCmU mUmCmGmGmUmCmCmCmAmGrAr cA1AT18_S2 l_TLock_RA R_U4G_OMe 004&F038_R1 4P11	152	UGAGACAUGCUCUUC CAGUACACGUCGU UGUACUCUGGGAC UUCGGUCCCAGAA GCUACAACGAUAA GGCUUCAUGCCGA AAUCAUGAGAUUC CAGGUGUACUGGA AGCA	153

LNP production

[0514] Briefly, LNPs were prepared using a molar ratio of 47.5% ionizable lipid: 39.95% cholesterol: 10% DSPC: 2.5% DMG-PEG2000 ; 0.05% galNAc lipid (Product #MV100100, Sussex Research Laboratories, Inc.) at a concentration in the ethanol phase corresponding to an N/P ratio of 6. The aqueous phase was composed of mRNA encoding the exemplary gene modifying polypeptide and the SERPINA1 template RNA at a mass ratio of 1:1 in 25 mM acetate buffer. The two phases were combined using a mixing tee at a ratio of 3:1 (aqueous phase: ethanol phase). Mixed LNPs were diluted with 1 volume of citrate-buffered saline and the resulting solution was neutralized by 1 vol% of 1M Tris-HCl, pH 8 after 15 minutes. LNPs were buffer exchanged into a suitable storage buffer and to remove ethanol using tangential flow filtration. The final LNPs were sterile filtered and were stored at -80C until use.

Administration of LNPs to cynomolgus macaques

[0515] Sexually mature male cynomolgus macaques (Biomere, Worcester, MA) were enrolled in the study and baseline weights and blood samples were collected. Between 1.5 to 2 hours prior to each test article dose administration, and at approximately 24 and 48-hours post-dose, animals were administered dexamethasone at a dose level of 2 mg/kg (1 mg/kg by intramuscular (IM) injection and 1 mg/kg by oral dose). For each animal, tLNPs were administered intravenously through a saphenous vein, cephalic vein, or tail vein with a temporary IV catheter on Day 1 at a dose of 3 mg/kg (dose volume of 5 mL/kg) via an approximate 60-minute infusion using an infusion pump, followed by about 0.5 mL of saline to flush the dose from the IV catheter. At 7 days after dosing, liver biopsies were collected and the samples were analyzed by using Amp-Seq to determine percent rewriting at the target locus in the genomic DNA and the expressed mRNA. To analyze gene editing activity, primers flanking the target locus were used to amplify across the locus in the genomic DNA or the cDNA of liver samples. Amplicons were analyzed via short read sequencing using an Illumina MiSeq.

Results

[0516] FIG. 38 shows that an average of about 36.3% of the SERPINA1 alleles in the hepatocytes of the dosed NHPs carried the precise edit following administration of the GalNAc-conjugated LNPs formulated with Lipid 232 comprising the exemplary gene modifying system, as assayed by Amp-Seq. An average of approximately 59.6 percent of the liver cDNA carried the intended precise edit at SERPINA1. Approximately 0.4% of the SERPINA1 alleles comprised unintended indels following administration of the Lipid 232 LNPs comprising the exemplary gene modifying system when compared to hepatocytes that were pre-dosed with the same LNPs except that they did not comprise the gene modifying system. The rewriting results are summarized in Table 42. The results show the exemplary Lipid 232 LNPs can effectively deliver a gene modifying system to NHP hepatocytes *in vivo* to introduce precise edits into the SERPINA1 gene locus.

Table 42: SERPINA 1 Rewriting Results in NHPs

Dose (mg/kg)	Animal	gDNA Ampseq (%)	Mean ± SD	cDNA Ampseq (%)	Mean ± SD	Indel (%)	Mean ± SD
3	1	45.3	36.3 ± 9.3	76.5	59.6 ± 17.9	0.4	0.4 ± 0.2
	2	26.8		40.9		0.2	

	3	36.8		61.5		0.6	
--	---	------	--	------	--	-----	--

Example 18: Lipid 232 is suitable for mRNA delivery to HSCs *in vivo*

[0517] **LNP Production:** LNPs were formulated with ionizable Lipid 232 at molar ratios of 47% ionizable lipid (Lipid 232): 42.5% cholesterol: 8.0% DSPC: 2.0% DMG-PEG2000: 0.5% DSPE-PEG2000-TCO using a N:P of 6. The aqueous phase was composed of mRNA encoding the eGFP in 25 mM acetate buffer. LNPs were formulated using microfluidic mixing using a Precision Nanosystems NanoAssemblr Ignite according to the manufacturer's protocol. A ratio of aqueous: organic solvent of 3:1 was maintained during mixing by using differential flow rates. After mixing, the LNPs were collected, diluted in citrate buffered solution pH = 6.0 (approximately 3:1 v/v). After 15 minutes, the diluted LNP solution was neutralized using 1 vol% Tris HCl pH = 8.0. Following neutralization, buffer exchange was conducted on the solution using dialysis in sucrose buffer (pH = 7.6, 9.0 wt% sucrose, 50 mM Tris, 50 mM salt) at 4°C. The buffer was replaced at 2 hours and dialysis continued overnight against a 10kDa filter. Each LNP formulation was concentrated using 100kDa Amicon Filters at 4C and 1500 RCF. The concentrated solution was then sterile filtered using a 0.2 µm syringe filter and measured using DLS. The filtered LNP samples were evaluated for mRNA content using the Ribogreen assay. Additionally, the pKa of LNPs was measured using the 2-(p-toluidinyl)naphthalene-6-sulfonic acid (TNS) assay. LNP solutions were diluted to the desired concentration and stored at -80°C for future use.

[0518] **Base Lipid (LNP) Screening in C57BL/6 mice:** Base LNPs formulated with Lipid 232. and an mRNA encoding GFP were administered to C57BL/6 mice as described in Example 9. Flow cytometry was then used to determine whether any of the Lipid 232 LNPs demonstrated GFP mRNA delivery to LSK cells and/or LT-HSCs relative.

[0519] Female C57BL/6 mice 6-8 weeks of age were purchased from Jackson Labs (Bar Harbor, ME; Catalog #000664). Following acclimation to the vivarium for 3-5 days, animals were randomized into groups of 4-5 and baseline weights were obtained. Animals were then treated via intravenous tail vein injection with 200 ul of candidate base LNPs formulated with GFP mRNA at a dose of 0.5 mg/kg. Following a 16-hour incubation, animals were euthanized via CO2 asphyxiation, weighed and their femurs were collected and then used to isolate bone marrow cells. In some cases, serum samples were collected to measure ALT and AST levels as biomarkers for liver damage.

[0520] **Flow cytometric analysis of C57BL/6 mouse bone marrow:** Bone marrow cells (1×10^6 in 0.1 ml PBS) from each test mouse were subsequently aliquoted to individual wells in 96 well plates and antibody fluorochrome conjugate mixtures (**Table 43**) were added to the cells in a total volume of 0.1 ml. Following a 30-minute incubation at room temperature, the plates were centrifuged for 5 minutes at 500G and supernatants were removed. The cells were then resuspended in 0.2 ml PBS and evaluated on an Aligent NovoCyte Quanteon (2010098) or Cyttek Northern Lights (N7-00008) flow cytometer using FlowJo software (Ashland, OR). Cells stained with relevant isotype control fluorochrome conjugated antibodies were used to set the negative control gates. HSCs and early progenitor cells were defined as hCD45+/Lin-/CD34+/CD38-, while long-term HSCs (LT-HSCs) were defined as hCD45+/Lin-/CD34+/CD38-/CD90+/CD45RA-.

Table 43. Flow cytometry Antibodies

Marker	Fluorophore	Clone	Vendor	Catalog
Live/Dead	Yellow		ThermoFisher	L34959
Lineage	PE	145-2C11 ; RB6-8C5 ; RA3-6B2 ; Ter-119 ; M1/70	Biolegend	133303
Sca1	APC	D7	Biolegend	108112
CD150	BV421	TC15-12F12.2	Biolegend	115926
CD48	PerCP-Cy5.5	HM48-1	Biolegend	103422
CD45	PEcy7	30-F11	Biolegend	103114
CD117	BV605	2B8	Biolegend	105847
CD11b	BV650	M1/70	Biolegend	101259

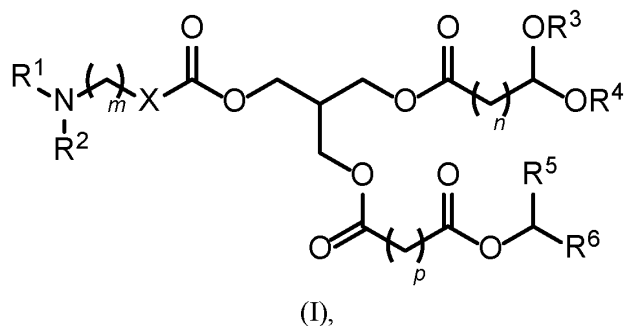
[0521] **Results:** Table 44 provides the results of this study. Base LNPs containing Lipid 232 delivered the mRNA payload to an average of 23% of the LSK cells and about 28% of the LT-HSCs in mice. The LSK cells and LT-HSCs exhibited higher GFP expression than the control cells.

Table 44: Lipid 232 Editing

Number	iLipid Tested	LNP Formulation (iLipid:chol:DSPC:PEG:PEG-TCO)	Average % GFP+ in LSK cells	Average GFP MFI in LSK cells	Average % GFP+ in LT-HSCs	Average GFP MFI in LT-HSCs
2	Control	-	0.35	1263	0.96	1129
4	Lipid 232	47:42.5:8:2:0.5	23.00	2019	28.08	2078

CLAIMS

1. A compound of formula (I):



5

or a pharmaceutically acceptable salt thereof, wherein:

X is -O- or -CH₂-;

m is 0, 1, 2, or 3;

10 R¹ is C₁₋₄alkyl;

R² is C₁₋₄alkyl;

n is 1, 2, 3, or 4;

R³ is C₄₋₁₀alkyl;

R⁴ is C₄₋₁₀alkyl;

15 p is 2, 3, 4, 5, or 6;

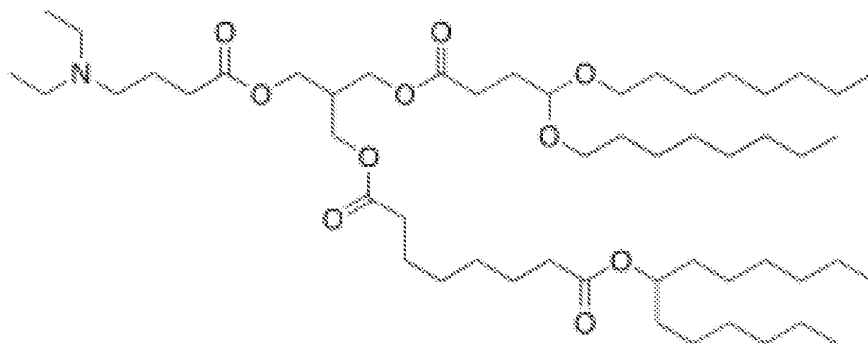
R⁵ is C₄₋₁₀alkyl; and

R⁶ is C₄₋₁₀alkyl.

2. A compound having a structure as set forth in Table 1.

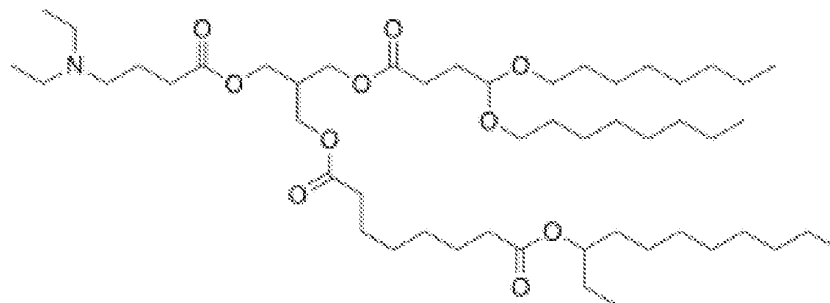
3. A compound having the structure:

20



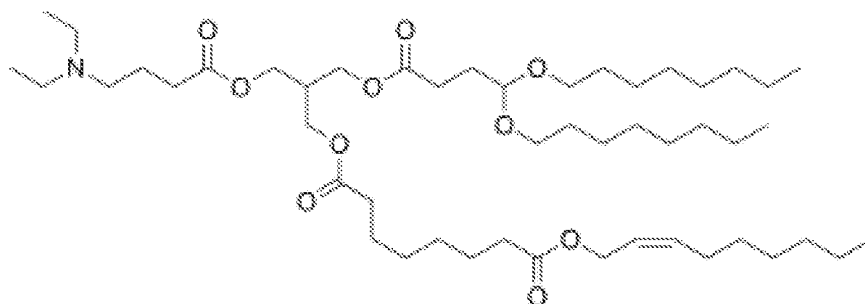
or a pharmaceutically acceptable salt thereof.

4. A compound having the structure:



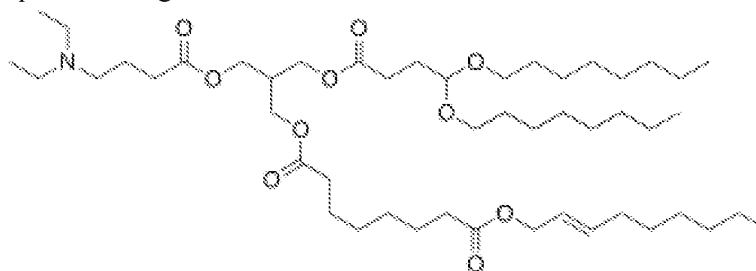
5 or a pharmaceutically acceptable salt thereof.

5. A compound having the structure:



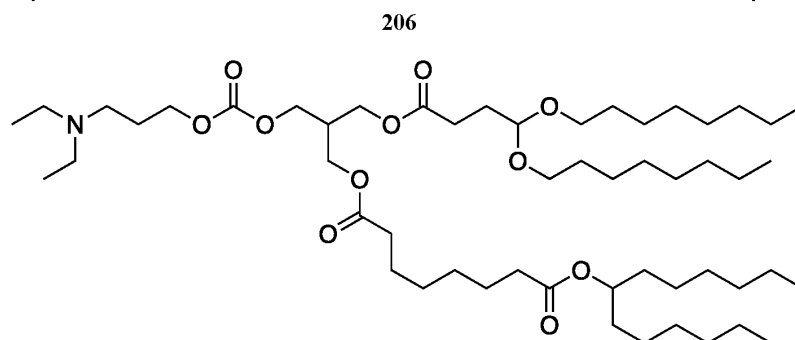
or a pharmaceutically acceptable salt thereof.

10 6. A compound having the structure:



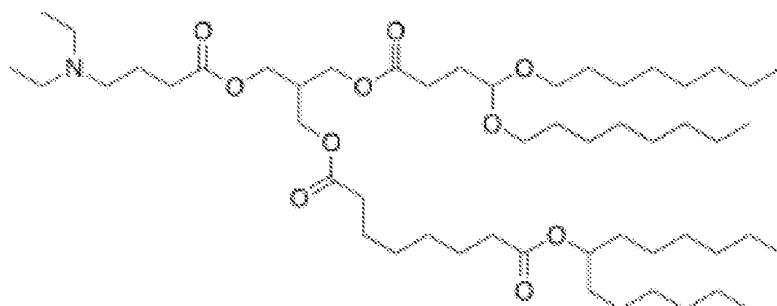
or a pharmaceutically acceptable salt thereof.

7. A compound having the structure:

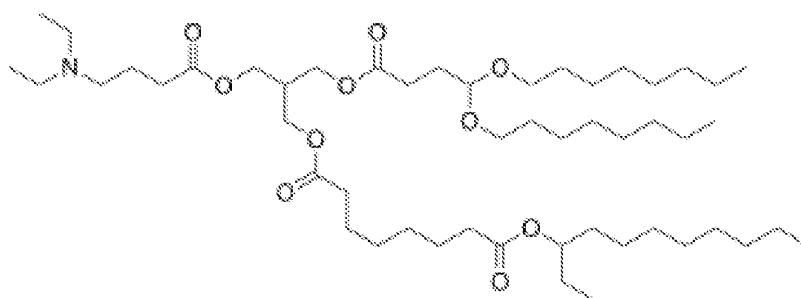


or a pharmaceutically acceptable salt thereof.

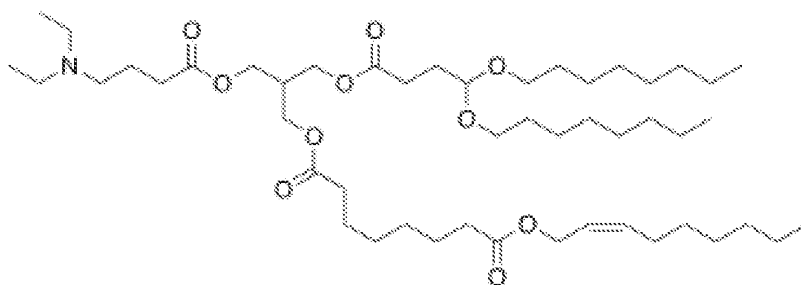
8. A lipid nanoparticle (LNP) comprising a compound of any one of claims 1-7.
9. The lipid nanoparticle of claim 8, further comprising a targeting moiety.
10. The lipid nanoparticle of claim 9, wherein the targeting moiety targets hematopoietic stem cells (HSCs).
11. The lipid nanoparticle of claim 9, wherein the targeting moiety targets immune cells.
12. The lipid nanoparticle of claim 11, wherein the immune cells are T cells.
13. The lipid nanoparticle of claim 9, wherein the targeting moiety targets liver cells.
14. A lipid nanoparticle (LNP) for delivery of a therapeutic agent to cells in vivo, the LNP comprising:
 - an ionizable lipid and a helper lipid, wherein the ionizable lipid is selected from the lipids in **Table 1**; and
 - a therapeutic agent encapsulated within the LNP.
15. The LNP of claim 14, wherein the ionizable lipid has the structure:



16. The LNP of claim 14, wherein the ionizable lipid has the structure:

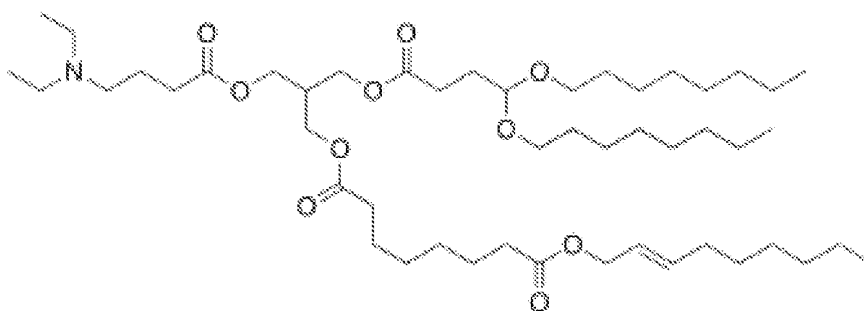


17. The LNP of claim 14, wherein the ionizable lipid has the structure:



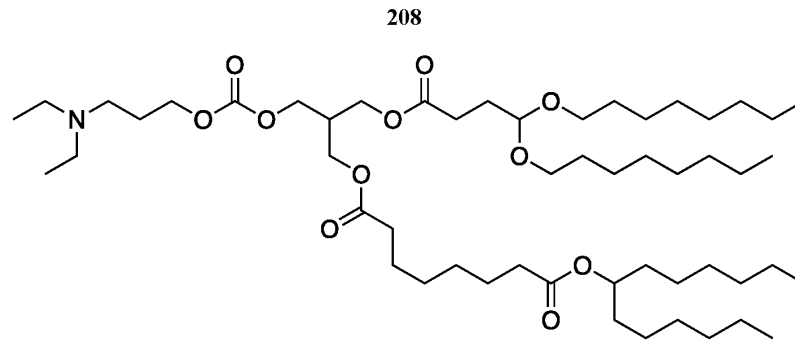
5

18. The LNP of claim 14, wherein the ionizable lipid has the structure:



19. The LNP of claim 14, wherein the ionizable lipid has the structure:

10



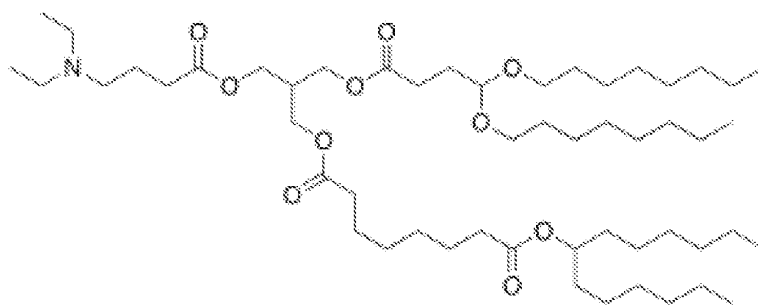
20. The LNP of any one of claims 14-19, wherein the cells are immune cells.
21. The LNP of claim 20, wherein the immune cells are T cells.
22. The LNP of any one of claims 14-19, wherein the cells are hematopoietic stem cells (HSCs).
23. The LNP of any one of claims 14-19, wherein the cells are liver cells.
24. The LNP of any one of claims 14-23, wherein the therapeutic agent comprises at least one nucleic acid molecule.
25. The LNP of claim 24, wherein the nucleic acid molecule is an mRNA molecule or a DNA molecule.
26. The LNP of claim 24, wherein the nucleic acid molecule is an mRNA molecule.
27. The LNP of claim 24, wherein the nucleic acid molecule is a siRNA or miRNA molecule.
28. The LNP of claim 24, wherein the therapeutic agent comprises two nucleic acid molecules.
29. The LNP of claim 28, wherein the two nucleic acid molecules are RNA molecules.
30. The LNP of claim 29, wherein one nucleic acid molecule comprises an mRNA molecule (e.g., an mRNA encoding a gene modifying polypeptide) and one nucleic acid molecule comprises an RNA molecule (e.g., a gRNA or template RNA).
31. The LNP of claim 30, wherein one nucleic acid molecule comprises an mRNA molecule and one nucleic acid molecule comprises a guide RNA.
32. The LNP of claim 31, wherein one nucleic acid molecule is an mRNA molecule and one nucleic acid molecule is a DNA molecule.

33. The LNP of any one of claims 28-32, wherein the at least one of the nucleic acid molecules encodes at least one component for altering a genome (e.g., a gene modifying polypeptide).
34. The LNP of claim 33, wherein the at least one component comprises an mRNA encoding a CRISPR-associated nuclease (Cas) and a guide RNA.
- 5 35. The LNP of claim 34, wherein the Cas is a Cas9 (e.g., spCas9) or a Cas12 (e.g., Cas12a).
36. The LNP any one of claims 28-35, wherein the mol% of ionizable lipids in the LNP ranges from about 35% to about 60%.
37. The LNP of claim 36, wherein the mol% of ionizable lipids in the LNP ranges from about 40% to about 50%.
- 10 38. The LNP of claim 37, wherein the mol% of ionizable lipids in the LNP ranges from about 45% to about 50%.
39. The LNP of any one of claims 28-38, wherein the helper lipid is selected from the group consisting of distearoyl-sn-glycero- phosphoethanolamine, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC),
 15 dioleoylphosphatidylglycerol (DOPG), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), dipalmitoylphosphatidylglycerol (DPPG), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyl-oleoylphosphatidylcholine (POPC), palmitoyl-oleoylphosphatidylethanolamine (POPE), dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane- 1 - carboxylate (DOPE-mal), dipalmitoyl
 20 phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidyl-ethanolamine (DSPE), monomethyl-phosphatidylethanolamine (such as 16-O-monomethyl PE), dimethyl- phosphatidylethanolamine (such as 16-O-dimethyl PE), 18-l-trans PE, 1-stearoyl-2-oleoyl- phosphatidylethanolamine (SOPE), hydrogenated soy phosphatidylcholine, egg phosphatidylcholine (EPC),
 25 dioleoylphosphatidylserine (DOPS), sphingomyelin (SM), dimyristoyl phosphatidylcholine (DMPC), dimyristoyl phosphatidylglycerol (DMPG), distearoylphosphatidylglycerol (DSPG), dierucoylphosphatidylcholine (DEPC), palmitoyl-oleoylphosphatidylglycerol (POPG), dielaidoyl-phosphatidylethanolamine (DEPE), lecithin, phosphatidylethanolamine, lysolecithin, lysophosphatidylethanolamine,
 30 phosphatidylserine, phosphatidylinositol, sphingomyelin, egg sphingomyelin (ESM),

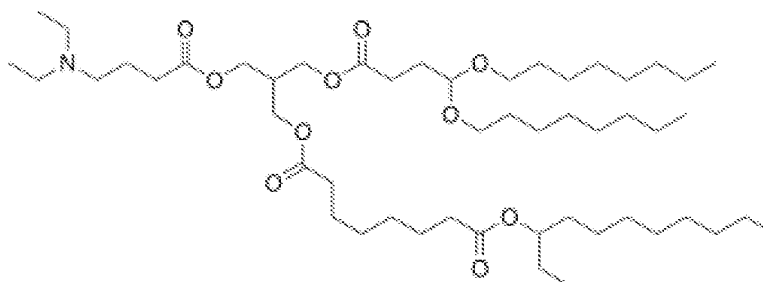
cephalin, cardiolipin, phosphatidic acid, cerebroside, dicetyl phosphate, lysophosphatidylcholine, dilinoleoylphosphatidylcholine, and mixtures thereof.

40. The LNP of claim 39 wherein the helper lipid is DSPC.
41. The LNP of claim 39, wherein the helper lipid is sphingomyelin.
- 5 42. The LNP of claim 41, wherein the sphingomyelin has a head group selected from phosphocholine, phosphoethanolamine and ceramide.
43. The LNP of claim 42, wherein the sphingomyelin is egg sphingomyelin.
44. The LNP of any one of claims 14-43, wherein the mol% of helper lipid is from about 18% to about 32%.
- 10 45. The LNP of claim 44, wherein the mol% of helper lipid is from about 20% to about 25%.
46. The LNP of claim 44, wherein the mol% of helper lipid is from about 22% to about 28%.
47. A method of treating and/or delaying the progression of a disease in a subject, the method comprising administering to the subject a therapeutically effective amount of an LNP of any one of claims 14-44.
- 15 48. The method of claim 47, wherein the disease is selected from any one of the diseases in Tables 18-23.
49. A conjugate for targeted delivery of a therapeutic agent to cells in vivo, the conjugate comprising:
- 20 a lipid nanoparticle (LNP) comprising an ionizable lipid and a helper lipid, wherein the ionizable lipid is selected from the lipids in **Table 1**;
- a plurality of targeting moieties conjugated to the LNP, wherein the plurality of targeting moieties is configured to target cells; and
- a therapeutic agent encapsulated within the LNP.
50. The conjugate of claim 49, wherein the ionizable lipid has the structure:

211

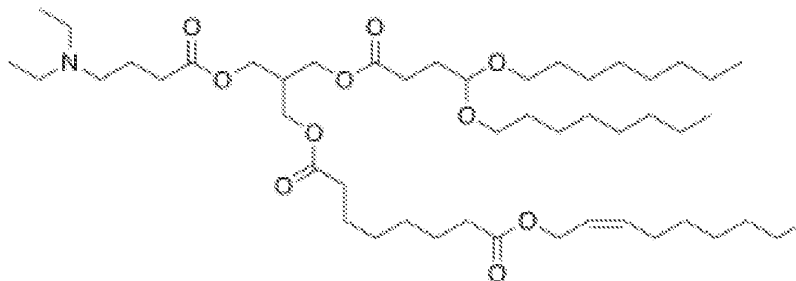


51. The conjugate of clam 49, wherein the ionizable lipid has the structure:

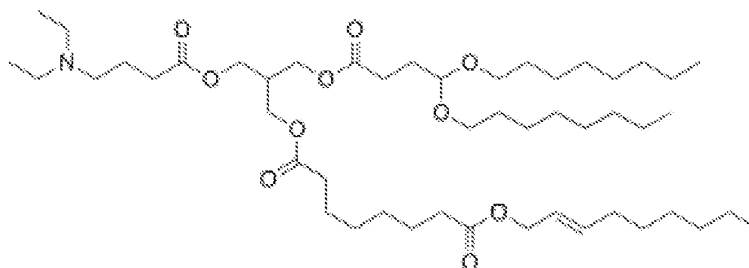


52. The conjugate of clam 49, wherein the ionizable lipid has the structure:

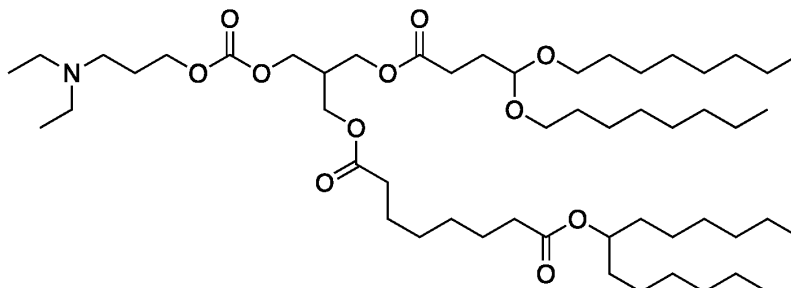
5



53. The conjugate of clam 49, wherein the ionizable lipid has the structure:



54. The conjugate of claim 49, wherein the ionizable lipid has the structure:



55. The conjugate of any one of claims 47-52, wherein the plurality of targeting moieties
5 is configured to target a receptor on the surface of HSCs.

56. The conjugate of claim 55, wherein the plurality of targeting moieties is configured to target a receptor on the surface of HSCs selected from the group consisting of CD33, CD34, CD38, CD43, CD59, CD105, CD123, CD164, CD338, CD71, CD117, CD50, CD49d, CD46, and CD184.

10 57. The conjugate of claim 56, wherein the plurality of targeting moieties is configured to target CD117.

58. The conjugate of claim 56, wherein the plurality of targeting moieties is configured to target CD71.

15 59. The conjugate of claim 56, wherein the plurality of targeting moieties is configured to target CD34.

60. The conjugate of any one of claims 49-54, wherein the plurality of targeting moieties is configured to target a receptor on the surface of immune cells.

61. The conjugate of claim 60, wherein the immune cells are T cells.

20 62. The conjugate of any one of claims 49-54, wherein the plurality of targeting moieties is configured to target a receptor on the surface of liver cells.

63. The conjugate of any one of claims 49-62, wherein the plurality of targeting moieties comprises antibodies or fragments thereof.

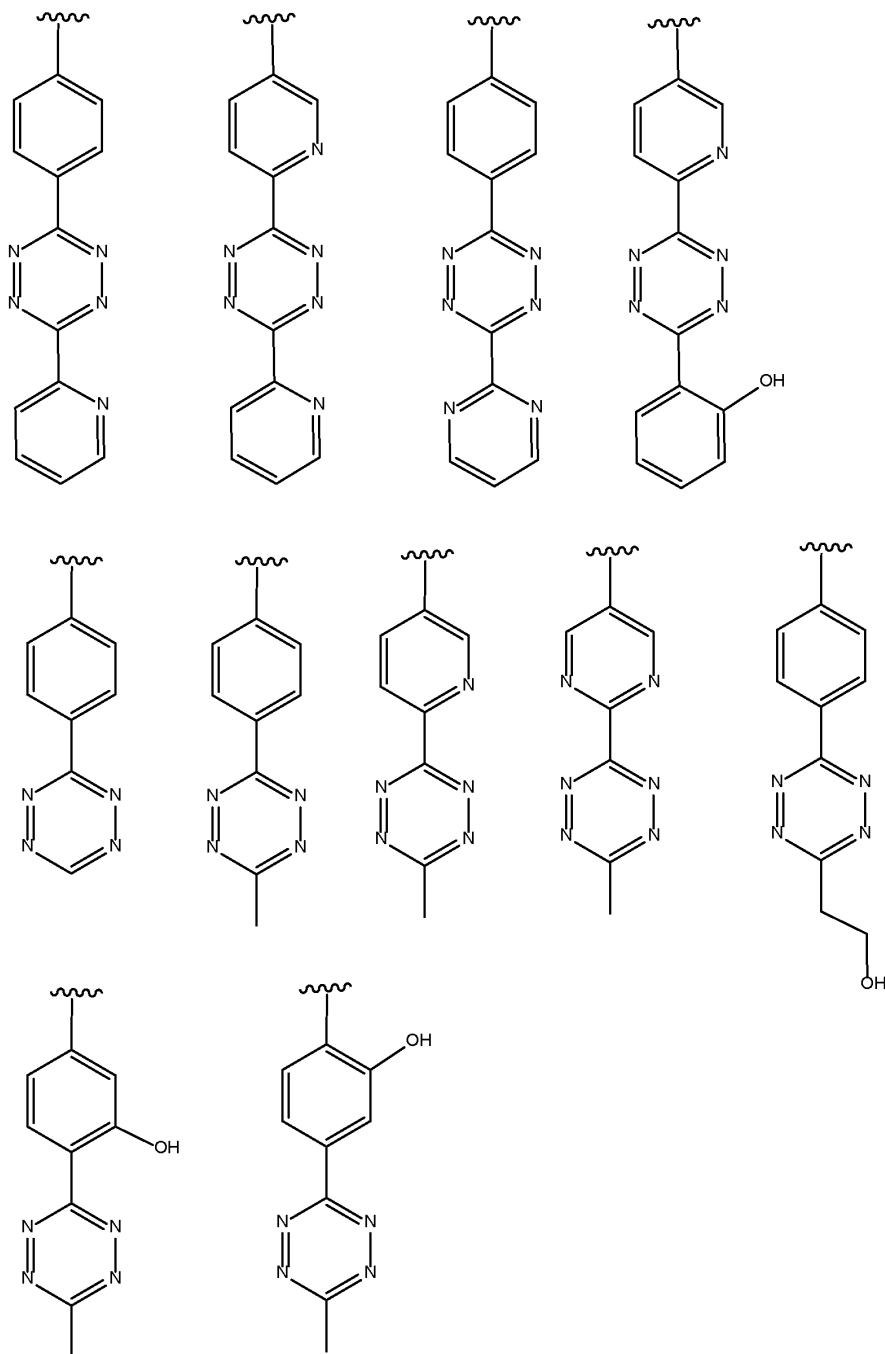
64. The conjugate of any one of claims 49-62, wherein the plurality of targeting moieties comprises Fab fragments.
65. The conjugate of any one of claims 49-62, wherein the plurality of targeting moieties comprises single chain variable fragments (scFv).
- 5 66. The conjugate of any one of claims 49-65, wherein the therapeutic agent comprises at least one nucleic acid molecule.
67. The conjugate of claim 66, wherein the nucleic acid molecule is an mRNA molecule or a DNA molecule.
68. The conjugate of claim 66, wherein the nucleic acid molecule is an mRNA molecule.
- 10 69. The conjugate of claim 66, wherein the nucleic acid molecule is a siRNA or miRNA molecule.
70. The conjugate of claim 66, wherein the therapeutic agent comprises two nucleic acid molecules.
71. The conjugate of claim 70, wherein the two nucleic acid molecules are RNA molecules.
- 15 72. The conjugate of claim 71, wherein one nucleic acid molecule comprises an mRNA molecule (e.g., an mRNA encoding a gene modifying polypeptide) and one nucleic acid molecule comprises an RNA molecule (e.g., a guide RNA or template RNA).
73. The conjugate of claim 71, wherein one nucleic acid molecule comprises an mRNA molecule and one nucleic acid molecule comprises a guide RNA.
- 20 74. The conjugate of claim 71, wherein one nucleic acid molecule is an mRNA molecule and one nucleic acid molecule is a DNA molecule.
75. The conjugate of any one of claims 68-74, wherein the at least one of the nucleic acid molecules encodes at least one component for altering a genome, e.g., a gene modifying polypeptide.
- 25 76. The conjugate of claim 75, wherein the at least one component comprises an mRNA encoding a CRISPR-associated nuclease (Cas) and a guide RNA.

77. The conjugate of claim 76, wherein the Cas is a Cas9 (e.g., spCas9) or a Cas12 (e.g., Cas12a).
78. The conjugate of claim 76, wherein the guide RNA is a single guide RNA.
79. The conjugate of any one of claims 49-76, wherein the mol% of ionizable lipids in the LNP
 5 ranges from about 35% to about 60%.
80. The conjugate of claim 79, wherein the mol% of ionizable lipids in the LNP ranges from about 40% to about 50%.
81. The conjugate of claim 79, wherein the mol% of ionizable lipids in the LNP ranges from about 45% to about 50%.
- 10 82. The conjugate of any one of claims 49-81, wherein the helper lipid is selected from the group consisting of distearoyl-sn-glycero- phosphoethanolamine, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylglycerol (DOPG), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), dipalmitoylphosphatidylglycerol
 15 (DPPG), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyl-oleoylphosphatidylcholine (POPC), palmitoyl-oleoylphosphatidylethanolamine (POPE), dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane- 1 -carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidyl-ethanolamine
 20 (DSPE), monomethyl-phosphatidylethanolamine (such as 16-O-monomethyl PE), dimethyl- phosphatidylethanolamine (such as 16-O-dimethyl PE), 18-l-trans PE, l-stearoyl-2-oleoyl- phosphatidylethanolamine (SOPE), hydrogenated soy phosphatidylcholine, egg phosphatidylcholine (EPC), dioleoylphosphatidylserine (DOPS), sphingomyelin (SM), dimyristoyl phosphatidylcholine (DMPC), dimyristoyl phosphatidylglycerol (DMPG),
 25 distearoylphosphatidylglycerol (DSPG), di-erucoylphosphatidylcholine (DEPC), palmitoyl-oleoylphosphatidylglycerol (POPG), di-erucoyl-phosphatidylethanolamine (DEPE), lecithin, phosphatidylethanolamine, lysolecithin, lysophosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, sphingomyelin, egg sphingomyelin (ESM), cephalin, cardiolipin, phosphatidic acid, cerebrosides, dicetylphosphate,
 30 lysophosphatidylcholine, dilinoleoylphosphatidylcholine, and mixtures thereof.

83. The conjugate of claim 82, wherein the helper lipid is DSPC.
84. The conjugate of claim 83, wherein the helper lipids is sphingomyelin.
85. The conjugate of claim 84, wherein the sphingomyelin has a head group selected from phosphocholine, phosphoethanolamine and ceramide.
- 5 86. The conjugate of claim 85, wherein the sphingomyelin is egg sphingomyelin.
87. The conjugate of any one of claims 49-86, wherein the mol% of helper lipid in the conjugate is from about 18% to about 32%.
88. The conjugate of claim 87, wherein the mol% of helper lipid in the conjugate is from about 20% to about 25%.
- 10 89. The conjugate of claim 88, wherein the mol% of helper lipid in the conjugate is from about 21% to about 23%.
90. The conjugate of claim 88, wherein the mol% of helper lipid in the conjugate is from about 22% to about 28%.
91. The conjugate of any one of claims 49-90, wherein the LNP further comprises one or more
- 15 PEGylated lipids.
92. The conjugate of claim 91, wherein the targeting moieties are conjugated to the LNP through the PEGylated lipids.
93. The conjugate of claim 91 or claim 92, wherein the PEGylated lipids each comprise a lipid component, wherein the lipid component of the PEGylated lipids is selected from the group
- 20 consisting of DMG, DPG, DSG, DTA, DOPE, DPPE, DMPE, DSPE, sphingosine, sphingomyelin, stearic acid, and any combination thereof.
94. The conjugate of claim 93 wherein the lipid component of the PEGylated lipids is DSPE.
95. The conjugate of any one of claims 91-94, wherein the mol% of PEGylated lipids in the LNP ranges from about 0.05% to about 2%.
- 25 96. The conjugate of any one of claims 49-95, wherein the plurality of lipid molecules comprises one or more cholesterol molecules.

97. The conjugate of claim 96, wherein the cholesterol molecules are selected from beta-sitosterol, hydroxycholesterol, stigmasterol, and any combination thereof.
98. The conjugate of claim 96 or claim 97, wherein the mol% of cholesterol molecules in the LNP ranges from about 35% to about 55%.
- 5 99. The conjugate of any one of claims 49-98, wherein the plurality of targeting moieties comprises more than 10 targeting moieties.
100. The conjugate of claim 99, wherein the plurality of targeting moieties comprises more than 20 targeting moieties.
101. The conjugate of claim 99, wherein the plurality of targeting moieties comprises more
10 than 50 targeting moieties.
102. The conjugate of claim 99, wherein the plurality of targeting moieties comprises more than 100 targeting moieties.
103. The conjugate of any one of claims 47-102, wherein the conjugate is formed via a Click reaction between a first Click handle on the targeting moiety and a second Click handle on
15 the LNP, thereby generating a Click product.
104. The conjugate of claim 103, wherein the second Click handle is covalently bonded to a pegylated lipid on the LNP.
105. The conjugate of claim 103 or claim 104, wherein the LNP is conjugated to the targeting moiety site specifically.
- 20 106. The conjugate of any one of claims 103-105, wherein the Click reaction is an inverse electron demand Diels-Alder reaction between a trans-cyclooctene (TCO) moiety on the first or second Click handle and a tetrazine ring on the first or second Click handle.
107. The conjugate of claim 106, wherein the first Click handle comprises a tetrazine ring and the second Click handle comprises a TCO moiety.
- 25 108. The conjugate of claim 106 or claim 107, wherein the tetrazine ring is unsubstituted.
109. The conjugate of claim 107 or claim 108, wherein the tetrazine ring is methyltetrazine.

110. The conjugate of claim 109, wherein the tetrazine ring is 6-methyltetrazine.
111. The conjugate of claim 110, wherein the tetrazine ring has one of the following structures:



112. The conjugate of any one of claims 106-111, wherein the first Click handle is conjugated to the targeting moiety through a sortase mediated ligation of a sortase recognition motif of the targeting moiety and a plurality of glycine residues conjugated to the first Click handle.
- 5 113. A method of treating and/or delaying the progression of a disease in a subject, the method comprising administering to the subject a therapeutically effective amount of a conjugate of any one of claims 49-112.
114. The method of claim 113, wherein the disease is selected from any one of the diseases in Tables 18-23.
- 10 115. A pharmaceutical composition comprising an LNP of any one of claims 8-46 or a conjugate of any one of claims 49-112.
116. A pharmaceutical composition comprising an LNP of claim 29 or a conjugate of claim 71.
117. The pharmaceutical composition of claim 116, wherein one RNA molecule comprises
15 a template RNA.
118. The pharmaceutical composition of claim 117, wherein the template RNA comprises a spacer of Table 5, Table 6, or Table 11.
119. The pharmaceutical composition of claim 117, wherein the template RNA comprises a heterologous sequence of Table 7, Table 8, or Table 12.
- 20 120. The pharmaceutical composition of claim 117, wherein the template RNA comprises a PBS sequence of Table 13.
121. The pharmaceutical composition of claim 117, wherein the template RNA comprises a heterologous object sequence and a PBS sequence of Table 7, Table 8 or Table 14.
122. The pharmaceutical composition of claim 117, wherein the template RNA comprises a
25 sequence of Table 9, Table 10, or Table 16.