

- (51) International Patent Classification:

A61K 47/69 (2017.01)

C12N 15/62 (2006.01)

A61K 9/127 (2025.01)

C12N 15/63 (2006.01)

C07K 14/005 (2006.01)

C12N 5/10 (2006.01)

C07K 14/47 (2006.01)

A61K 39/395 (2006.01)

C07K 19/00 (2006.01)
- (21) International Application Number:

PCT/US2024/033139
- (22) International Filing Date:

07 June 2024 (07.06.2024)
- (25) Filing Language:

English
- (26) Publication Language:

English
- (30) Priority Data:

63/506,776

07 June 2023 (07.06.2023)

US
- (71) Applicant: NVELOP THERAPEUTICS, INC. [US/US];

100 Technology Square, Floor 5, Cambridge, Massachusetts 02139 (US).
- (72) Inventors: CABECEIRAS, Peter; 100 Technology Square, Floor 5, Cambridge, Massachusetts 02139 (US).

PAJEROWSKI, Dave; 100 Technology Square, Floor 5, Cambridge, Massachusetts 02139 (US).

YOUNG, Matthew; 100 Technology Square, Floor 5, Cambridge, Massachusetts 02139 (US).

ADAMS, Noah; 100 Technology Square, Floor 5, Cambridge, Massachusetts 02139 (US).
- (74) Agent: HUANG, Wei et al.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, California 94304 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available):

AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

(54) Title: COMPOSITIONS AND METHODS FOR PROMOTING PAYLOAD DELIVERY IN CELLS

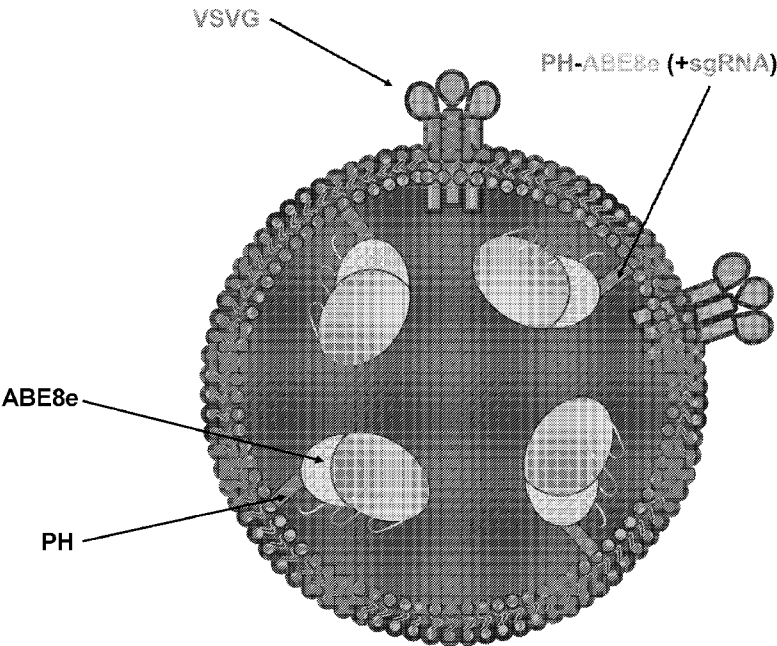


FIG. 4

(57) Abstract: Disclosed herein, in some aspects, are compositions of lipid delivery particles used to deliver a payload to a target cell and methods of producing the like. Lipid delivery particles can comprise ectosomes, exosomes, nanoparticles, or other classes of extracellular vesicles suitable for transporting a payload. Lipid delivery particles can further comprise a chimeric protein comprising a plasma recruitment element and one, two, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif; and/or (iv) a multimerization motif. Additionally, disclosed herein, in some aspects, are nucleic acid molecules encoding the same.

HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

COMPOSITIONS AND METHODS FOR PROMOTING PAYLOAD DELIVERY IN CELLS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application No. 63/506,776, filed on June 7, 2023, the entirety of which is incorporated herein by reference.

BACKGROUND

[0002] Vehicles for intercellular delivery, such as lipid delivery particles, enhance in vivo intercellular communication. These lipid delivery particles can be used to deliver a payload to a target cell. Developed gene editing agents and other therapeutic agents have broad therapeutic applications in the treatment of diseases and disorders. The delivery of these payloads to the cytosol of target cells however has posed significant challenges and there is a need for novel delivery particles that demonstrate efficient assembly and payload recruitment.

SUMMARY

[0003] In an aspect, provided herein is a lipid delivery particle that comprises: (a) a lipid membrane; and (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload; wherein the chimeric protein further comprises: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif that promotes assembly of the lipid delivery particle and/or promotes loading of the payload into the lipid delivery particle; or (iv) a multimerization motif.

[0004] In an aspect, provided herein is a lipid delivery particle that comprises: (a) a lipid membrane; and (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload; wherein the chimeric protein further comprises: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif selected from the group consisting of: myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, and farnesylation motif; or (iv) a multimerization motif.

[0005] In some embodiments, the lipid delivery particle is an ectosome, an exosome, or a viral-like particle.

[0006] In some embodiments, the lipid delivery particle is an ectosome, an exosome, or a viral-like particle.

[0007] In some embodiments, the plasma membrane recruitment element is a pleckstrin homology (PH) domain.

[0008] In some embodiments, the PH domain is from a protein selected from a group consisting of human phospholipase C δ 1, human Akt1, human Akt1 with E17K substitution, human 3-phosphoinositide-dependent protein kinase 1, human Daap1, mouse Grp1, human Grp1, human OSBP, human Btk1, human FAPP1, human CERT, human PKD, human PHLPP1, human SWAP70, and human MAPKAP1.

[0009] In some embodiments, the plasma membrane recruitment element is linked to a retroviral gag protein. In some embodiments, the plasma membrane recruitment element is linked to an N-terminus of the retroviral gag protein. In some embodiments, the plasma membrane recruitment element is linked to a gag-payload fusion construct. In some embodiments, the plasma membrane recruitment element is linked via its C-terminus to an N-terminus of the payload. In some embodiments, the plasma membrane recruitment element is linked via its N-terminus to a C-terminus of the payload.

[0010] In some embodiments, the payload comprises a therapeutic agent selected from the group consisting of an antibody, a cytokine, a therapeutic protein, a small molecule therapeutic, a therapeutic DNA, and a therapeutic RNA. In some embodiments, the payload comprises a ribonucleoprotein, a zinc-finger, a transcription activator-like effector (TALE), or a CRISPR-based genome editing or modulating protein.

[0011] In some embodiments, the chimeric protein comprises the particle budding motif.

[0012] In some embodiments, the particle budding motif is PTAP, PPXY, LPYX, PPEE, LYPL, LYPSL, PPPY, PPEY, PSAP, LYPAL, PPAP, PPPE, YMYL, YRKL, YQCL, YCYL, LYRTL, YPX_nL, or any combination thereof, and n is an integer from 1 to 5.

[0013] In some embodiments, the particle budding motif is linked to the payload or the plasma membrane recruitment element. In some embodiments, is linked between the payload and the plasma membrane recruitment element. In some embodiments, the particle budding motif is linked to the chimeric protein based on a construct arrangement listed in Table 12 or Table 15.

[0014] In some embodiments, the chimeric protein comprises the membrane penetrating peptide.

[0015] In some embodiments, the membrane penetrating peptide is linked to the payload or the plasma membrane recruitment element. In some embodiments, the membrane penetrating peptide is linked between the payload and the plasma membrane recruitment element. In some embodiments, the membrane penetrating peptide is linked to the chimeric protein based on a construct arrangement listed in Table 13.

[0016] In some embodiments, the chimeric protein comprises the post-translational modification motif.

[0017] In some embodiments, the chimeric protein comprises the post-translational modification motif to one or more amino acids of the payload or one or more amino acids of the plasma

membrane recruitment element. In some embodiments, the chimeric protein comprises the post-translational modification motif to one or more amino acids of a region in the chimeric protein outside of the payload and the plasma membrane recruitment element. In some embodiments, the chimeric protein comprises the post-translational modification motif to one or more amino acids of a region in the chimeric protein between the payload and the plasma membrane recruitment element. In some embodiments, the post translational modification motif is linked to the chimeric protein based on a construct arrangement listed in Table 14 or Table 15.

[0018] In some embodiments, the chimeric protein comprises the multimerization motif.

[0019] In some embodiments, the multimerization motif comprises a leucine zipper oligomerization motif. In some embodiments, the multimerization motif is linked to the payload or the plasma membrane recruitment element. In some embodiments, the multimerization motif is linked between the payload and the plasma membrane recruitment element. In some embodiments, the multimerization motif is linked to the chimeric protein based on a construct arrangement listed in Table 16.

[0020] In some embodiments, the chimeric protein comprises the particle budding motif and the membrane penetrating peptide.

[0021] In some embodiments, the chimeric protein comprises the particle budding motif and the post-translational modification motif.

[0022] In some embodiments, the chimeric protein comprises the particle budding motif and the multimerization motif.

[0023] In some embodiments, the chimeric protein comprises the membrane penetrating peptide and the post-translational modification motif.

[0024] In some embodiments, the chimeric protein comprises the membrane penetrating peptide and the multimerization motif.

[0025] In some embodiments, the chimeric protein comprises the post-translational modification motif and the multimerization motif.

[0026] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the post-translational modification motif.

[0027] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the multimerization motif.

[0028] In some embodiments, the chimeric protein comprises the particle budding motif, the post-translational modification motif, and the multimerization motif.

[0029] In some embodiments, the chimeric protein comprises the post-translational modification motif, the membrane penetrating peptide, and the multimerization motif.

[0030] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, and the multimerization motif.

[0031] In some embodiments, the chimeric protein comprises a tandem repeat of the particle budding motif, the membrane penetrating peptide, or the post-translational modification motif. In some embodiments, the tandem repeat of the particle budding motif further comprises a linker sequence.

[0032] In some embodiments, the chimeric protein further comprises a nuclear localization signal (NLS) or a nuclear export signal (NES).

[0033] In some embodiments, the nuclear localization signal is linked to the payload or the plasma membrane recruitment element.

[0034] In some embodiments, the nuclear localization signal comprises a sequence having at least about 80% sequence identity to any one of SEQ ID NOs: 268-280. In some embodiments, the nuclear localization signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 268-280. In some embodiments, the nuclear localization signal comprises a sequence as set forth in any one of SEQ ID NOs: 268-280.

[0035] In some embodiments, the lipid delivery particle comprises 2, 3, 4, 5, or more NLS sequences.

[0036] In some embodiments, the nuclear export signal is linked to the payload or the plasma membrane recruitment element. In some embodiments, the nuclear export signal is linked between the payload and the plasma membrane recruitment element. In some embodiments, the nuclear export signal is linked to the particle budding motif, the membrane penetrating particle, and/or the post-translational modification motif. In some embodiments, the nuclear export signal is linked to a cleavage site.

[0037] In some embodiments, the nuclear export signal comprises a sequence having at least about 80% sequence identity to any one of SEQ ID NOs: 87-267. In some embodiments, the nuclear export signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 87-267. In some embodiments, the nuclear export signal comprises a sequence as set forth in any one of SEQ ID NOs: 87-267.

[0038] In some embodiments, the lipid delivery particle comprises 2, 3, 4, 5, or more NES sequences.

[0039] In some embodiments, the chimeric protein comprises a linker sequence. In some embodiments, the linker sequence comprises a cleavable linker.

[0040] In some embodiments, the particle budding motif comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 1-32. In some embodiments, the particle budding motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 1-32. In some embodiments, the particle budding motif comprises a sequence as set forth in any one of SEQ ID NOs: 1-32.

[0041] In some embodiments, the membrane penetrating peptide comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 33-56. In some embodiments, the membrane penetrating peptide comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 33-56. In some embodiments, the membrane penetrating peptide comprises a sequence as set forth in any one of SEQ ID NOs: 33-56.

[0042] In some embodiments, the post-translational modification motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 57-86. In some embodiments, the post-translational modification motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 57-86. In some embodiments, the post-translational modification motif comprises a sequence as set forth in SEQ ID NOs: 57-86.

[0043] In some embodiments, the multimerization motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 281-285. In some embodiments, the multimerization motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 281-285. In some embodiments, the multimerization motif comprises a sequence as set forth in SEQ ID NOs: 281-285.

[0044] In some embodiments, the particle budding motif interacts with an endosomal sorting complex required for transport (ESCRT).

[0045] In some embodiments, the lipid delivery particle further comprises an envelope protein.

[0046] In some embodiments, the envelope protein comprises a VSV-G glycoprotein, a human immunodeficiency virus (HIV) GP160 glycoprotein, a Baboon Endogenous Retrovirus (BaEVTR) glycoprotein, a fusion protein of Vesicular stomatitis Indiana virus and Rabies virus glycoprotein (FuG-E), an ecotropic Murine Leukemia Virus envelope protein (MLV ENV ecotropic), a human T-cell lymphotropic virus type 1 (HTLV-1) glycoprotein, an amphotropic Murine Leukemia Virus envelope protein (MLV ENV amphotropic), a moloney murine leukemia virus 10A1 strain glycoprotein (MLV 10A1), a Baculovirus envelope glycoprotein (GP64), a pantropic MLV envelope protein, a xenotropic MLV envelope protein, a xenotropic murine

leukemia virus (XMLV) envelope protein, a Moloney murine leukemia virus (MMLV) envelope protein, a Moloney murine sarcoma virus (MoMSVg) envelope protein, a simian endogenous type D retrovirus protein (RD-114), a gibbon ape leukemia virus (GALV) envelope protein, a feline leukemia virus (FLV) envelope protein, a mouse mammary tumor virus (MMTV) envelope protein, an avian leukosis virus envelope protein, a Rous Sarcoma virus envelope protein, or an endogenous feline virus envelope protein (RD114 ENV), or a mutant thereof.

[0047] In some embodiments, the envelope protein comprises a HIV-1 envelope protein, a respiratory syncytial virus glycoprotein, a SARS-CoV glycoprotein, a rabies glycoprotein, a Mokola virus glycoprotein, a Semliki Forest virus glycoprotein, a Venezuelan equine encephalitis virus glycoprotein, an Ebola virus glycoprotein, a Marburg virus glycoprotein, or a mutant thereof.

[0048] In some embodiments, the envelope protein comprises a glycoprotein of a mammalian endogenous retrovirus or mutant thereof. In some embodiments, the glycoprotein of the mammalian endogenous retrovirus is a glycoprotein of a human endogenous retrovirus (hERV). In some embodiments, the glycoprotein of a hERV comprises a hENVH1, a hENVH2, a hENVH3, a hENVK1, a hENVK2, a hENVK3, a hENVK4, a hENVK5, a hENVK6, a hENVV, a hENVW, a hENVFRD, a hENVR, a hENVR(b), a hENVR(c)1, a hENVR(c)2 or a hENVK_{con}, or a mutant thereof.

[0049] In some embodiments, the envelope protein is fused to a targeting moiety, optionally a scFv.

[0050] In some embodiments, the envelope protein comprises a non-viral envelope protein.

[0051] In an aspect, provided herein is a chimeric protein comprising: (a) a plasma membrane recruitment element; (b) a payload; and (c) one, two, three, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif that promotes assembly of the lipid delivery particle and/or promotes loading of the payload into the lipid delivery particle; or (iv) a multimerization motif.

[0052] In an aspect, provided herein is a chimeric protein comprising: (a) a plasma membrane recruitment element; (b) a payload; and (c) one, two, three, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif selected from the group consisting of: myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, and farnesylation motif; or (iv) a multimerization motif.

[0053] In some embodiments, the plasma membrane recruitment element is a pleckstrin homology (PH) domain. In some embodiments, the plasma membrane recruitment element is a mutated PH domain or an engineered PH domain.

[0054] In some embodiments, the PH domain is from a protein selected from a group consisting of human phospholipase C δ 1, human Akt1, human Akt1 with E17K substitution, human 3-

phosphoinositide-dependent protein kinase 1, human Daap1, mouse Grp1, human Grp1, human OSBP, human Btk1, human FAPP1, human CERT, human PKD, human PHLPP1, human SWAP70, and human MAPKAP1.

[0055] In some embodiments, the plasma membrane recruitment element is linked to a retroviral gag protein. In some embodiments, the plasma membrane recruitment element is linked to an N-terminus of the retroviral gag protein. In some embodiments, the plasma membrane recruitment element is linked to a gag-payload fusion construct. In some embodiments, the plasma membrane recruitment element is linked via its C-terminus to an N-terminus of the payload. In some embodiments, the plasma membrane recruitment element is linked via its N-terminus to a C-terminus of the payload.

[0056] In some embodiments, the payload comprises a therapeutic agent selected from the group consisting of an antibody, a cytokine, a therapeutic protein, a small molecule therapeutic, a therapeutic DNA, and a therapeutic RNA. In some embodiments, the payload comprises a ribonucleoprotein, a zinc-finger, a transcription activator-like effector (TALE), or a CRISPR-based genome editing or modulating protein.

[0057] In some embodiments, the chimeric protein comprises the particle budding motif.

[0058] In some embodiments, the particle budding motif is PTAP, PPXY, LPYX, PPEE, LYPL, LYPSL, PPPY, PPEY, PSAP, LYPAL, PPAP, PPPE, YMYL, YRKL, YQCL, YCYL, LYRTL, YPX_nL, or any combination thereof, and n is an integer from 1 to 5.

[0059] In some embodiments, the particle budding motif is linked to the payload or the plasma membrane recruitment element. In some embodiments, the particle budding motif is linked between the payload and the plasma membrane recruitment element. In some embodiments, the particle budding motif is linked to the chimeric protein based on a construct arrangement listed in Table 12 or Table 15.

[0060] In some embodiments, the chimeric protein comprises the membrane penetrating peptide.

[0061] In some embodiments, the membrane penetrating peptide is linked to the payload or the plasma membrane recruitment element. In some embodiments, the membrane penetrating peptide is linked between the payload and the plasma membrane recruitment element. In some embodiments, the membrane penetrating peptide is linked to the chimeric protein based on a construct arrangement listed in Table 13.

[0062] In some embodiments, the chimeric protein comprises the post-translational modification motif.

[0063] In some embodiments, the chimeric protein comprises the post-translational modification motif to one or more amino acids of the payload or one or more amino acids of the plasma membrane recruitment element. In some embodiments, the chimeric protein comprises the post-

translational modification motif to one or more amino acids of a region in the chimeric protein outside of the payload and the plasma membrane recruitment element. In some embodiments, the chimeric protein comprises the post-translational modification motif to one or more amino acids of a region in the chimeric protein between the payload and the plasma membrane recruitment element. In some embodiments, the post-translational modification motif is linked to the chimeric protein based on a construct arrangement listed in Table 14 or Table 15.

[0064] In some embodiments, the chimeric protein comprises the multimerization motif. In some embodiments, the multimerization motif comprises a leucine zipper oligomerization motif.

[0065] In some embodiments, the multimerization motif is linked to the payload or the plasma membrane recruitment element. In some embodiments, the multimerization motif is linked between the payload and the plasma membrane recruitment element. In some embodiments, the multimerization motif is linked to the chimeric protein based on a construct arrangement listed in Table 16.

[0066] In some embodiments, the chimeric protein comprises the particle budding motif and the membrane penetrating peptide.

[0067] In some embodiments, the chimeric protein comprises the particle budding motif and the post-translational modification motif.

[0068] In some embodiments, the chimeric protein comprises the particle budding motif and the multimerization motif.

[0069] In some embodiments, the chimeric protein comprises the membrane penetrating peptide and the post-translational modification motif.

[0070] In some embodiments, the chimeric protein comprises the membrane penetrating peptide and the multimerization motif.

[0071] In some embodiments, the chimeric protein comprises the post-translational modification motif and the multimerization motif.

[0072] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the post-translational modification motif.

[0073] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the multimerization motif.

[0074] In some embodiments, the chimeric protein comprises the particle budding motif, the post-translational modification motif, and the multimerization motif.

[0075] In some embodiments, the chimeric protein comprises the post-translational modification motif, the membrane penetrating peptide, and the multimerization motif.

[0076] In some embodiments, the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, and the multimerization motif.

[0077] In some embodiments, the chimeric protein comprises a tandem repeat of the particle budding motif, the membrane penetrating peptide, or the post-translational modification motif. In some embodiments, the tandem repeat of the particle budding motif further comprises a linker sequence.

[0078] In some embodiments, the chimeric protein further comprises a nuclear localization signal (NLS) or a nuclear export signal (NES).

[0079] In some embodiments, the nuclear localization signal is linked to the payload or the plasma membrane recruitment element.

[0080] In some embodiments, the nuclear localization signal comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 268-280. In some embodiments, the nuclear localization signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in SEQ ID NOs: 268-280. In some embodiments, the nuclear localization signal comprises a sequence as set forth in SEQ ID NOs: 268-280.

[0081] In some embodiments, wherein the chimeric protein comprises 2, 3, 4, 5, or more NLS sequences.

[0082] In some embodiments, the nuclear export signal is linked to the payload or the plasma membrane recruitment element. In some embodiments, the nuclear export signal is linked between the payload and the plasma membrane recruitment element. In some embodiments, the nuclear export signal is linked to the particle budding motif, the membrane penetrating particle, the post-translational modification motif, and/or the multimerization motif. In some embodiments, the nuclear export signal is linked to a cleavage site.

[0083] In some embodiments, the nuclear export signal comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 87-267. In some embodiments, the nuclear export signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in SEQ ID NOs: 87-267. In some embodiments, the nuclear export signal comprises a sequence as set forth in SEQ ID NOs: 87-267.

[0084] In some embodiments, the chimeric protein comprises 2, 3, 4, 5, or more NES sequences.

[0085] In some embodiments, the chimeric protein further comprises a linker sequence. In some embodiments, the linker sequence comprises a cleavable linker.

[0086] In some embodiments, the post-translational modification comprises myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, farnesylation motif, or any combination thereof.

[0087] In some embodiments, the particle budding motif comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 1-32. In some embodiments, the particle budding motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 1-32. In some embodiments, the particle budding motif comprises a sequence as set forth in any one of SEQ ID NOs: 1-32.

[0088] In some embodiments, the membrane penetrating peptide comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 33-56. In some embodiments, the membrane penetrating peptide comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 33-56. In some embodiments, the membrane penetrating peptide comprises a sequence as set forth in any one of SEQ ID NOs: 33-56.

[0089] In some embodiments, the post-translational modification motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 57-86. In some embodiments, the post-translational modification motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 57-86. In some embodiments, the post-translational modification motif comprises a sequence as set forth in SEQ ID NOs: 57-86.

[0090] In some embodiments, the multimerization motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 281-285. In some embodiments, the multimerization motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 281-285. In some embodiments, the multimerization motif comprises a sequence as set forth in SEQ ID NOs: 281-285.

[0091] In some embodiments, the particle budding motif interacts with an endosomal sorting complex required for transport (ESCRT).

[0092] In an aspect, provided herein is a method for preparing a lipid delivery particle comprising a payload, the method comprising: (a) providing a producer cell that produces the lipid delivery particle described herein; and (b) collecting the lipid delivery particle from the producer cell.

[0093] In some embodiments, the producer cell comprises a nucleic acid sequence encoding the chimeric protein.

[0094] In some embodiments, the method further comprises expressing in the producer cell the nucleic acid molecule encoding the chimeric protein. In some embodiments, the method further comprises transiently expressing in the producer cell the nucleic acid molecule encoding the

chimeric protein. In some embodiments, the method further comprises permanently expressing in the producer cell the nucleic acid molecule encoding the chimeric protein.

[0095] In some embodiments, the method further comprises culturing the producer cell in a medium and maintaining the producer under conditions to produce the lipid delivery particle.

[0096] In some embodiments, the method further comprises harvesting the medium and purifying the extracellular vesicle. In some embodiments, the purifying retains the structural integrity of the extracellular vesicle.

[0097] In an aspect, provided herein is a lipid delivery particle produced by the method described herein.

[0098] In an aspect, provided herein is a nucleic acid sequence encoding the chimeric protein described herein.

[0099] In an aspect, provided herein is a composition comprising: (a) a first nucleic acid sequence encoding an envelope protein; and (b) a second nucleic acid sequence encoding a chimeric protein, wherein the chimeric protein comprises: (i) a plasma membrane recruitment element; (ii) a payload; and (iii) one, two, three, or all of: a particle budding motif; a membrane penetrating peptide; a post-translational modification motif; or a multimerization motif.

[0100] In an aspect, provided herein is a vector comprising the nucleic acid sequence described herein.

[0101] In some embodiments, the vector is a vector selected from a plasmid, a cosmid, a bacterial vector, a viral vector, an artificial chromosome.

[0102] In an aspect, provided herein is a cell comprising the lipid delivery particle described herein.

[0103] In an aspect, provided herein is a cell produced by the method described herein.

[0104] In an aspect, provided herein is a pharmaceutical formulation comprising the lipid delivery particle described herein and a pharmaceutically acceptable carrier, excipient, adjuvant or diluent.

[0105] In an aspect, provided herein is a kit comprising: (a) the lipid delivery particle described herein, or the pharmaceutical formulation described herein; and (b) instructions for use of the lipid delivery particle or the pharmaceutical formulation.

INCORPORATION BY REFERENCE

[0106] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0107] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0108] **FIGs. 1A-1E** are schematic representations of constructs encoding exemplary chimeric protein according to some embodiments of the present disclosure. **FIG. 1A** shows a construct with a promoter sequence (pCMV), sequence encoding a payload, and sequence encoding a plasma membrane recruitment element such as a PH domain (new PH or mutant PH), which is located upstream of (5' to) or downstream of (3' to) the nucleic acid encoding the payload. **FIG. 1B** shows a construct with nucleic sequences encoding a payload, a PH domain, and a particle budding motif (also named as budding domain, or BD). **FIG. 1C** shows a construct with nucleic acid sequences encoding a payload, a PH domain, and a nuclear localization signal (NLS) or nuclear export signal (NES). **FIG. 1D** shows a construct with nucleic acid sequences encoding a payload, a PH domain, and a membrane penetrating peptide (MPP) motif. **FIG. 1E** shows a construct with nucleic acid sequences encoding a payload, a PH domain, and a post-translational modification (PTM) motif. **FIG. 1F** shows a chimeric protein containing a plasma membrane recruitment element (PMRE), a repeat of two particle budding motifs (BD), a multimerization motif, such as leucine zipper oligomerization motif (LZ), a third particle budding motif, and a payload.

[0110] **FIG. 2** is an exemplary process diagram for manufacturing lipid delivery particles of the present disclosure that can be used to deliver a payload.

[0111] **FIG. 3** is a schematic representation of an exemplary construct for a pleckstrin homology domain (PH domain of protein lipase C, PLC PH) fused to a late domain via a linker.

[0112] **FIG. 4** is a schematic representation of a lipid delivery particle. The particle comprises a VSVG envelope protein and a pleckstrin homology domain (PH) fused to a payload (e.g., ABE8e, which is coupled with sgRNA).

[0113] **FIG. 5A** is a schematic representation of constructs for the example lipid delivery particle of **FIG. 4**. The three constructs comprise (i) a pU6 plasmid encoding sgRNA, (ii) a plasmid with a CMV promoter, followed by sequences encoding adenine base editor payload (ABE8e) and PH domain, and (iii) a plasmid with a CMV promoter followed by sequence encoding VSVG envelope.

[0114] **FIG. 5B** is a schematic representation of the mechanism for formation of the lipid delivery particle. The PH domain fused to the nuclease buds from the 293T cell forming a particle with VSVG envelope protein and delivers its payload into the transduced cell.

[0115] **FIG. 6A** is a schematic representation of an exemplary chimeric protein comprising a payload and PH domain (e.g., AKT (also named as Akt1) with E17K substitution). **FIG. 6B**

depicts results of experiments testing effects of various nuclear localization signal (NLS)/nuclear export signal (NES) on base editing efficiency of the resultant lipid delivery particles. Percent adenine to guanine change of Bcl1a gene in 293T cells was measured as a readout of the tests following delivery of ABE8e.

[0116] FIGs. 7A-7B depict results of experiments assessing effects of various budding domains on base editing efficacy. **FIG. 7A** depicts results of high dose of delivery particles (30 μ l 100x PEG-concentrated particles). **FIG. 7B** depicts results of low dose of delivery particles (1 μ l 50x PEG-concentrated particles). Both the high-dose and low-dose of delivery particles showed that PPPY + PSAP motif had the greatest editing percentage.

[0117] FIG. 8A-8B depict the structure and results of an experiment assessing effects of budding domains on delivery particles on base editing efficacy. **FIG. 8A** shows schematic representations of a subset of tested constructs. ABE8e-PH-NLS, having wildtype AKT PH domain, corresponds to AKT WT in **FIG. 8B**, ABE8e-PH(E17K)-NLS, having AKT PH domain with E17K substitution, corresponds to AKT E17K in **FIG. 8B**, and ABE8e-PH(E17K)-BD-NLS, having Akt1 PH domain with E17K substitution and budding domain, corresponds to AKT E17K 3x BD in **FIG. 8B**. **FIG. 8B** shows the results of the experiment measuring A to G conversion. At 1.0 μ l, addition of the budding domain (BD) improved editing percentage compared to controls (ABE8e with no PH and AKT WT).

[0118] FIG. 9 depicts a schematic representation of an exemplary payload construct comprising a multimerization domain (MD). The multimerization domain has a fusion at the C-terminus.

[0119] FIGs. 10A-10B depict the structure and results of an experiment to assess effects of multimerization domains on delivery particles on base editing efficacy. **FIG. 10A** shows schematic representations of a subset of tested constructs. ABE8e-PH-NLS, having wildtype AKT PH domain, corresponds to AKT WT in **FIG. 10B**, ABE8e-PH(E17K)-NLS, having AKT PH domain with E17K substitution, corresponds to AKT E17K in **FIG. 10B**, and ABE8e-PH(E17K)-MD-NLS, having wildtype AKT PH domain with E17K substitution and a multimerization domain, corresponds to MD Dimer in **FIG. 10B**. **FIG. 10B** shows the results of the analysis measuring A to G conversion with MD domains. At 1.0 μ l dose, addition of the dimerization domain (MD) improved editing percentage compared to controls (ABE8e with no PH and AKT WT).

[0120] FIGs 11A-11B depict constructs and results examining effects of addition of post-translational modifications to delivery particles on base editing efficacy. **FIG. 11A** shows an exemplary construct comprising a post-translational modification (PTM) motif at the N-terminus of the payload. The construct comprises from N-terminus to C-terminus PTM, payload (ABE8e), AKT PH domain with E17K mutation, and NLS. **FIG. 11B** shows results of the experiment to

assess editing efficiency across original delivery particles (Original), particles with multimerization domain (MD), particles with budding domain (BD), and particles with post-translational modification (e.g., myristoylation; PTM). Particles were tested at doses of 0.2, 2.0 and 20.0 μ l. At 2.0 μ l dose, addition of PTM, BD, and MD to the particle increased the editing percentage compared to that from the original particle (far right bars).

DETAILED DESCRIPTION

[0121] Disclosed herein, in some aspects, are lipid delivery particles comprising a phospholipid bilayer and a chimeric protein. The chimeric protein can comprise a plasma membrane recruitment element and a payload, and can further comprise: (i) a particle budding motif; (ii) a membrane penetrating peptide motif; or (iii) a post-translational modification motif, or (iv) a multimerization motif. In some cases, the lipid delivery particle is an extracellular vesicle. In some cases, the lipid delivery particle is an ectosome, an exosome, or a viral-like particle. In some cases, the chimeric protein can comprise one, two, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide motif; (iii) a post-translational modification motif; or (iv) a multimerization motif.

[0122] Without wishing to be bound by a certain theory, presence of one or more of (i) a particle budding motif; (ii) a membrane penetrating peptide motif; (iii) a post-translational modification motif; or (iv) a multimerization motif in the chimeric protein of the present disclosure can improve production of the lipid delivery particles. Without wishing to be bound by theory, presence of (i) a particle budding motif; (ii) a membrane penetrating peptide motif; (iii) a post-translational modification motif; or (iv) a multimerization motif in the chimeric protein of the present disclosure can increase the total number of particles produced, increase the amount of payload encapsulated per particle, and/or increase the efficiency of payload release. In some cases, additional motifs are present in the chimeric protein of the lipid delivery particle. The motifs of the present disclosure can also include polypeptide sequences that are additionally fused to the fusion of the plasma membrane recruitment element and the payload, or in some cases replacing or modifying elements of the fusion of the plasma membrane recruitment element and payload.

[0123] The plasma membrane recruitment elements and motifs can be placed at the N-terminus, C-terminus, or between the N- and C-terminus of the construct. In some cases, the plasma membrane recruitment elements and motifs can be present in repeating fashion or at various segments of the DNA.

[0124] In some cases, the plasma membrane recruitment element comprises a pleckstrin homology (PH) domain. In some cases, the PH domain is an engineered PH domain or a mutated PH domain. Without wishing to be bound by theory, the PH domain can assist recruiting a payload into the lipid delivery particle and enhance release of the payload into a target cell. Accordingly, the PH domain can be fused to the payload. In some cases, the chimeric protein further comprises

a gag protein, *e.g.*, a retroviral gag protein. In some cases, in the chimeric protein, the PH domain is fused to a gag-payload fusion construct.

[0125] The payload disclosed herein can comprise therapeutic agents. In some cases, the payload comprises gene editing agent including, but not limited to, a ribonucleoprotein, a zinc-finger, a transcription activator-like effector (TALE), or a CRISPR-based genome editing or modulating protein. Alternatively or additionally, the payload can comprise a biomarker, an imaging agent, a sugar or polysaccharide, or a small molecule.

[0126] The lipid delivery particle can comprise an envelope protein. In some cases, the envelope protein is of a viral origin, *e.g.*, having a sequence identical to a wildtype viral protein, or being a variant thereof (*e.g.*, having at least 80% sequence identity to the wildtype viral protein sequence). In some cases, the envelope protein is a VSV-G glycoprotein, a human immunodeficiency virus (HIV) GP160 glycoprotein, or a human T-cell lymphotropic virus type 1 (HTLV-1) glycoprotein. In some cases, the envelope protein is derived from a mammalian endogenous retrovirus. In some cases, the mammalian endogenous retrovirus is a koala retrovirus (KoRV) or a Jaagsiekte sheep retrovirus (enJSRV). In some cases, the envelope protein is derived from a human endogenous retrovirus. In some cases, the envelope protein is a non-viral protein.

[0127] In another aspect, the disclosure features a chimeric protein comprising a plasma membrane recruitment element, a payload, and one, two, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif; or (iv) a multimerization motif.

[0128] In another aspect, the disclosure features a nucleic acid molecule encoding a chimeric protein described herein.

[0129] In another aspect, the disclosure features a vector, *e.g.*, an expression vector, comprising the nucleic acid molecules described herein.

[0130] In some cases, the vector comprises a plasmid, a cosmid, a bacterial vector, a viral vector, an artificial chromosome, a liposome, a lipid nanoparticle, or an exosome.

[0131] In another aspect, the disclosure features a cell comprising a lipid delivery particle, a nucleic acid molecule or a vector described herein.

[0132] In another aspect, the disclosure features a method of making, *e.g.*, producing, a lipid delivery particle described herein, comprising providing a producer cell that produces the lipid delivery particle as described in the present disclosure.

[0133] In another aspect, the disclosure features a method of making, *e.g.*, producing, a lipid delivery particle described herein, comprising culturing a producer cell described herein under suitable conditions, *e.g.*, conditions suitable for gene expression.

[0134] In another aspect, the disclosure features a pharmaceutical formulation comprising a lipid delivery particle described herein.

[0135] In another aspect, the disclosure features a kit comprising a lipid delivery particle, a nucleic acid molecule or a pharmaceutical formulation as described herein.

[0136] Accordingly, provided herein are, *inter alia*, lipid delivery particles (*e.g.*, ectosomes, exosomes, or viral-like particles) that include the aforesaid motifs, nucleic acids encoding the same, and methods of producing the aforesaid particles.

DEFINITIONS

[0137] Certain specific details of this description are set forth in order to provide a thorough understanding of various embodiments. However, one skilled in the art will understand that the present disclosure may be practiced without these details. In other instances, well-known structures have not been shown or described in detail to avoid unnecessarily obscuring descriptions of the disclosure.

[0138] Unless the context requires otherwise, throughout the specification and claims which follow, the word “comprise” and variations thereof, such as, “comprises” and “comprising” are to be construed in an open, inclusive sense, that is, as “including, but not limited to.” Further, headings provided herein are for convenience only and do not interpret the scope or meaning of the claimed disclosure.

[0139] As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the content clearly dictates otherwise. The use of the words “a” or “an” when used in conjunction with the term “comprising” herein may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.”

[0140] It should also be noted that the term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise.

[0141] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below.

[0142] The term “about” when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 20\%$ or in some instances $\pm 10\%$, or in some instances $\pm 5\%$, or in some instances $\pm 1\%$, or in some instances $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods. As used herein, “about” and “approximately” generally mean an acceptable degree of error for the quantity measured given

the nature or precision of the measurements. Exemplary degrees of error are within 20 percent (%), typically, within 10%, and more typically, within 5% of a given range of values.

[0143] Calculations of homology or sequence identity between sequences (the terms are used interchangeably herein) are performed as follows. To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In some cases, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, 60%, and even more preferably at least 70%, 80%, 90%, 100% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid “identity” is equivalent to amino acid or nucleic acid “homology”).

[0144] The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In some cases, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch ((1970) J. Mol. Biol. 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In some cases, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A particularly preferred set of parameters (and the one that should be used unless otherwise specified) are a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frame shift gap penalty of 5.

[0145] The percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of E. Meyers and W. Miller ((1989) CABIOS, 4:11-17) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. The nucleic acid and protein sequences described herein can be used as a “query sequence” to perform a search against public databases to, for

example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J. Mol. Biol. 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecule of the disclosure. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the disclosure. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

[0146] It is understood that the molecules of the present disclosure may have additional conservative or non-essential amino acid substitutions, which do not have a substantial effect on their functions.

[0147] The terms “peptide,” “polypeptide,” and “protein” are used interchangeably, and refer to a compound comprised of amino acid residues covalently linked by peptide bonds. A protein or peptide must contain at least two amino acids, and no limitation is placed on the maximum number of amino acids that can comprise a protein’s or peptide’s sequence. Polypeptides include any peptide or protein comprising two or more amino acids joined to each other by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. “Polypeptides” include, for example, biologically active fragments, substantially homologous polypeptides, oligopeptides, homodimers, heterodimers, variants of polypeptides, modified polypeptides, derivatives, analogs, fusion proteins, among others. A polypeptide includes a natural peptide, a recombinant peptide, or a combination thereof.

[0148] The term “expression vector” refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an in vitro expression system. Expression vectors include all those known in the art, including cosmids, plasmids (e.g., naked or contained in liposomes) and viruses (e.g., lentiviruses, retroviruses, adenoviruses, and adeno-associated viruses) that incorporate the recombinant polynucleotide.

[0149] The term “promoter” refers to a DNA sequence recognized by the transcription machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a polynucleotide sequence.

[0150] An example of a promoter is the immediate early cytomegalovirus (CMV) promoter sequence. This promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. In some cases, a polynucleotide comprises a CMV promoter. In some cases, a polynucleotide may comprise a promoter selected from the group consisting of, but not limited to, a simian virus 40 (SV40) early promoter, a mouse mammary tumor virus (MMTV), a human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, a MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcomavirus (RSV) promoter, an actin promoter, a myosin promoter, an elongation factor-1a promoter, a hemoglobin promoter, and a creatine kinase promoter. In some cases, a polynucleotide comprises a sequence encoding a poly(A) tail. In some cases, a polynucleotide comprises a 3' UTR sequence. In some cases, a polynucleotide comprises a 5' UTR sequence. In some cases, a polynucleotide comprises multiple smaller and discrete nucleotide sequences that are typically heterologous and exhibit different and measurable function(s), a promoter sequence, a translation initiating sequence, a start codon, a polyadenylation sequence, a stop codon, and/or a linker sequence.

[0151] The term “subject” is intended to include living organisms in which an immune response can be elicited (*e.g.*, mammals, human). Exemplary subjects include humans, monkeys, dogs, cats, mice, rats, cows, horses, camels, goats, rabbits, and sheep. In certain embodiments, the subject is a human. A “patient” is a subject suffering from or at risk of developing a disease, disorder or condition or otherwise in need of the compositions and methods provided herein.

[0152] The term “therapeutic” as used herein means a treatment. A therapeutic effect is obtained by reduction, suppression, remission, or eradication of a disease state.

[0153] The term “prophylaxis” as used herein means the prevention of or protective treatment for a disease or disease state.

[0154] The term “transfected” or “transformed” or “transduced” refers to a process by which exogenous nucleic acid is transferred or introduced into the host cell. A “transfected” or “transformed” or “transduced” cell is one which has been transfected, transformed or transduced with exogenous nucleic acid. The cell includes the primary subject cell and its progeny. The term “transformed cell” or “transfected cell” can be a transformed or transfected cell in which inserted DNA can replicate either as an autonomously replicating plasmid or as part of the host chromosome. In some cases, a transfected cell or transformed cell can express the inserted DNA or RNA transiently (*e.g.*, for brief periods of time).

[0155] The terms “nucleic acid,” “nucleic acid sequence,” “nucleotide sequence,” or “polynucleotide sequence,” and “polynucleotide” are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or

analogues thereof. The polynucleotide may be either single-stranded or double-stranded, and if single-stranded may be the coding strand or non-coding (antisense) strand. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. The nucleic acid may be a recombinant polynucleotide, or a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which either does not occur in nature or is linked to another polynucleotide in a non-natural arrangement.

[0156] The term “cell line” or “producer cell line”, as used herein, refers to cultured cells that can be passed (e.g., divided) more than once. In some cases, a cell line can be passed more than 2 times, more than 5 times, more than 10 times, more than 50 times, more than 100 times, or more than 150 times. In some cases, a cell line can be passed between 2-250 times. In some cases, a cell line can be passed up to 200 times.

[0157] The term “lipid delivery particle” refers to lipid vesicles secreted by the plasma membrane of a cell having a membrane or envelope that surrounds a central internal space. Lipid delivery particles can be produced via budding out of the cell’s plasma membrane, secretion from lysosomes, or export by the Golgi complex. Lipid delivery particles can promote intercellular communication *in vitro* and *in vivo*. Exemplary lipid delivery particles can include, but are not limited to, exosomes, ectosomes (e.g., microvesicles and shedding vesicles), apoptotic bodies, microsomes, micelles, oncosomes, and liposomes.

[0158] Lipid delivery particles can have a range of size, such as a cross-sectional diameter, wherein the cross-sectional diameter has a range from about 1 nm to about 1200 nm, about 10 nm to about 1200 nm, about 25 nm to about 1200 nm, about 50 nm to about 1200 nm, about 100 nm to about 1200 nm, about 10 nm to about 1000 nm, about 25 nm to about 1000 nm, about 50 nm to about 1000 nm, about 100 nm to about 1000 nm, about 10 nm to about 500 nm, about 25 nm to about 500 nm, about 50 nm to about 500 nm or about 100 nm to about 500 nm (all inclusive). Lipid delivery particles can comprise a vesicle size of greater than 10 nm. Lipid delivery particles can comprise a vesicle size of less than 1000 nm.

[0159] The term “module” or “modification”, used interchangeably, can refer to an addition or a motif that is added to a payload or a chimeric protein within lipid delivery particles (e.g., extracellular vesicles and/or virus-like particles). The additions can enhance characteristics of particle assembly and/or payload recruitment. A module can occur in combination or alone. A module can be linked to a payload or another element (e.g., domain) in a construct.

[0160] The term “purifying,” “purified,” and “purify,” can be used interchangeably and refer to a preparation phase of a lipid delivery particle. A preparation of lipid delivery particles can comprise

a known or unknown concentration of a substance. In some cases, purifying produced lipid delivery particles comprises removing, either partially removing or wholly removing, a portion of the produced lipid delivery particles from a sample containing one or more biological components (*e.g.*, producer cells). In some cases, a composition comprising lipid delivery particles, as described herein, that have been purified are enriched as compared to a starting sample. Enrichment of the lipid delivery particles can comprise 5%, 10%, 25%, 50%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% enrichment compared to a starting sample. In some cases, purified lipid delivery particles can be free or substantially free of residual biological products. Residual biological products can include, but are not limited to, unwanted nucleic acids, proteins, lipids, and/or metabolites or abiotic materials such as including chemicals. In some cases, lipid delivery particles free of residual biological products comprise lipid delivery particles containing no detectable producer cells in the composition.

[0161] Ranges: throughout this disclosure, various aspects of the present disclosure can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the present disclosure. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. As another example, a range such as 95-99% identity, includes something with 95%, 96%, 97%, 98% or 99% identity, and includes subranges such as 96-99%, 96-98%, 96-97%, 97-99%, 97-98% and 98-99% identity. This applies regardless of the breadth of the range.

LIPID DELIVERY PARTICLE

[0162] In an aspect, described herein is a lipid delivery particle comprising a membrane and a chimeric protein. The chimeric protein can promote targeting of a payload to a recipient cell (*e.g.*, target cell). The lipid delivery particle can comprise an extracellular vesicle secreted by a cell. A lipid delivery particle can be secreted by cells *in vivo* or if cultured *in vitro*. Following secretion, the lipid delivery particle can deliver a payload and release its components into a recipient cell. In some cases, the lipid delivery particle can comprise an exosome, an ectosome (*e.g.*, a microvesicle or a shedding vesicle), an apoptotic body, a microsome, a micelle, an oncosome, or a liposome. In some cases, the lipid delivery particle can comprise a nanoparticle.

[0163] Ectosomes can be small heterogenous membrane-bound particles that are shed from the surface of the cell membrane in the extracellular space. Ectosomes can be shed by budding or

outward blebbing from the plasma membrane of the cell. Ectosomes can be shed from plasma membrane micro-domains (*e.g.*, lipid rafts or caveolae domains). Ectosomes have a series of biological functions, including, but not limited to, regulation of nucleic acid metabolism, signal transduction, transport, protein metabolism, intercellular communication, apoptosis, and cell growth. As described herein, ectosomes can comprise a payload that can be delivered to a target cell. Ectosomes can be variable in size. In some cases, an ectosome can be between 25 and 1000 nm. In some cases, the ectosome can be 30 nm, 50 nm, 75 nm, 100 nm, 200 nm, 300 nm, 400 nm, 500 nm, 600 nm, 700 nm, 800 nm, 900 nm, or 1000 nm in size. In some cases, the ectosome can be between about 80 and about 200 nm in size. In some cases, the ectosome is about 80 nm, about 90 nm, about 100 nm, about 110 nm, about 120 nm, about 130 nm, about 140 nm, about 150 nm, about 160 nm, about 170 nm, about 180 nm, about 190 nm, or about 200 nm in size. Formation of ectosomes can be regulated via the endosomal sorting complex required for transport (ESCRT), and/or by non-ESCRT-related mechanisms, including tetraspanins and membrane lipids.

[0164] Exosomes can be small vesicular particles formed by intraluminal vesicles within multivesicular bodies. Exosome can range from 30-125 nm. In some cases, an exosome can comprise 30 nm, 40 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, or 100 nm in size. Formation of exosomes can be regulated via the endosomal sorting complex required for transport (ESCRT), and/or by non-ESCRT-related mechanisms, including tetraspanins and membrane lipids.

[0165] Virus-like particles (VLPs) can be assemblies of viral proteins that can infect cells but lack viral general material. VLPs can be synthesized through expression of viral structural proteins. VLPs can be formed in vivo through assembly inside a producer cell via recombinant protein expression or in vitro through self-assembly in a vessel using quantities of purified protein. VLPs can be designed to deliver payloads to targeted cells and tissues. VLPs can be used to deliver to a recipient cell an agent (*e.g.*, prophylactic agent, therapeutic agent or diagnostic agent) or an enclosed circular or linear DNA or RNA molecule. Formation of VLPs can be regulated via the endosomal sorting complex required for transport (ESCRT), and/or by non-ESCRT-related mechanisms, including tetraspanins and membrane lipids.

[0166] Nanoparticles can be synthesized particles between 1 and 200 nm in size that can encapsulate payload for delivery to a recipient cell. In some cases, nanoparticles are engineered for a specific cell type and trigger endocytosis and later endosome lysis. In some cases, a nanoparticle can be about 5 nm, about 10 nm, about 15 nm, about 20 nm, about 25 nm, about 30 nm, about 35 nm, about 40 nm, about 45 nm, about 50 nm, about 60 nm, about 70 nm, about 80 nm, about 90 nm, about 100 nm, about 105 nm, about 110 nm, about 115 nm, about 120 nm, about 125 nm, about 130 nm, about 135 nm, about 140 nm, about 145 nm, about 150 nm, about 160 nm,

about 170 nm, about 180 nm, about 190 nm, or about 200 nm in size. Without wishing to be bound by theory, the small size can allow a higher intracellular uptake compared to larger molecules.

[0167] For example, a lipid delivery particle can comprise: (a) a lipid membrane; and (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload; wherein the chimeric protein further comprises: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif that promotes assembly of the lipid delivery particle and/or promotes loading of the payload into the lipid delivery particle; (iv) a multimerization motif, or (v) any combination thereof.

[0168] As another example, a lipid delivery particle can comprise: (a) a lipid membrane; and (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload; wherein the chimeric protein further comprises: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif selected from the group consisting of: myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, and farnesylation motif; (iv) a multimerization motif, or (v) any combination thereof.

[0169] As shown in **FIG. 4**, a lipid delivery particle may comprise an envelope protein (e.g., VSV-G), a payload (e.g., base editor ABE8e), and a plasma membrane recruitment element (e.g., PH). The payload and plasma membrane recruitment element may be fused.

ENVELOPE PROTEIN

[0170] In some aspects, the lipid delivery particle provided herein comprises an envelope protein. The envelope protein can be associated with the outside boundary or the surface of the lipid delivery particle, for example, the membrane or envelope of the lipid delivery particle.

[0171] The membrane of the lipid delivery particle can comprise a lipid layer, such as a single layer or a lipid bilayer. In some cases, the membrane of the lipid delivery particle is from plasma membrane, endoplasmic reticulum, or a combination thereof. In some cases, the membrane of the lipid delivery particle is from Golgi complex, ER Golgi intermediate compartment, or nuclear envelope. In some cases, the membrane of the lipid delivery particle is from plasma membrane. In some cases, the membrane of the lipid delivery particle is a phospholipid bilayer.

[0172] The envelope protein can be associated with the membrane of the lipid delivery particle in various manners. For example, the envelope protein can be anchored or attached to the external membrane of the particle or anchored or attached to the internal membrane of the particle. The envelope protein can be embedded or inserted in the membrane, spanning through the membrane, with certain portions located at the outside of the membrane, or certain portions extending to the inside of the particle, or both. The envelope protein within the lipid delivery particle described

herein can be overexpressed from an exogenous source, such as plasmids or stably integrated transgenes, in the production cells.

[0173] The envelope protein can play a role in the delivery of the lipid delivery particle to a target cell and release of the components of the lipid delivery particle within the target cell. The envelope protein can contact with the surface of a target cell and participate in the fusion of the lipid delivery particle and the membrane of the target cell. The envelope protein can participate in the fusion of the lipid delivery particle with the membrane of the target cell via any appropriate mechanism, such as those described in White et al. Crit Rev Biochem Mol Biol. 2008; 43(3): 189–219. One example of the fusion mechanisms is unifying Trimer-of-Hairpins Fusion Mechanism. Membrane fusion can occur after allosteric priming by binding to a target receptor. In some cases, membrane fusion occurs after proteolysis. In some cases, membrane fusion occurs after isomerization of disulfide bridges. In some cases, membrane fusion occurs by internalization and then priming of fusion via (i) cathepsin-mediated proteolysis, or (ii) low pH/acidification. The cathepsin-mediated proteolysis can be pH dependent or pH independent. Other fusion triggering mechanisms can include low PH, binding to target cell receptors, and a receptor followed by low pH. The envelope protein can also play a role in the formation of the lipid delivery particle. The envelope protein can interact with another component within the lipid delivery particle and participate in the assembly of the lipid delivery particle, for example, in a producer cell. The envelope protein can make contact with another envelope protein and form an oligomer embedded within the membrane. The envelope protein can be a glycoprotein, for example, a transmembrane glycoprotein. In some cases, envelope protein comprises multiple membrane-spanning regions. These multiple membrane-spanning regions can oligomerize and form channels in the membrane.

[0174] In some cases, the envelope protein is fused with a targeting moiety. In some cases, the targeting moiety recognizes a specific molecule (*e.g.*, antigen, receptor, or other membrane protein) on the surface of a target cell to allow targeted cell entry with more specificity. In some cases, the targeting moiety is specific for a certain cell type or is specific for a certain target cell. The targeting moiety can be fused to the envelope protein at a position that is located at an outside of the lipid delivery particle. For example, the targeting moiety includes scFvs, antibody variable regions, nanobodies, T-cell receptor variable regions, other antigen-binding fragments or their mimetics, such as DARPins. In some cases, the targeting moiety is a protein ligand from the human ligandome. The targeting moiety can be a natural peptide or a synthetic peptide. In some cases, the targeting moiety is not fused with the envelope protein and is attached to the membrane of the lipid delivery particle from the outside, for example, via a transmembrane domain.

[0175] A targeting moiety can include, *e.g.*, an antibody or an antigen-binding fragment thereof (*e.g.*, Fab, Fab', F(ab')₂, Fv fragments, scFv antibody fragments, disulfide-linked Fvs (sdFv), a Fd

fragment consisting of the VH and CH1 domains, linear antibodies, single domain antibodies such as sdAb (either VL or VH), nanobodies, or camelid VHH domains), an antigen-binding fibronectin type III (Fn3) scaffold such as a fibronectin polypeptide minibody, a ligand, a cytokine, a chemokine, or a T cell receptor (TCRs). Membrane-fusion proteins can be re-targeted by non-covalently conjugating a targeting moiety to the membrane-fusion protein or targeting protein (*e.g.* the hemagglutinin protein). For example, the membrane-fusion protein can be engineered to bind the Fc region of an antibody that targets an antigen on a target cell, redirecting the membrane fusion activity towards cells that display the antibody's target.

[0176] In some cases, the targeting moiety linked to the membrane-fusion protein binds a cell surface marker on the target cell, *e.g.*, a protein, glycoprotein, receptor, cell surface ligand, agonist, lipid, sugar, class I transmembrane protein, class II transmembrane protein, or class III transmembrane protein.

[0177] In some cases, the lipid delivery particles disclosed herein display targeting moieties that are not conjugated to the membrane-fusion protein or other proteins in order to redirect the fusion activity of the lipid delivery particle towards a cell that is bound by the targeting moiety, or to affect tropism of the lipid delivery particle toward the target cell.

Envelope protein of viral origin

[0178] In some cases, an envelope protein has a viral origin. For example, a suitable envelope protein is from a DNA virus, an RNA virus, or a retrovirus. The envelope protein can be envelope protein from Herpesviruses, Avian sarcoma leukosis virus, Poxviruses, Hepadnaviruses, Asfarviridae, Flaviviruses, Alphaviruses, Togaviruses, Coronaviruses, Hepatitis D, Orthomyxoviruses, Rhabdovirus, Bunyaviruses, Filoviruses, Oncoretroviruses, lentiviruses, Spumaviruses. In some cases, envelope protein can be envelope protein from lentiviruses, for example, human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV) and equine infectious anemia virus (EIAV). In some cases, an envelope protein is a fusion of two different envelope proteins, wherein each comes from a different virus. Additional suitable envelope proteins that are from viral origins and their functions are described in White JM *et al.*, Crit Rev Biochem Mol Biol. 2008 May-Jun;43(3):189-219.

[0179] In some cases, the envelope protein is a vesicular stomatitis virus glycoprotein (VSVG) or a biologically active mutant thereof. A “biologically active mutant” disclosed herein in connection with a reference protein can refer to a mutant of the reference protein that remains displaying one or more biological activities that are of same nature as the reference protein, which are relevant to the context in which the reference protein is used in the lipid delivery particle disclosed herein, while the level of the one or more biological activities of the biologically active mutant can be either similar as or different than the reference protein. For instance, the biologically active mutant

of a VSVG in the context of an envelope protein remains displaying the biological activities of an envelope protein, *e.g.*, mediating membrane fusion, tropism of the lipid delivery particle toward a target cell, or both. Unless otherwise noted, a mutant as described in the present disclosure is equivalent to a biologically active mutant. In some cases, the envelope protein is a Human immunodeficiency virus GP160 or a biologically active mutant thereof. In some cases, the envelope protein is a Baboon Endogenous Retrovirus (BaEVTR) glycoprotein or a biologically active mutant thereof. In some cases, the envelope protein is a modified Baboon Endogenous Retrovirus (BaEVTRless) glycoprotein or a biologically active mutant thereof. In some cases, the envelope protein is the fusion protein of Vesicular stomatitis Indiana virus and Rabies virus Glycoproteins (FuG-E) or a biologically active mutant thereof. In some cases, the envelope protein is the pantropic murine leukemia virus envelope protein (MLV) or a biologically active mutant thereof. In some cases, the envelope protein is a modified Fusion protein of Vesicular stomatitis Indiana virus and Rabies virus Glycoproteins (FuG-E P440E) or a biologically active mutant thereof. In some cases, the envelope protein is an ecotropic Murine Leukemia Virus envelope protein (MLV ENV ecotropic) or a biologically active mutant thereof. In some cases, the envelope protein is an amphotropic Murine Leukemia Virus envelope protein (MLV ENV amphotropic) or a biologically active mutant thereof. In some cases, the envelope protein is a Moloney murine leukemia virus envelope protein (MMLV) or a biologically active mutant thereof. In some cases, the envelope protein is a Moloney murine sarcoma virus envelope protein (MoMSVg) or a biologically active mutant thereof. In some cases, the envelope protein is a moloney murine leukemia virus 10A1 strain Glycoprotein (MLV 10A1) or a biologically active mutant thereof. In some cases, the envelope protein is a xenotropic murine leukemia virus envelope protein (MLV ENV xenotropic) or a biologically active mutant thereof. In some cases, the envelope protein is a xenotropic murine leukemia virus-related envelope protein (XMRV) or a biologically active mutant thereof. In some cases, the envelope protein is a Baculovirus envelope glycoprotein (GP64) or a biologically active mutant thereof. In some cases, the envelope protein is an endogenous feline virus envelope protein (RD114ENV) or a biologically active mutant thereof. In some cases, the envelope protein is a mammalian endogenous retrovirus protein, or a biologically active mutant thereof. The mammalian endogenous retrovirus protein can be a koala retrovirus protein (KoRV) or a Jaagsiekte sheep retrovirus protein (enJSRV), or a biologically active mutant thereof.

[0180] In some cases, the envelope protein is a simian endogenous type D retrovirus protein (RD-114) or a biologically active mutant thereof. In some cases, the envelope protein is a gibbon ape leukemia virus envelope protein (GALV) or a biologically active mutant thereof. In some cases, the envelope protein is a feline leukemia virus envelope protein (FLV) or a biologically active

mutant thereof. In some cases, the envelope protein is a mouse mammary tumor virus envelope protein (MMTV) or a biologically active mutant thereof. In some cases, the envelope protein is an avian leukosis virus envelope protein or a biologically active mutant thereof. In some cases, the envelope protein is a rous sarcoma virus envelope protein or a biologically active mutant thereof.

[0181] In some cases, the envelope protein can direct the lipid delivery particles to fuse with a certain type of target cells rather than other cells. For example, based on the specific type of envelope protein associated with the membrane of the lipid delivery particle, the lipid delivery particle can preferentially target different cell types (*i.e.*, tropisms of the lipid delivery particles), such as liver cells, ocular cells, nerve cells, lung cells, immune cells, muscle cells, and any other cell types of interest. For example, to fuse with a target liver cells, the envelope protein can be a glycoprotein from human hepatitis viruses or a biologically active mutant thereof, *e.g.*, Hepatitis B virus (HBV) or hepatitis C virus (HCV), VSV-G glycoprotein or a biologically active mutant thereof, a Marburg virus glycoprotein or a biologically active mutant thereof, an Ebola virus glycoprotein or a biologically active mutant thereof. To fuse with a target muscle cell, for example, a skeletal muscle cell, the envelope protein can be a Ross River virus glycoprotein or a biologically active mutant thereof, or a VSV-G or a biologically active mutant thereof. To fuse with a target ocular cell, for example, a photoreceptor cell or a retinal cell, the envelope protein can be an Ebola virus glycoprotein or a biologically active mutant thereof, a Marburg virus glycoprotein or a biologically active mutant thereof, or a VSV-G or a biologically active mutant thereof. To fuse with a target immune cell, for example, CD8⁺ T cell, an HTLV-1 glycoprotein or a biologically active mutant thereof, or a VSV-G glycoprotein or a biologically active mutant thereof. To fuse with a target immune cell, for example, CD4⁺ T cell, the envelope protein can be a HIV-1 envelope or a biologically active mutant thereof, a HTLV-1 glycoprotein or a biologically active mutant thereof, or a VSV-G glycoprotein or a biologically active mutant thereof. To fuse with a target lung cells, the envelope protein can be a respiratory syncytial virus glycoprotein or a biologically active mutant thereof, or a SARS-CoV glycoprotein or a biologically active mutant thereof. To fuse with a target nerve cell, such as a cell from the central nervous system cell (*e.g.*, neurons, glial cells including oligodendrocytes, astrocytes and microglia), the envelope protein can be a rabies glycoprotein or a biologically active mutant thereof, a Mokola virus glycoprotein or a biologically active mutant thereof, a Semliki Forest virus glycoprotein or a biologically active mutant thereof, a Venezuelan equine encephalitis virus glycoprotein or a biologically active mutant thereof, or a VSV-G or a biologically active mutant thereof. To fuse with a target sensory cell, such as an auditory cell, including hair cells, cochlear cells, etc., the envelope protein can be

an Ebola virus glycoprotein or a biologically active mutant thereof, a Marburg virus glycoprotein or a biologically active mutant thereof, or a VSV-G or a biologically active mutant thereof.

[0182] In some cases, the envelope protein comprises the sequences set forth in **Table 1**. In some cases, the envelope protein comprises the sequences set forth in **Table 1** with at least one amino acid substitution, deletion, or insertion. For instance, N-terminal methionine can be absent from the envelope protein of the lipid delivery particle provided herein relative to the wild-type viral envelope protein. In some cases, the envelope protein comprises the sequences set forth in **Table 1** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0183] In some cases, the envelope protein comprises one or more of the sequences set forth in **Table 1** with at least one amino acid substitution, deletion, or insertion. For instance, N-terminal methionine can be absent from the envelope protein of the lipid delivery particle provided herein relative to the wild-type viral envelope protein. In some cases, the envelope protein comprises one or more of the sequences set forth in **Table 1** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0184] In some cases, the envelope protein comprises any one of the sequences set forth in **Table 1** with at least one amino acid substitution, deletion, or insertion. For instance, N-terminal methionine can be absent from the envelope protein of the lipid delivery particle provided herein relative to the wild-type viral envelope protein. In some cases, the envelope protein comprises any one of the sequences set forth in **Table 1** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0185] In some cases, the envelope protein comprises an amino acid sequence that has at least about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in **Table 1**. In some cases, the envelope protein comprises an amino acid sequence that has at least about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 50% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 60% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 70% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 75% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a

sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 80% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 85% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 90% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 95% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 96% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 97% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 98% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307. In some cases, the envelope protein comprises an amino acid sequence that has at least about 99% sequence identity to a sequence set forth in any one of SEQ ID NOs: 286-307.

Table 1. Exemplary envelope proteins from virus origin

Exemplary envelope proteins from virus origin	Amino Acid Sequence	SEQ ID NO.
vesicular stomatitis virus glycoprotein (VSVG)	MKCLLYLAFLFIGVNCKFTTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPKSHKAIQADGWMCHASKWVTTCDFRWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPQSCGYATVTDAAEAVIVQVTPHHVLVDEYTGWVDSQFINGKCSNYICPTVHNSTTWHS DYKVKGLCDSNLI SMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRLPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISSGTTTERELWDDWAPYEDVEIGPNGVLRRTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDESLFFGDTGLSKNPIELVEGWFFSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCIKLKHTKKRQIYTDIEMNRLGK	286
	KCLLYLAFLFIGVNCKFTTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPKSHKAIQADGWMCHASKWVTTCDFRWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPQSCGYATVTDAAEAVIVQVTPHHVLVDEYTGWVDSQFINGKCSNYICPTVHNSTTWHS DYKVKGLCDSNLI SMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRLPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISSGTTTERELWDDWAPYEDVEIGPNGVLRRTSSGYKFPLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDESLFFGDTGLSKNPIELVEGWFFSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCIKLKHTKKRQIYTDIEMNRLGK	287
Human immunodeficiency	MRVKEKYQH LWRWGWRWGTM L LGM L M I C S A T E K L W V T V Y Y G V P V W K E A T T T L F C A S D A K A Y D T E V H N V W A T H A C V P T D P N P Q E V V L V N V T E N F N M W K N D M V E Q M H E D I I S L W D Q S L K P C V K L T P L C V S L K C T D L K N D T N T N S S G R M I M E K G E I K N	288

ency virus - GP160	CSFNI STSIRGKVQKEYAFFYKLDIIPIDNDTT SYKLTS CNTSVITQACPKV SFEPIPIHYCAPAGFAILKCNNKTFNGTG PCTNVSTVQCTHGIRPVVSTQLL LNGSLAEEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRG PGRAFTIGKIGNMRQAHNCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQS SGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTGSDTITL PCRIKQIINMWQKVGKAMYAPPISGQIRC SSNITGLLLTRDGGNSNNESEIF RPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGALFL GFLGAAGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVWG IKQLQARILAVERYLKDQQLLGIWGC SGKLICTTAVPWNASWSNKSLEQIWN HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNIT NWLWYIKLFI MIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTPRGPD RPEGIEEEGGERDRDRSIRLVNGSLALIWD DLRS LCLFSYHRLRDL LLIVTR IVELLGRRGWEALKYWWNLLQYWSQELKN SAVS LLNATAI AVAEGTDRVIEV VQGACRAIRHIPRRIRQGLERILL	
	RVKEYQHLWRWGWRWGTMLLGMLMICSATEKLWTVYYGVPVWKEATTTLF CASDAKAYDTEVHNWATHACVPTD PNPQEVVLVNVTENFNMWKN DMVEQMH EDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKNC SFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVS FEPIPIHYCAPAGFAILKCNNKTFNGTG PCTNVSTVQCTHGIRPVVSTQLLL NGSLAEEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRG GRAFTIGKIGNMRQAHNCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSS GGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTGSDTITLP CRIKQIINMWQKVGKAMYAPPISGQIRC SSNITGLLLTRDGGNSNNESEIFR PGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGALFLG FLGAAGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVWGI KQLQARILAVERYLKDQQLLGIWGC SGKLICTTAVPWNASWSNKSLEQIWNH TTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITN WLWYIKLFI MIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTPRGPDR PEGIEEEGGERDRDRSIRLVNGSLALIWD DLRS LCLFSYHRLRDL LLIVTRI VELLGRGWEALKYWWNLLQYWSQELKNSAVSLNATAI AVAEGTDRVIEV VQGACRAIRHIPRRIRQGLERILL	289
Baboon Endogenous Retrovirus glycoprotein - BaEVTR	MGFTTKIIFLYNLVLVYAGFDDPRKAIELVQKRYGRPCDCSGGQVSEPPSDR VSQVTC SGKTAYLMPDQRWKCKSIPKDTSPSGPLQECPCNSYQSSVHSSCYT SYQQCRSGNKTYTATLLKTQTGGTSDVQVLGSTNKL IQSPCNGIKGQSI CW STTAPIHVS DGGGPLDTTRIKSVQRKLEE IHKALYPELQYHPLAIPKVRDNL MVDAQTLNINL NATYNLL LMSNTSLVDDCWLC LKLG PPTPLAIPNFLLSYVTR SSDNISCLII PPLLVQPMQFSNSSCLFSPSYNSTEEIDLGHVAFSNCTSITN VTGPICAVNGSVFLCGNNMAYTYLPTNWTGLCVLATLLPDIDII PGDEPVPI PAIDHFIYRPKRAIQFI PLLAGLGITAAFTTGATGLGVSVTQYTKLSNQLIS DVQILSSTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYV NKSGIVRDKIKTLQEELERRRKDLASNPLWTGLQGLLPYLLPFLGPLLLTLLL LLTIGPCI FNRLVQFVKDRISVVQALVLTQQYHQLKPLEYEP	290
	GFTTKIIFLYNLVLVYAGFDDPRKAIELVQKRYGRPCDCSGGQVSEPPSDR SQVTC SGKTAYLMPDQRWKCKSIPKDTSPSGPLQECPCNSYQSSVHSSCYTS YQQCRSGNKTYTATLLKTQTGGTSDVQVLGSTNKL IQSPCNGIKGQSI CWS TTAPIHVS DGGGPLDTTRIKSVQRKLEE IHKALYPELQYHPLAIPKVRDNL VDAQTLNINL NATYNLL LMSNTSLVDDCWLC LKLG PPTPLAIPNFLLSYVTRS SDNISCLII PPLLVQPMQFSNSSCLFSPSYNSTEEIDLGHVAFSNCTSITNV TGPICAVNGSVFLCGNNMAYTYLPTNWTGLCVLATLLPDIDII PGDEPVPI AIDHFIYRPKRAIQFI PLLAGLGITAAFTTGATGLGVSVTQYTKLSNQLISD VQILSSTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYVN KSGIVRDKIKTLQEELERRRKDLASNPLWTGLQGLLPYLLPFLGPLLLTLLL LTIGPCI FNRLVQFVKDRISVVQALVLTQQYHQLKPLEYEP	291
Modified Baboon Endogenous Retrovirus	MGFTTKIIFLYNLVLVYAGFDDPRKAIELVQKRYGRPCDCSGGQVSEPPSDR VSQVTC SGKTAYLMPDQRWKCKSIPKDTSPSGPLQECPCNSYQSSVHSSCYT SYQQCRSGNKTYTATLLKTQTGGTSDVQVLGSTNKL IQSPCNGIKGQSI CW STTAPIHVS DGGGPLDTTRIKSVQRKLEE IHKALYPELQYHPLAIPKVRDNL MVDAQTLNINL NATYNLL LMSNTSLVDDCWLC LKLG PPTPLAIPNFLLSYVTR	292

glycoprotein - BaEVTRless	SSDNI SCLII PPLLVQPMQFSNSSCLFSPSYNSTEEI DLGHVAFSNCTS ITN VTGPICAVNGSVFLCGNNMAYTYLPTNWTGLCVLATLLPDIDII PGDEPVPI PAIDHFIYRPKRAIQFI PLLAGLGITAAFTTGATGLGVSVTQYTKLSNQLIS DVQILSSTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYV NKSGIVRDKIKTLQEELERRRKDLASNPLWTGLQGLLPYLLPFLGPLLTLLL LLTIGPCIFNRLTAFINDKLNIIHAM	
	GFTTKIIIFLYNLVLYAGFDDPRKAIELVQKRYGRPCDCSGGQVSEPPS DRV SQVTC SGKTAYLMPDQRWKCKSIPKDTSPSGPLQECPCNSYQSSVHSSCYTS YQQRSGNKTYTATLLKTQTGGTS DVQVLGSTNKLIQSPCNGIKGQSI CWS TTAPIHVS DGGGLDTRIKSVQRKLEEI HKALYPELQYHPLAIPKVRDNL M VDAQTLN ILNATYNLLLSMNTSLVDDCWLCLKLGPPTPLAI PNFLLSYVTRS SDNISCLII PPLLVQPMQFSNSSCLFSPSYNSTEEIDLGHVAFSNCTSITNV TGPICAVNGSVFLCGNNMAYTYLPTNWTGLCVLATLLPDIDII PGDEPVPI P AIDHFIYRPKRAIQFI PLLAGLGITAAFTTGATGLGVSVTQYTKLSNQLIS D VQILSSTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYV N KSGIVRDKIKTLQEELERRRKDLASNPLWTGLQGLLPYLLPFLGPLLTLLL LLTIGPCIFNRLTAFINDKLNIIHAM	293
Fusion protein of Vesicular stomatitis Indiana virus and Rabies virus Glycoproteins - FuG-E	MVPQVLLFVLLLGFSLCFGKFPITYTIPDELGPWSPIDIHHLSCPNNLVVEDE GCTNLSEFSYMELKVGYSIAIKVNGFTCTGVVTEAETYTNFVG YVTTTFKRK HFRPTPDACRAAYNWKMAGDPRYEE SLHNPYPDYHWRRTVTRTTKESLIIISP SVTDLDPYDKSLHSRVFPGGKCSGITVSSTYCS TNHDYTIWMPENPRPRTPC DIFTNSRGKRASNGNKT CGFVDERGLYKSLKGACRLKLCGVLGLRLMDGTWV AMQTSDETKWCPPDQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIMTT KSVSFRRLSHLRKLVPGFGKAYTIFNKTLMEADAHYKSVRTWNEIIPSKGCL KVGGRCHPHVNGVFFNGIILGPDDHVLIPEMQSSLLQQHMELLES SVIPLMH PLADPSTVFKEGDEAEDFVEVHLPKNPIELVEGW FSSWKSSIASFFFIIGLI IGLFLVLRVGIHLCKIKLKHKKRQIYTDIEMNRLGK	294
	VPQVLLFVLLLGFSLCFGKFPITYTIPDELGPWSPIDIHHLSCPNNLVVEDE GCTNLSEFSYMELKVGYSIAIKVNGFTCTGVVTEAETYTNFVG YVTTTFKRKH FRPTPDACRAAYNWKMAGDPRYEE SLHNPYPDYHWRRTVTRTTKESLIIISP SVTDLDPYDKSLHSRVFPGGKCSGITVSSTYCS TNHDYTIWMPENPRPRTPC DIFTNSRGKRASNGNKT CGFVDERGLYKSLKGACRLKLCGVLGLRLMDGTWVA MQTSDETKWCPPDQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIMTTK SVSFRRLSHLRKLVPGFGKAYTIFNKTLMEADAHYKSVRTWNEIIPSKGCLK VGGGRCHPHVNGVFFNGIILGPDDHVLIPEMQSSLLQQHMELLESSVIPLMHP LADPSTVFKEGDEAEDFVEVHLPKNPIELVEGW FSSWKSSIASFFFIIGLI IGLFLVLRVGIHLCKIKLKHKKRQIYTDIEMNRLGK	295
Modified Fusion protein of Vesicular stomatitis Indiana virus and Rabies virus Glycoproteins - FuG-E (P440E)	MVPQVLLFVLLLGFSLCFGKFPITYTIPDELGPWSPIDIHHLSCPNNLVVEDE GCTNLSEFSYMELKVGYSIAIKVNGFTCTGVVTEAETYTNFVG YVTTTFKRK HFRPTPDACRAAYNWKMAGDPRYEE SLHNPYPDYHWRRTVTRTTKESLIIISP SVTDLDPYDKSLHSRVFPGGKCSGITVSSTYCS TNHDYTIWMPENPRPRTPC DIFTNSRGKRASNGNKT CGFVDERGLYKSLKGACRLKLCGVLGLRLMDGTWV AMQTSDETKWCPPDQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIMTT KSVSFRRLSHLRKLVPGFGKAYTIFNKTLMEADAHYKSVRTWNEIIPSKGCL KVGGRCHPHVNGVFFNGIILGPDDHVLIPEMQSSLLQQHMELLES SVIPLMH PLADPSTVFKEGDEAEDFVEVHLEKNPIELVEGW FSSWKSSIASFFFIIGLI IGLFLVLRVGIHLCKIKLKHKKRQIYTDIEMNRLGK	296
	VPQVLLFVLLLGFSLCFGKFPITYTIPDELGPWSPIDIHHLSCPNNLVVEDE GCTNLSEFSYMELKVGYSIAIKVNGFTCTGVVTEAETYTNFVG YVTTTFKRKH FRPTPDACRAAYNWKMAGDPRYEE SLHNPYPDYHWRRTVTRTTKESLIIISP SVTDLDPYDKSLHSRVFPGGKCSGITVSSTYCS TNHDYTIWMPENPRPRTPC DIFTNSRGKRASNGNKT CGFVDERGLYKSLKGACRLKLCGVLGLRLMDGTWVA MQTSDETKWCPPDQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIMTTK SVSFRRLSHLRKLVPGFGKAYTIFNKTLMEADAHYKSVRTWNEIIPSKGCLK VGGGRCHPHVNGVFFNGIILGPDDHVLIPEMQSSLLQQHMELLESSVIPLMHP LADPSTVFKEGDEAEDFVEVHLEKNPIELVEGW FSSWKSSIASFFFIIGLI IGLFLVLRVGIHLCKIKLKHKKRQIYTDIEMNRLGK	297

Ecotropic Murine Leukemia Virus - MLV ENV (ecotropic)	MARSTLSKPLKNKVNPRGPLIPLILLMLRGVSTASPGSSPHQVYNITWEVTN GDRETVWATSGNHPLWTWWPDLTDLCLMLAHHGPSYWGLEYQSPFSSPPGPP CCSGGSSPGCSRDCHEEPLTSLTPRCNTAWNRLKLDQTTTHKSNEGFYVCPGPH RPRESKSCGGPDSFYCAYWGCETTGRAYWKPPSSSWDFITVNNNLTS DQAVQV CKDNKWCNPLVIRFTDAGRRTSWTTGHHYGLRLYVSGQDPGLTFGIRLRYQ NLGPRVPIGPNPVLADQQPLSKPKPVKSPSVTKPPSGTPLSPTQLPPAGTEN RLNLNLVDGAYQALNLTSPDKTQECWLCLVAGPPYYEGVAVLGTYSNHTSAPA NCSVASQHKLTLS EVTGGQLCIGAVPKTHQALCNTTQTSSRGSYYLVAPTGT MWACSTGLTPCISTTILNLTTDYCVLVELWPRVTYHSPSYVYGLFERSNRHK REPVS LTLALLLGGLTMGGIAAGIGTGTTALMATQQFQQLQAAVQDDLREVE KSISNLEKSLTSLSEVVLQNRRLD LFLKEGGLCAALKEECCFYADHTGLV RDSMAKLRLRERLNQRQKLFESTQGWFEGLFNRS PWFTT LISTIMGPLIVLLMI LLFGPCILNRLVQFVKDRISVVQALVLTQQYHQLKPIEYEP	298
	ARSTLSKPLKNKVNPRGPLIPLILLMLRGVSTASPGSSPHQVYNITWEVTNG DRETVWATS GNHPLWTWWPDLTDLCLMLAHHGPSYWGLEYQSPFSSPPGPPC CSGGSSPGCSRDCHEEPLTSLTPRCNTAWNRLKLDQTTTHKSNEGFYVCPGPHR PRESKSCGGPDSFYCAYWGCETTGRAYWKPPSSSWDFITVNNNLTS DQAVQVC KDNKWCNPLVIRFTDAGRRTSWTTGHHYGLRLYVSGQDPGLTFGIRLRYQN LGPRVPIGPNPVLADQQPLSKPKPVKSPSVTKPPSGTPLSPTQLPPAGTENR LLNLNLVDGAYQALNLTSPDKTQECWLCLVAGPPYYEGVAVLGTYSNHTSAPAN CSVASQHKLTLS EVTGGQLCIGAVPKTHQALCNTTQTSSRGSYYLVAPTGT WACSTGLTPCISTTILNLTTDYCVLVELWPRVTYHSPSYVYGLFERSNRHKKR EPVS LTLALLLGGLTMGGIAAGIGTGTTALMATQQFQQLQAAVQDDLREVEK SISNLEKSLTSLSEVVLQNRRLD LFLKEGGLCAALKEECCFYADHTGLVR DSMAKLRLRERLNQRQKLFESTQGWFEGLFNRS PWFTT LISTIMGPLIVLLMIL LFGPCILNRLVQFVKDRISVVQALVLTQQYHQLKPIEYEP	299
Amphotrophic Murine Leukemia Virus - MLV ENV (amphotropic)	MARSTLSKPPQDKINPWKPLIVMGVLLGVGMAESPHQVFNVNTRVNTNMTGR TANATSL LGTVQDAFPKLYFDLCLDLVGEEDWPSDQEPYVGYGCKYPAGRQRT RTDFYVCPGHTVKSGCGGPGEGYCGKWGCETTQAYWKPTSSWDLISLKRGN NTPWDTGCSKVACGPCYDLSKVSNSFQGATRGGRCNPLVLEFTDAGKKANWD GPKSWGLRLYRTGTDPI TMFSLTRQVLNVGPRVPIGPNPVL PDQRLPSSPIE IVPAPQPPSPLNTSYPPSTTSTPSTSPSPVPQPPPGTGDRLLALVKGAYQ ALNLTNPDKTQECWLCLVSGPPYYEGVAVVGTYTNHSTAPANCTATSQHKLT LSEVTGGQLCMGAVPKTHQALCNTTQSAGSGSYLAAPAGTMWACSTGLTPC LSTTVLNLTTDYCVLVELWPRVIYHSPDYMYGQLEQRTKYKREPVS LTLALL LGGLTMGGIAAGIGTGTTALIKTQQFEQLHAAIQTDLNEVEKSITNLEKSLT SLSEVVLQNRRLD LFLKEGGLCAALKEECCFYADHTGLVRDSMAKLRLR NQRQKLFETGQGWFEGLFNRS PWFTT LISTIMGPLIVLLLI LLFGPCILNRL VQFVKDRISVVQALVLTQQYHQLKPIEYEP	300
	ARSTLSKPPQDKINPWKPLIVMGVLLGVGMAESPHQVFNVNTRVNTNMTGR ANATSL LGTVQDAFPKLYFDLCLDLVGEEDWPSDQEPYVGYGCKYPAGRQRT RTDFYVCPGHTVKSGCGGPGEGYCGKWGCETTQAYWKPTSSWDLISLKRGN TPWDTGCSKVACGPCYDLSKVSNSFQGATRGGRCNPLVLEFTDAGKKANWDG PKSWGLRLYRTGTDPI TMFSLTRQVLNVGPRVPIGPNPVL PDQRLPSSPIE VPAPQPPSPLNTSYPPSTTSTPSTSPSPVPQPPPGTGDRLLALVKGAYQA LNLTNPDKTQECWLCLVSGPPYYEGVAVVGTYTNHSTAPANCTATSQHKLT LSEVTGGQLCMGAVPKTHQALCNTTQSAGSGSYLAAPAGTMWACSTGLTPC LSTTVLNLTTDYCVLVELWPRVIYHSPDYMYGQLEQRTKYKREPVS LTLALL LGGLTMGGIAAGIGTGTTALIKTQQFEQLHAAIQTDLNEVEKSITNLEKSLT LSEVVLQNRRLD LFLKEGGLCAALKEECCFYADHTGLVRDSMAKLRLR NQRQKLFETGQGWFEGLFNRS PWFTT LISTIMGPLIVLLLI LLFGPCILNRL VQFVKDRISVVQALVLTQQYHQLKPIEYEP	301
Moloney murine leukemia virus 10A1 strain	MARSTLSKPLKDKINPWKSLMVMGVLLRVGMAESPHQVFNVNTRVNTNMTGR TANATSL LGTVQDAFPRLYFDLCLDLVGEEDWPSDQEPYVGYGCKYPGGRKRT RTDFYVCPGHTVKSGCGGPGREGYCGEWGCETTQAYWKPTSSWDLISLKRGN NTPWDTGCSKMACGPCYDLSKVSNSFQGATRGGRCNPLVLEFTDAGKKANWD GPKSWGLRLYRTGTDPI TMFSLTRQVLNIGPRIPIGPNPVITGQLPPSRPVQ IRLPRPPQPPTGAASIVPETAPPSQQPGTGDRLLNLVEGAYRALNLTNPDK	302

Glycoprotein - MLV 10A1	TQECWLCLVSGPPYYEGVAVVGTYTNHSTAPASCTAT SQHKLTLSEVTGQGL CMGAVPKTHQALCNTTQSAGSGSYLAAPAGTMWACSTGLTPCLSTTMLNLT TDYCVLVELWPRI IYHSPDYMYGQLEQRTKYKREPVS LTLALLLGGLTMGGI AAGIGTGTTALIKTQQFEQLHAAIQTDLNEVEKSITNLEKSLTSLSEVVLQN RRGLDLLFLKEGGLCAALKEECCFYADHTGLVRDSMAKLRERLNQKQLFES GQGWFEGLFNRSPWFTT LIST IMGPLIVLLLLILFGPCILNRLVQFVKDRIS VVQALVLTQQYHQLKPIEYEP	303
	ARSTLSKPLKDKINPWKSLMVMGVL LRVGMAES PHQVFNVTRVNTNLMTGRT ANATSL LGTVQDAFPRLYFDL CDLVGEEWDPSDQEPYVGYGCKYPGGRKRTR TFDFYVCPGHTVKSGCGGPREGYCGEWGCETTQAYWKPTS SWDLISLKRGN TPWDTGCSKMACGPCYDLSKVSNSFQGATRGRCNPLVLEFTDAGKKNWDG PKSWGRLRLYRTGTDPTMFSLTRQVLNIGPRIP IGPNPVITGQLP PSRPVQI RLPRPQPPPTGAASIVPETAPPSQQPGTGDRLLNLVEGAYRALNLTNPDKT QECWLCLVSGPPYYEGVAVVGTYTNHSTAPASCTATS QHKLTLSEVTGQGLC MGAVPKTHQALCNTTQSAGSGSYLAAPAGTMWACSTGLTPCLSTTMLNLT DYCVLVELWPRI IYHSPDYMYGQLEQRTKYKREPVS LTLALLLGGLTMGGIA AGIGTGTTALIKTQQFEQLHAAIQTDLNEVEKSITNLEKSLTSLSEVVLQNR RGLDLLFLKEGGLCAALKEECCFYADHTGLVRDSMAKLRERLNQKQLFESG QGWFEGLFNRSPWFTT LIST IMGPLIVLLLLILFGPCILNRLVQFVKDRISV VQALVLTQQYHQLKPIEYEP	
Baculovirus envelope glycoprotein - GP64	MVSAIVLYVLLAAAAHSAFAAEHCNAQMKTGPYKIKNLDITPPKETLQKDVE ITIVETDYNENVI IGYKGYQAYAYNGGSLDPNTRVEETMKT LNVGKEDLLM WSIRQQCEVGEELIDRWGSDSDDCFRDNEGRGQWVGKELVKRQNNNHFAHH TCNKS WRCGISTSKMYSRLECQDDTDECQVYILDAEGNPINVTVDTVLHRDG VSMILKQKSTFTTRQIKAACLLIKDDKNNPESVTREHCLIDNDIYDLSKNTW NCKFNRCIKRKVEHRVKKRPPTWRHNVRAYTEGDTATKGDLMHIQEELMYE NDLLKMNIELMHAHINKLNNMLHDLIVSVAKVDERLIGNLMNNSVSTFLSD DTFLMPCTNPPAHTSNCYNNSIYKEGRVVVANTDSSQCIDFSNYKELAI DD DVEFWIPTIGNTTYHDSWKDASGWSFIAQQKSNLITTMENTKFGGVGTSLS ITSMAEGELAAKLTSEFMFGHVNFVILIVILFLYCMIRNRNRQY	304
	VSAIVLYVLLAAAAHSAFAAEHCNAQMKTGPYKIKNLDITPPKETLQKDVEI TIVETDYNENVI IGYKGYQAYAYNGGSLDPNTRVEETMKT LNVGKEDLLMW SIRQQCEVGEELIDRWGSDSDDCFRDNEGRGQWVGKELVKRQNNNHFAHHT CNKSWRCGISTSKMYSRLECQDDTDECQVYILDAEGNPINVTVDTVLHRDGV SMILKQKSTFTTRQIKAACLLIKDDKNNPESVTREHCLIDNDIYDLSKNTWN CKFNRCIKRKVEHRVKKRPPTWRHNVRAYTEGDTATKGDLMHIQEELMYEN DLLKMNIELMHAHINKLNNMLHDLIVSVAKVDERLIGNLMNNSVSTFLSDD TFLMPCTNPPAHTSNCYNNSIYKEGRVVVANTDSSQCIDFSNYKELAI DDD VEFWIPTIGNTTYHDSWKDASGWSFIAQQKSNLITTMENTKFGGVGTSLS ITSMAEGELAAKLTSEFMFGHVNFVILIVILFLYCMIRNRNRQY	
Endogenous feline virus - RD114 ENV	MKLPTGMVILCSLIIVRAGFDDPRKAIALVQKQHGKPCCESSGGQVSEAPNS IQQVTCPGKTAYLMTNQKWKCRVTPKISP SGGELQNCPCNTFQDSMHSSCYT EYRQCRINKTYYTATLLKIRSGSLNEVQILQNPQLLQSPCRGSINQPV CW SATAPIHISDGGGPLDTKRVVTVQKRLEQIHKAMTPELQYHPLALPKVRDD LSLDARTFDILNTTFRLLQMSNFSLAQDCWLCLKLGTPTPLAIP TPSLTYSL ADSLANASCQIIPPLLVQPMQFSNS SCLSSPFINDTEQIDLGAVTFTNCTSV ANVSSPLCALNGSVFLCGNNMAYTYLPQNVVTRL CVQASLLPDIDINPGDEP VPIPAIDHYIHRPKRAVQFIPLLAGLGITAAFTTGATGLGVSVTQYTKLSHQ LISDVQVLSGTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKC FYANKSGIVRNKIRTLQEELQKRRESLATNPLVVTGLQGFLPYLLPLLGPLL TLLILITIGPCVFSRLMAFINDRLNVVHAMVLAQQYQALKAEEEEAQD	306
	KLPTGMVILCSLIIVRAGFDDPRKAIALVQKQHGKPCCESSGGQVSEAPPNSI QQVTC PGKTAYLMTNQKWKCRVTPKISP SGGELQNCPCNTFQDSMHSSCYTE YRQCRINKTYYTATLLKIRSGSLNEVQILQNPQLLQSPCRGSINQPV CWS ATAPIHISDGGGPLDTKRVVTVQKRLEQIHKAMTPELQYHPLALPKVRDDL SLDARTFDILNTTFRLLQMSNFSLAQDCWLCLKLGTPTPLAIP TPSLTYSLA DSLANASCQIIPPLLVQPMQFSNS SCLSSPFINDTEQIDLGAVTFTNCTSV ANVSSPLCALNGSVFLCGNNMAYTYLPQNVVTRL CVQASLLPDIDINPGDEPV	

	PIPAI DHI IHRPKRAVQFIPLLAGLGITAAFTTGATGLGVSVTQYTKLSHQL ISDVQVLSGTIQDLQDQVDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCF YANKSGIVRNKIRTLQEELQKRRESLATNPLVVTGLQGFLPYLLPLLGPLLT LLLILTIGPCVFSRLMAFINDRLNVVHAMVLAQQYQALKAEFEAQD	
--	---	--

Envelope protein of human origin

[0186] In some aspects, the envelope protein in the lipid delivery particle described herein has a human origin, *e.g.*, has significant sequence similarity to a human wild-type protein, such as at least 90%, at least 95%, at least 98%, or at least 99%. Using an envelope protein of a human origin can have benefits such as providing a minimized immunogenicity and better tolerance in a human subject receiving the lipid delivery particles. The lipid delivery particle comprising an envelope protein of a human origin can comprise another component that is from human origin or from non-human origin (*e.g.*, a payload or a plasma membrane recruitment element). An envelope protein that is from human origin can include, example, envelope proteins or glycoproteins of human endogenous retroviruses (HERVs), other human endogenous envelope proteins, or other human endogenous proteins that serve a similar function of recognizing and/or fusing with membrane of a target cell (*e.g.*, clathrin adaptor protein complex-1, CHMP4C, Proteolipid protein 1, TSAP6, immunoglobulin variable domains, or a mutant thereof).

[0187] In some cases, the envelope protein is a HERV envelope protein such as any one of those listed in **Table 2**. In some cases, the envelope protein is a hENVH1 or a mutant thereof. In some cases, the envelope protein is a hENVH2 or a mutant thereof. In some cases, the envelope protein is a hENVH3 or a mutant thereof. In some cases, the envelope protein is a hENVK1 or a mutant thereof. In some cases, the envelope protein is a hENVK2 or a mutant thereof. In some cases, the envelope protein is a hENVK3 or a mutant thereof. In some cases, the envelope protein is a hENVK4 or a mutant thereof. In some cases, the envelope protein is a hENVK5 or a mutant thereof. In some cases, the envelope protein is a hENVK6 or a mutant thereof. In some cases, the envelope protein is a hENVV or a mutant thereof. In some cases, the envelope protein is a hENVFRD or a mutant thereof. In some cases, the envelope protein is a hENVR or a mutant thereof. In some cases, the envelope protein is a hENVR(b) or a mutant thereof. In some cases, the envelope protein is a hENVR(c)2 or a mutant thereof. In some cases, the envelope protein is a hENVR(c)1 or a mutant thereof. In some cases, the envelope protein is a hENVKcon or a mutant thereof. In some cases, the envelope protein is a truncated HERV protein.

Table 2. Exemplary human HERV envelope proteins

#	HERV envelope	Gene name	Accession no.	Position in sequence entry (a)
1.	hENVH1	envH/p62	AJ289709.1	6313-8067(+)

2.	hENVH2	envH/p60	AJ289710.2	5393-7084(+)
3.	hENVH3	envH/p59	AJ289711.1	5204-6871(+)
4.	hENVK1	envK1	AC074261.3	93508-95604(+)
5.	hENVK2	envK2/HML-2.HOM	AC072054.10	30365-32464(-)
6.	hENVK3	envK3/C19	Y17833.1	5581-7680(+)
7.	hENVK4	envK4/K109	AF164615.1	6412-8508(+)
8.	hENVK5	envK5/K113	AY037928.1	6451-8550(+)
9.	hENVK6	envK6/K115	AY037929.1	6442-8541(+)
10.	hENVT	envT	AC078899.1	154738-156618(+)
11.	hENVW	Syncytin-1	AC000064.1	35879-37495(+)
12.	hENVFRD	Syncytin-2	AL136139.6	21355-22972(-)
13.	hENVR	erv-3	AC073210.8	54963-56978(-)
14.	hENVR(b)	envRb	AC093488.1	78681-80225(+)
15.	hENVF(c)2	envFc2	AC016222.4	85216-86963(+)
16.	hENVF(c)1	envFc1	AL354685.2	46744-48717(-)
*17.	hENVK _{con}	N/A	N/A	N/A

(a) '+' and '-' refer to the orientation within the sequence entry

*hENVK_{con} is a consensus sequence from ten proviral ENV sequences. The ENV sequences used to derive this consensus ENV sequence are from the following HERVs: HERV-K113, HERV-K101, HERV-K102, HERV-K104, HERV-K107, HERV-K108, HERV-K109, HERV-K115, HERV-K11p22, and HERV-K12q13.

[0188] In some cases, the envelope protein comprises the sequences set forth in **Table 3**. In some cases, the envelope protein comprises the sequences set forth in **Table 3** with at least one amino acid substitution, deletion, or insertion. For example, for those amino acid sequences start with a N-terminal methionine, the N-terminal methionine can be absent. In some cases, the envelope protein comprises the sequences set forth in **Table 3** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0189] In some cases, the envelope protein comprises one or more of the sequences set forth in **Table 3** with at least one amino acid substitution, deletion, or insertion. For instance, N-terminal methionine can be absent from the envelope protein of the lipid delivery particle provided herein relative to the wild-type viral envelope protein. In some cases, the envelope protein comprises one or more of the sequences set forth in **Table 3** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0190] In some cases, the envelope protein comprises any one of the sequences set forth in **Table 3** with at least one amino acid substitution, deletion, or insertion. For instance, N-terminal methionine can be absent from the envelope protein of the lipid delivery particle provided herein relative to the wild-type viral envelope protein. In some cases, the envelope protein comprises any one of the sequences set forth in **Table 3** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0191] In some cases, the envelope protein comprises an amino acid sequence that has at least about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 50% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 60% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 70% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 75% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 80% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 85% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 90% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 95% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 96% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 97% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 98% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341. In some cases, the envelope protein comprises an amino acid sequence that has at least about 99% sequence identity to a sequence set forth in any one of SEQ ID NOs: 308-341.

[0192]

Table 3. Exemplary sequences for human HERV envelope proteins

HERV envelope	Amino Acid Sequence	SEQ ID NO:
hENVT	MGPEAWVRPLKTA PKPGEAIRLILFIYLS CFFLPVMS SEPSYSFLLTSFTT GRVFANTTWRAGT SKEVSFAVDLCVLFPE PARTHEEQHNLPVIGAGSVDLA AGFGHSGSQTGCGSSKGA EKGLQNVDFYLC PGNHDPDASCRDTYQFFCPDWT CVTLATYSGGSTRSSTLSISRVPHPKLCTRKNCNPLTITVHDPNAAQWY YG	308

	MSWGLRLYI PGFDVGTMTFTIQKKILVSWSSPKPIGPLTDLGDPFIQKHPDK VDLTVPLPFLVPRPQLQQOHLQPSLMSILGGVHHLLNLTQPKLAQDCWLC KAKPPYYVGLGVEATLKRGP LSCHTRPRALTIGDVSGNASCLISTGYNL SA SPFOATCNQSLT SISTSVSYQAPNNTWLACTSGLTRCINGTEPGPLLCVL VHVLPQVYVYSGPEGRQLIAPPELHPR LHQAVPLLVPLLAGLSIAGSAAIG TAALVQGETGLISLSQQVDADFSNLQSAIDILHSQVESLAEVVLQNCRC LD LFLSQQGLCAALGESCCFYANQSGVIKGT VKKVRENLD RHQQERENNI PW YQSMFNWNPWLTT LITGLAGPL LILLLSLIFGPCILNSFLNFIKQRIASVK LTYLKTQYDTLVNN	
	GPEAWVRPLKTAPKPGEAIRLILFIYLSCFFLPVMSSEPSYSFLLTSFTTG RVFANTTW RAGTS KEVS FAVDLCVLFPEPARTHEEQHNLPVIGAGSVDLAA GFGHSGSQT GCGS SKGA EKGLQNVDFYLC PGNH PDAS CRDTYQFFCPDWTC VTLATYSGGSTRS STLS ISRVPHPKLCTRKN CNPLTI TVHDPNAAQWYYGM SWGLRLYI PGFDVGTMTFTIQKKILVSWSSPKPIGPLTDLGDPFIQKHPDKV DLTVPLPFLVPRPQLQQOHLQPSLMSILGGVHHLLNLTQPKLAQDCWLC LK AKPPYYVGLGVEATLKRGP LSCHTRPRALTIGDVSGNASCLISTGYNL SAS PFQATCNQSLT SISTSVSYQAPNNTWLACTSGLTRCINGTEPGPLLCVL LV HVLPQVYVYSGPEGRQLIAPPELHPR LHQAVPLLVPLLAGLSIAGSAAIGT AALVQGETGLISLSQQVDADFSNLQSAIDILHSQVESLAEVVLQNCRC LDL LFLSQQGLCAALGESCCFYANQSGVIKGT VKKVRENLD RHQQERENNI PWY QSMFNWNPWLTT LITGLAGPL LILLLSLIFGPCILNSFLNFIKQRIASVK LTYLKTQYDTLVNN	309
hENVW	MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEF LWRMQRPGNIDAPS YRSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV TVCWTYFTQTGMS DGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSK LHETLRTHTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSI PVPE QWNNFSTEINTTSVLVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWV TPPTQIVCLPSGIF FVCGTSA YRCLNGSS ESMCFLSFLVPPMTIYTEQDLY SYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTTSTQFYKLSQELNGD MERVADSLVTLQDQLNSLA AVVLQNRALDLLTAERGGTCLFLGEECCY YV NQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILPFLGPLAAII L LLLFGPCI FNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASPRS DVNDIKGTPPEEISAAQPLLRPNSAGSS;	310
	ALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEF LWRMQRPGNIDAPSY RSLSKGTPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV TVCWTYFTQTGMS DGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSK LHETLRTHTRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSI PVPEQ WNNFSTEINTTSVLVGPLVSNLEITHTSNLTCVKFSNTTYTTNSQCIRWV TPPTQIVCLPSGIF FVCGTSA YRCLNGSS ESMCFLSFLVPPMTIYTEQDLY SYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTTSTQFYKLSQELNGD MERVADSLVTLQDQLNSLA AVVLQNRALDLLTAERGGTCLFLGEECCY YV NQSGIVTEKVKEIRDRIQRRAEELRNTGPWGLLSQWMPWILPFLGPLAAII L LLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASPRSD VNDIKGTPPEEISAAQPLLRPNSAGSS	311
hENVH1	MI FAGKAPSNTSTLMKFYSLLLYSLLFSFPFLCHPLPLPSYLHHTINLTHS LLAASNPSLVNNCWLCLISLSSAYTAVPAVQTDWATSPISLHLRTSFNSPH LYPPEELIYFLDRSSKTS PDISHQQAALLRTYLKNLSPYINSTPPIFGPL TTQTTIPVAAPLCISWQRPTGIP LGNLSPSRCSFTLHLRSPTTNINETIGA FQLHITDKPSINTDKLKNISSNYCLGRHLPCISLHPWLSSPCSSDSPRPS SCLLI PS PENNSERLVDTRRFLIHENRTFPSTQLPHQSP LQPLTAAALA GSLGVWVQDTPFSTPSHLFTLHLQFCLAQGLFFLCGSTYMCLPANWTGTC TLVFLTPKI QFANGTEELPVP LMTPTQQKRVIPLIPLMVGLGLSASTVALG TGIAGISTSVMTFRSLSNDFSASITDISQTL SVLQAQVDSLAAVVLQNR RG LDLLTA EKGLCI FLNEECCFYLNQSGLVYDNIKKLKDRAQKLANQASNYA EPPWALS NWSWVLP IVSPLIPIF LLLLFGPCI FRLVSQFIQNR IQAITNH SIRQMFLLTSPQYHPLPQDLPSA;	312
	IFAGKAPSNTSTLMKFYSLLLYSLLFSFPFLCHPLPLPSYLHHTINLTHSL LLAASNPSLVNNCWLCLISLSSAYTAVPAVQTDWATSPISLHLRTSFNSPHL	313

	YPPEELIYFLDRS SKTS PDISHQQAALLRTYLKNLS PYINSTPPIFGPLT TQTTI PVAAPLCI SWQRPTGI PLGNLSPSRCSFTLHLRSPTTNINETIGAF QLHITDKPS INTDKLKNISSNYCLGRHLP CISLHPWLSPPCSDSPRPSS CLLIPSPENNSERLLVDTRRFLIHHENRT FPSTQLPHQSPLQPLTAAALAG SLGVWVQDTPFSTPSHLFTLHLQFC LAQGLFFLCGSSTYMCLPANWTGTCT LVFLT PKIQFANGTEEL PVPLMTPTQQRVIPLIPLMVGLGLSASTVALGT GIAGISTSVMTFRSLNDFSASITDISQTL SVLQAQVDSLAAVVLQNRRL DLLTAEKGGLCIFLNEECCFYLNQSGLVYDNIKKLKDRAQKLANQASNYAE PPWALS NWMSWVLP IVS PLIP IFLLLLFGPCIFRLVSQFIQNRIQAITNHS IRQMFLLTS PQYHPLPQDLPSA	
hENVH2	MIFAGRASSNTSTLMKFYSLLLYSLLFSFPILCHPLPLPSYLHHTINLTHS LLAVSNPSLAKNCWLCISLPSSAYPAVPALQTDWGTS PVSPHLRTSFNSPH LYPPEKLIYFLDRSSKTS PDISHQQAALLCTYLKNLS PYINSTPPTFGPL TTQTTI PVAAPLCISRQRPTGI PLGNLSPSRCSFTLHLRSPTTHITETNGA FQLHITDKPS INTDKLKNVSSNYCLGRHLS CISLHPWLFSPCSDSPRPSS SCLLI PSKNNSESLVDAQRF LIYHENRTSPSTQLPHQSPLQPLTAAPLG GSLRVWVQDTPFSTPSHLFTLHLQFCLVQSLFFLCGSSTYMCLPANWTGTCT TLVFLT SKIQFANGTEELPVPLMTPTRQKRVIPLIPLMVGLGLSASTVALG TGIAGISTSVTTFRILSNDFSASITDISQTL SGLQAQVDSSAAVVLQNRQG LDLLTAEKGGLCIFLNEESYFYLNQSGLVYDNIKKLKDKAQNLANQASNYA EPPWPLSNWMSWVLPILSPLIP IFLLLLFFRPCIFHLVSQFIQNHIQAITDHS SI	314
	IFAGRASSNTSTLMKFYSLLLYSLLFSFPILCHPLPLPSYLHHTINLTHSL LAVSNPSLAKNCWLCISLPSSAYPAVPALQTDWGTS PVSPHLRTSFNSPHL YPPEKLIYFLDRS SKTS PDISHQQAALLCTYLKNLS PYINSTPPTFGPLT TQTTI PVAAPLCISRQRPTGI PLGNLSPSRCSFTLHLRSPTTHITETNGAF QLHITDKPS INTDKLKNVSSNYCLGRHLS CISLHPWLFSPCSDSPRPSS CLLIPSPKNNSESLVDAQRF LIYHENRTSPSTQLPHQSPLQPLTAAPLGG SLRVWVQDTPFSTPSHLFTLHLQFCLVQSLFFLCGSSTYMCLPANWTGTCT LVFLT SKIQFANGTEELPVPLMTPTRQKRVIPLIPLMVGLGLSASTVALGT GIAGISTSVTTFRILSNDFSASITDISQTL SGLQAQVDSSAAVVLQNRQGL DLLTAEKGGLCIFLNEESYFYLNQSGLVYDNIKKLKDKAQNLANQASNYAE PPWPLSNWMSWVLPILSPLIP IFLLLLFFRPCIFHLVSQFIQNHIQAITDHS I	315
hENVH3	MILAGRAPSNTSTLMKFYSLLLYSLLFSFPFLYHPLPLPSYLHHTINLTHS LPAASNPSLANNCWLCISLSSSAYIAVPTLQTDRTSPVSLHLRTSFNSPH LYPPEELIYFLDRSSKTS PDISHQPAALLHIY LKNLS PYINSTPPIFGPL TTQTTI PVAAPLCISRQRPTGI PLGNISPSRCSFTLHLQSPTTHVTETIGV FQLHIDKPS INTDKLKNVSSNYCLGRHLPYISLHPWLPSPCSDSPRPSS SCLLT PSQNNSERLLVDTRQRF LIHHENRTSSSMQLAHQSPLQPLTAAALA GSLGVWVQDTPFSTPSHPFSLHLQFCLTQGLFFLCGSSTYMCLPANWTGTCT TLVFLT PKIQFANGTKELPVPLMTLTTPQKRVIPLIPLMVGLGLSASTIALS TGIAGISTSVTTFRSPSNDFSASITDISQTL SVLQAQVDSLAAVVLQNRRL LGLSILLNEECCFYLNQSGLVYENIKKLKDRAQKLANQASNYAES PWALS NWMSWVLPILSPLIP IFLLLLFGPCIFHLVSQFIQNRIQAITNHSI ;	316
	ILAGRAPSNTSTLMKFYSLLLYSLLFSFPFLYHPLPLPSYLHHTINLTHSL PAASNPSLANNCWLCISLSSSAYIAVPTLQTDRTSPVSLHLRTSFNSPHL YPPEELIYFLDRS SKTS PDISHQPAALLHIY LKNLS PYINSTPPIFGPLT TQTTI PVAAPLCISRQRPTGI PLGNISPSRCSFTLHLQSPTTHVTETIGVF QLHIIDKPS INTDKLKNVSSNYCLGRHLPYISLHPWLPSPCSDSPRPSS CLLT PSQNNSERLLVDTRQRF LIHHENRTSSSMQLAHQSPLQPLTAAALAG SLGVWVQDTPFSTPSHPFSLHLQFCLTQGLFFLCGSSTYMCLPANWTGTCT LVFLT PKIQFANGTKELPVPLMTLTTPQKRVIPLIPLMVGLGLSASTIALST GIAGISTSVTTFRSPSNDFSASITDISQTL SVLQAQVDSLAAVVLQNRRL GLSILLNEECCFYLNQSGLVYENIKKLKDRAQKLANQASNYAES PWALS NWMSWVLPILSPLIP IFLLLLFGPCIFHLVSQFIQNRIQAITNHSI	317
hENVK1	MHPSEMQRKAPRRRRHRNRAPLTHKMNMVTS EQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTPESMLLAALMIVSMVVS LPMPAGAAAANYTN	318

	WAYVPFPPLIRAVTWMDNPPIEVYVNDVWVHGPIDDRCPAKPEEEGMMINI SIGYHYPPICLGRAPGCLMPAVQNWLVPTVSPISRFTYNMVSGMSLRPR VNYLQDFSYQRS LKFRPKGKPCPKEIPKESKNTEVLVWEECVANSVVILQN NEFGTIIIDWAPRGQFYHNCSGQTQSCPSAQVSPA VDSDLTESLDKHKHKKL QSFY PWEWGEKGI STPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCR LLTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSIHLTEVLKGV L NRSKRFIFTLIAVIMGLIAVTAMAAVAGVALHSFVQSVNFVNDWQKNSTR L WNSQSSIDQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITP QIYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTEA IAGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRFTQQLRRD SYHRE RAMMTMVVLSKRKGGNVGKSKRDQIVTVSV;	
	HPSEMQRKAPPRRRRHRNRAPLTHKMNMVMTSEQMKL PSTKKAEPPTWAQ L KKLTQLATKYLENTKVTQTPE SMLLAALMIVSMVSLPMPAGAAAANYTNW AYVPFPPLIRAVTWMDNPPIEVYVNDVWVHGPIDDRCPAKPEEEGMMINIS IGYHYPPICLGRAPGCLMPAVQNWLVPTVSPISRFTYNMVSGMSLRPRV NLYLQDFSYQRS LKFRPKGKPCPKEIPKESKNTEVLVWEECVANSVVILQNN EFGTIIIDWAPRGQFYHNCSGQTQSCPSAQVSPA VDSDLTESLDKHKHKKLQ SFY PWEWGEKGI STPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL E TRDRKPFYTVDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCR L LTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSIHLTEVLKGV LN RSKRFIFTLIAVIMGLIAVTAMAAVAGVALHSFVQSVNFVNDWQKNSTR L W NSQSSIDQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQ IYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTEAI AGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRFTQQLRRDS YHRE RAMMTMVVLSKRKGGNVGKSKRDQIVTVSV	319
hENVK2	MNPSEMQRKAPPRRRRHRNRAPLTHKMNMVMTSEEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQTPE SMLLAALMIVSMVSLPMPAGAAAANYT YWAYVPFPPLIRAVTWMDNPTEVYVNDVWVPGPIDDRCPAKPEEEGMMIN ISIGYHYPPICLGRAPGCLMPAVQNWLVPTVSPICRFTYH MVSGMSLRP RVNYLQDFSYQRS LKFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQ NNEFGTIIIDWAPRGQFYHNCSGQTQSCPSAQVSPA VDSDLTESLDKHKHKK LQSFY PWEWGEKGI STPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL LETRDRKPFYTVDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENC RLLTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCIT PQIYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTE AIAGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRCTQQLRR DSDHRE RAMMTMAVLSKRKGGNVGKSKRDQIVTVSV;	320
	NPSEMQRKAPPRRRRHRNRAPLTHKMNMVMTSEEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQTPE SMLLAALMIVSMVSLPMPAGAAAANYTY WAYVPFPPLIRAVTWMDNPTEVYVNDVWVPGPIDDRCPAKPEEEGMMINI SIGYHYPPICLGRAPGCLMPAVQNWLVPTVSPICRFTYH MVSGMSLRP VNYLQDFSYQRS LKFRPKGKPCPKEIPKESKNTEVLVWEECVANSVILQN NEFGTIIIDWAPRGQFYHNCSGQTQSCPSAQVSPA VDSDLTESLDKHKHKKL QSFY PWEWGEKGI STPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCR LLTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV L NRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSVQSVNFVNDWQKNSTR L WNSQSSIDQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITP QIYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTEA IAGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRCTQQLRRD SDHRE RAMMTMAVLSKRKGGNVGKSKRDQIVTVSV	321
hENVK3	MNPSEMQRKAPPRRRRHRNRAPLTHKMNMVMTSEEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQTPE SMLLAALMIVSMVSLPMPAGAAAANYT YWAYVPFPPLIRAVTWMDNPPIEVYVNDVWVPGPTDDHCPAKPEEEGMMIN ISIGYHYPPICLGRAPGCLMPAVQNWLVPTVSPISRFTYH MVSGMSLRP	322

	RVNYLQDFS YQRS FKFRPKGKPCPKEI PKESKNTEVLVWEECVANS AVILQ NNEFGTII DWAPRGQFYHNCSGQTQSCPSAQVS PAVDSDLTESLDKHKHKK LQSFY PWEWGEKGISTPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQT LETRDRKPFYTVDLNSSVTVPLQSC IKPPYMLVVGNI VIKPDSQITCENC RL LTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWETSPSIHTL TEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFSIT PQIYNESEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTE AIAGVADGLANLNPVTWVKTI GSTTIIINL ILILVCLFCLLLVCRCTQQLRR DSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV;	
	NPSEMQRKAPPRRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAAANYTY WAYVPFPPLIRAVTWMDNPIEVYVND SVWVPGPTDDHCPAKPEEEGMMINI SIGYRYPICLGRAPGCLMPAVQNWLV E VPTVSPISRFTYH MVSGMSLRPR VNYLQDFS YQRSFKFRPKGKPCPKEI PKESKNTEVLVWEECVANS AVILQ NEFGTII DWAPRGQFYHNCSGQTQSCPSAQVS PAVDSDLTESLDKHKHKKL QSFY PWEWGEKGISTPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSVTVPLQSC IKPPYMLVVGNI VIKPDSQITCENC RL LTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWETSPSIHTL TEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHS SVQSVNFVNDWQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFSITP QIYNESEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTEA IAGVADGLANLNPVTWVKTI GSTTIIINL ILILVCLFCLLLVCRCTQQLRRD SDHREERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV	323
hENVK4	MNPSEMQRKAPPRRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAAANYT NWAYVPFPPLIRAVTWMDNPIEVYVND SVWVPGPIDDRCPAKPEEEGMMIN ISIGYRYPICLGRAPGCLMPAVQNWLV E VPIVSPICRFTYH MVSGMSLRPR VNYLQDFS YQSLKFRPKGKPCPKEI PKESKNTEVLVWEECVANS AVILQ NEFGTII DWTPQGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKL QSFY PWEWGEKGISTPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSLTLP LQSCVKPPYMLVVGNI VIKPDSQITCENC RL LTCIDSTFNWQHRILLVRAREGVWIPVSM DRPW EASPSIHL TEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHS SVQSVNFVNDGQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCITP QIYNESEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTEA IAGVADGLANLNPVTWVKTI GSTTIIINL ILILVCLFCLLLVCRCTQQLRRD SDHREERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV;	324
	NPSEMQRKAPPRRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAAANYTN WAYVPFPPLIRAVTWMDNPIEVYVND SVWVPGPIDDRCPAKPEEEGMMINI SIGYRYPICLGRAPGCLMPAVQNWLV E VPIVSPICRFTYH MVSGMSLRPRV NYLQDFS YQSLKFRPKGKPCPKEI PKESKNTEVLVWEECVANS AVILQ NNEFGTII DWTPQGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQ SFY PWEWGEKGISTPRPKIISPVS GPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSLTLP LQSCVKPPYMLVVGNI VIKPDSQITCENC RL LTCIDSTFNWQHRILLVRAREGVWIPVSM DRPW EASPSIHL TEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDGQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCITPQ IYNESEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTEAI AGVADGLANLNPVTWVKTI GSTTIIINL ILILVCLFCLLLVCRCTQQLRRDS DHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV	325
hENVK5	MNPSEMQRKAPPRRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAAANYT YWAYVPFPPLIRAVTWMDNPIEIYVND SVWVPGPTDDCCPAKPEEEGMMIN ISIGYRYPICLGRAPGCLMPAVQNWLV E VPTVSPISRFTYH MVSGMSLRP RVNYLQDFS YQRS LKFRPKGKPCPKEI PKESKNTEVLVWEECVANS AVILQ NNEFGTLIDWAPRGQFYHNCSGQTQSCPSAQVS PAVDSDLTESLDKHKHKK	326

	LQSFY PWEWGEKG ISTARPKI I SPVSGPEHPELWRLTVASHHIRI WSGNQTL LETRDRKPFY TIDLNSS LTVPLQSCVKPP YMLVVGNI VIKPDSQT ITCENC RL LTCIDST FNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQNNSTR LWNSQSS IDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCIT PQIYNESEHHWDMVRCHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTE AIAGVADGLANLNTVTWVKTI GSTT IINL ILILVCLFCLLLVYRCTQQLRR DSDHRERAMMTMVVLSKRKGGNVGKSKRDQIVTVSV;	
	NPSEMQRKAPPRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAAANYTY WAYVPFPPLIRAVTWMDNPIEIVVND SVWVPGPTDDC CPAKPEEEGMMINI SIGRYPPICLGRAPGCLMPAVQNWLV E VPTVSPISRFTYH MVSGMSLRPR VNYLQDFS YQSLKFRPKGKPCPKEIPKESKNTEVLVWEECVANS AVILQN NEFGTLIDWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKL QSFY PWEWGEKG ISTARPKI I SPVSGPEHPELWRLTVASHHIRI WSGNQTL ETRDRKPFY TIDLNSS LTVPLQSCVKPP YMLVVGNI VIKPDSQT ITCENC RL LTCIDST FNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQNNSTR LWNSQSS IDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCIT PQIYNESEHHWDMVRCHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTE AIAGVADGLANLNTVTWVKTI GSTT IINL ILILVCLFCLLLVYRCTQQLRR SDHREAMMTMVVLSKRKGGNVGKSKRDQIVTVSV	327
hENVK6	MNPSEMQRKAPPRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAVANYT NWAYVPFPPLIRAVTWMDNPIEIVVND SVWVPGPIDDRCPAKPEEEGMMIN ISIGRYPPICLGRAPGCLMPAVQNWLV E VPTVSPISRFTYH MVSGMSLRP RVNYLQDFS YQRS LKFRPKGKPCPKEIPKESKNTEVLVWEECVANS AVILQ NNEFGTII DWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKK LQSFY PWEWGEKRISTPRPKIVSPVSGPEHPELWRLTVASHHIRI WSGNQTL LETRDRKPFY TVDLNSS LTLPLQSCVKPP YMLVVGNI VIKPDSQT ITCENC RL LTCIDST FNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDGQKNSTR LWNSQSS IDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCIT PQIYNDSEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTE AIAGVADGLANLNPVTWVKTI GSTT IINL ILILVCLFCLLLVCRCTQQLRR DSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV;	328
	NPSEMQRKAPPRRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVVS LPMPAGAAVANYTN WAYVPFPPLIRAVTWMDNPIEIVVND SVWVPGPIDDRCPAKPEEEGMMINI SIGRYPPICLGRAPGCLMPAVQNWLV E VPTVSPISRFTYH MVSGMSLRPR VNYLQDFS YQSLKFRPKGKPCPKEIPKESKNTEVLVWEECVANS AVILQN NEFGTII DWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKL QSFY PWEWGEKRISTPRPKIVSPVSGPEHPELWRLTVASHHIRI WSGNQTL ETRDRKPFY TVDLNSS LTLPLQSCVKPP YMLVVGNI VIKPDSQT ITCENC RL LTCIDST FNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFI FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDGQKNSTR LWNSQSS IDQKLANQINDLRQTVIWMGDR LMSLEHRFQLQCDWNTSDFCIT PQIYNDSEHHWDMVRRHLQGREDNLTLDI SKLKEQIFEASKAHLNLVPGTE AIAGVADGLANLNPVTWVKTI GSTT IINL ILILVCLFCLLLVCRCTQQLRR SDHREAMMTMAVLSKRKGGNVGKSKRDQIVTVSV	329
hENVFRD	MGLLLLVLI LTPSLAAYRHPDFLLEKAQQLLQSTGSPYSTNCWLCTSSST ETPGTAYPASPREWTSIEAELHISYRWDPNLKG LMRPANSLLSTVKQDFPD IRQKPPIFGPIFTNINLMGIA PICVM AKRKNGTNVGTL PSTVCNVTFTVDS NQQT YQTYTHNQFRHQPRFPKPPNITFPQGTLLDKSSRFCQGRPS SCSTRN FWFRPADYNQCLQISNLSSTA EWVLLDQTRNSLFWENKTGANQSQT PCVQ VLAGMTIATSYLGISAVSEFFGTS LTPLFHFHISTCLKTQGA FYICQS IH QCLPSNWTGTCTIGYVTPDIF IAPGNLSLPIPIYGNS PLPRVRAIHFIPL LAGLGILAGTGTG IAGITKASLTYSQLSKEIANNIDTMAKALTTMQEQIDS	330

	LAAVVLQNRRLDMLTAAQGGICLALDEKCCFWVNQSGKVQDNIRQLLNQAS SSLRERATQGWLNWEGTWKWF SWVL PLTG PLVSL LLLLLLFGPCLLNLTITQF VSSRLQAIKLQTNLSAGRHPN IQES PF;	
	GLLLLLVLILTPSLAAYRHPDFPLEKAQQLLQSTGSPYSTNCWLCTSSSTE TPGTAYPASPREWTSIEAELHISYRWDPNLKGMLRPANSLSTVKQDFPDI RQKPPIFGPIFTNINLMGIAPICVMAKRKNGTNVGTLPSTVCNVTFTVDSN QQTYQTYTHNQFRHQPRFPKPNITFPQGTLLDKSSRFCQGRPSSCSTRNF WFRPADYNQCLQISNLSSTA EWVLLDQTRNSLFWENKTKGANQSQTPCVQV LAGMTIATSYLGI SAVSEFFGTSLT PLFHFHISTCLKTQGAFYICGQSIHQ CLPSNWTGTCTIGYVTPDIFIAPGNLSLPIPIYGN SPLPRVRAIHFIPLL AGLGI LAGTGTGIAGITKASLTYSQLSKEIANNIDTMAKALTTMQEQIDSL AAVVLQNRRLDMLTAAQGGICLALDEKCCFWVNQSGKVQDNIRQLLNQAS SLRERATQGWLNWEGTWKWF SWVL PLTG PLVSL LLLLLLFGPCLLNLTITQFV SSRLQAIKLQTNLSAGRHPN IQES PF	331
hENVR	MLGMNMLLITLFLLLPLSMLKGEPWEGCLHCTHTTWSGNIMTKTLLYHTYY ECAGTCLGTCTHNQTTY SVCDPGRGQPYVCYDPKSSPGTWFEIHVGSKEGD LLNQTKVFPSGKDVVSLYFDVCQIVSMGSLFPVIFSSMEYYSSCHKNRYAH PACSTDSPVTTTCWDCTTWSTNQQSLGPIMLT KI PLEPDCKTSTCNSVNLTI LEPDQPIWTTGLKAPLGARVSGEEIGPGAYVYLYIIKKTRTRSTQQFRVFE SFYEHVNQKLPEPPPLASNLFAQLAENIASSLHVASCYVCGGMNMGDQWPW EARELMPQDNFTLTASSLEPAPSSQSIWFLKTSIIGKFCIARWGKAFTDPV GELTCLGQQYYNETLGKTLWRGKSNNSESPHPS PFSRFP SLNHSWYQLEAP NTWQAPSGLYWICGPQAYRQLPAKWSGACVLGTIRPSFFLMP LKQGEALGY PIYDETKRKS KRGITIGDWKDNEWPPERIIQYYGPATWAEDGMWGYRTPVY MLNRIIRLQAVLEIITNETAGALNLLAQQATKMRNVIYQNRALDYL LAQE EGVCGKFNL TNCCLELDDEGKVIKEITAKIQKLAHIPVQTWKG;	332
	LGMNMLLITLFLLLPLSMLKGEPWEGCLHCTHTTWSGNIMTKTLLYHTYYE CAGTCLGTCTHNQTTYSVCDPGRGQPYVCYDPKSSPGTWFEIHVGSKEGDL LNQTKVFPSGKDVVSLYFDVCQIVSMGSLFPVIFSSMEYYSSCHKNRYAHP ACSTDSPVTTTCWDCTTWSTNQQSLGPIMLT KI PLEPDCKTSTCNSVNLTI EPDQPIWTTGLKAPLGARVSGEEIGPGAYVYLYIIKKTRTRSTQQFRVFES FYEHVNQKLPEPPPLASNLFAQLAENIASLHVASCYVCGGMNMGDQWPWE ARELMPQDNFTLTASSLEPAPSSQSIWFLKTSIIGKFCIARWGKAFTDPVG ELTCLGQQYYNETLGKTLWRGKSNNSESPHPS PFSRFP SLNHSWYQLEAPN TWQAPSGLYWICGPQAYRQLPAKWSGACVLGTIRPSFFLMP LKQGEALGY PIYDETKRKS KRGITIGDWKDNEWPPERIIQYYGPATWAEDGMWGYRTPVYM LNRIIRLQAVLEIITNETAGALNLLAQQATKMRNVIYQNRALDYL LAQEE GVCGKFNL TNCCLELDDEGKVIKEITAKIQKLAHIPVQTWKG	333
hENVR(b)	MDPLHTIEKVPARRNIHDRGHQGHMMDGT PGRPKISVQOMTRFSLIIFFL SAPFVVNASTSNVFLQWAHSYADGLQQGDP CWVCGSLPVTNTMELPWWVSP LOGKDWVFFQSFIGDLKQWTGAQMTGVTRKNISEWPINKTLNEPGHDKPFS VNETRDKVIAFAI PLLDTKV FVQTSRPQNTQYRNGFLQIWDGFIWLTATKG HLSQIAPLCWEQRNHS LDNWPNTTRVMGWIPPGQCRHTILLQQRD LFATDW SQQPGLNWYAPNGTQWLCSPNLWPWLP SGWLGCC TLGIPWAQGRWVKTMEV YPYLP HVVNQGTRAIVHRNDHLPTIFMPSVGLGTVIQHI EALANFTQRALN DSLQSI SLMNAEVYMHEDILQNRMALDILTAAEGGT CALIKTECCVYIPN NSRNISLALEDTCRQIQVISSSALS LHDWIASQFSGRPSWWQKILIVLATL WSVGIALCCGLYFCRMFSQHI PQTHSII FQOELPLSPPSQEHYQSQRDI FH SNAP;	334
	DPLHTIEKVPARRNIHDRGHQGHMMDGT PGRPKISVQOMTRFSLIIFFLS APFVVNASTSNVFLQWAHSYADGLQQGDP CWVCGSLPVTNTMELPWWVSP LOGKDWVFFQSFIGDLKQWTGAQMTGVTRKNISEWPINKTLNEPGHDKPFSV NETRDKVIAFAI PLLDTKV FVQTSRPQNTQYRNGFLQIWDGFIWLTATKG LSQIAPLCWEQRNHS LDNWPNTTRVMGWIPPGQCRHTILLQQRD LFATDWS QQPGLNWYAPNGTQWLCSPNLWPWLP SGWLGCC TLGIPWAQGRWVKTMEV YPYLP HVVNQGTRAIVHRNDHLPTIFMPSVGLGTVIQHI EALANFTQRALND SLQSI SLMNAEVYMHEDILQNRMALDILTAAEGGT CALIKTECCVYIPNN SRNISLALEDTCRQIQVISSSALS LHDWIASQFSGRPSWWQKILIVLATLW	335

	SVGIALCCGLYFCRMFSQHIPPQTHSIIIFQQELPLSPPSQEHYQSQRDIFHS NAP	
hENVF(c)1	MARPSPLCLLLLLTLLTPIVPSNSLLTEPPFRWRFYLHETWTQGNRLSTVT LATVDCQPHGCQAQVTFNFTSFKSVLRGWSNPTICFVYDQTHSNCRDYWVD TNGGCPYAYCRMHVTLHTAKKLQHTYRLTSDGRTTYFLTIPDPWDSRWVS GVTGRLYRWPTDSYPVGKLRIFLTYIRVIQVLSNLKDQADNIKHQEEVIN TLVQSHPKADMVTYDDKAEAGPFSWITLVRHGARLVNMAGLVNLSHCFLCT ALSQPLVAVPLPQAFNTSGNHTAHPSGVFSEQVPLFRDPLQPQFPFCYTT PNSSWCNQTYSGSLSNLSAPAGGYFWCNFTLTKHLNISSNNTLSRNLCLPI SLVPRLTLYSEAELSSLVNPPMRQKRAVFPPLVIGVSLTSSLVASGLGTGA IVHFISSSQDLSIKLQMAIEASAESLASLQRQITSVAKVAMQNRALDLIT ADKGGTCMFLGEECCYYINESGLVETSLTLTDKIRDGLHRPSTPNYGGGW WQSPLTTWIIPFISPILIIICLLLLIAPCVLKFIKNRISEVSRVTVNQMLLH PYSRLPTSEDHYDDALTQQAAR;	336
	ARPSPLCLLLLLTLLTPIVPSNSLLTEPPFRWRFYLHETWTQGNRLSTVT ATVDCQPHGCQAQVTFNFTSFKSVLRGWSNPTICFVYDQTHSNCRDYWVD TNGGCPYAYCRMHVTLHTAKKLQHTYRLTSDGRTTYFLTIPDPWDSRWVSG VTGRLYRWPTDSYPVGKLRIFLTYIRVIQVLSNLKDQADNIKHQEEVIN LVQSHPKADMVTYDDKAEAGPFSWITLVRHGARLVNMAGLVNLSHCFLCTA LSQPLVAVPLPQAFNTSGNHTAHPSGVFSEQVPLFRDPLQPQFPFCYTT PNSSWCNQTYSGSLSNLSAPAGGYFWCNFTLTKHLNIS SNNTLSRNLCLPI SLVPRLTLYSEAELSSLVNPPMRQKRAVFPPLVIGVSLTSSLVASGLGTGA IVHFISSSQDLSIKLQMAIEASAESLASLQRQITSVAKVAMQNRALDLIT ADKGGTCMFLGEECCYYINESGLVETSLTLTDKIRDGLHRPSTPNYGGGW WQSPLTTWIIPFISPILIIICLLLLIAPCVLKFIKNRISEVSRVTVNQMLLH PYSRLPTSEDHYDDALTQQAAR	337
hENVF(c)2	MNSPCDRLQQFIQVLLSESWSPSFANTLHWPENLLSYIDELVWQGSQNF HQHEVRFDKPPRLPLTGFSSLTENWSSRQAVS SRLVATAASPPAGCQAPI AFLGLKFSS LGPARKNPALCFLYDQSNKCNSTSWVKENVGCPWHWCNIHEA LIRTEKGS DPMFYVNTSTGGRDGFNGFNLQISDPWDPRWASGVDGGLYEHK TFMYPVAKIRIARTLKT TVTGLSDLASSIQSAEKELTSQLQPAADQAKSR FSWLT LISEGAQLLQSTGVQNL SHCFLCAALRRPPLVAVPLPTPFNYTINS STPIPPVPKGQVPLFSDPIRHKFPFCYSTPNASWCNQTRMLTSTPAPPRGY FWCNSTLT KVLNSTGNHTLCLPISLIPGLTLYSQDEL SHLLAWTEPRPQNK SKWAI FLPLVLGISLASLVASGLGKALT HSIQTSQDLSTHLQLAIEASA ESLDSLQRQITTVAAQVAAQNRQALDLLMAEKGR TCLFLQEECCYYL NESGV VENSLQTLKKKKSKRS ;	338
	NSPCDRLQQFIQVLLSESWSPSFANTLHWPENLLSYIDELVWQGSQNFH QHEVRFDKPPRLPLTGFSSLTENWSSRQAVSSRLVATAASPPAGCQAPIA AFLGLKFSS LGPARKNPALCFLYDQSNKCNSTSWVKENVGCPWHWCNIHEA LIRTEKGS DPMFYVNTSTGGRDGFNGFNLQISDPWDPRWASGVDGGLYEHKT TFMYPVAKIRIARTLKT TVTGLSDLASSIQSAEKELTSQLQPAADQAKSRF SWLT LISEGAQLLQSTGVQNL SHCFLCAALRRPPLVAVPLPTPFNYTINSS TPIPPVPKGQVPLFSDPIRHKFPFCYSTPNASWCNQTRMLTSTPAPPRGYF WCNSTLT KVLNSTGNHTLCLPISLIPGLTLYSQDELSHLLAWTEPRPQNK KWAI FLPLVLGISLASSLVASGLGKALT HSIQTSQDLSTHLQLAIEASAE SLDSLQRQITTVAAQVAAQNRQALDLLMAEKGR TCLFLQEECCYYL NESGV VENSLQTLKKKKSKRS	339
hENVK_{con}	MNPSEMQRKAPPRRRHRNRAPLTHKMNMVTS EEQMKLPSTKKAEPPTWA QLKKLTQLATKYLENTKVTQT PESMLLAALMIVSMVSLPMPAGAAAANYT YWAYVPFPPLIRAVTWM DNPIEVYVND SVWVPGPIDDRCPAKPEE EGMMIN ISIGYRYPICLGRAPGCLMPAVQNWLVVPTVSPISRFTYHMSGMSLRP RVNYLQDFS YQRS LKFRPKGKPCPK EIPKESKNTEVLVWEECVANS AVILQ NNEFGTII DWAPRGQFYHNCSGQTQSCPSAQVS PAVDSDLTESLDKHKHKK LQSFY PWEWGEKGISTPRPKIVSPVSGPEHP ELWRLTVASHHIRI WSGNQ LETDRKPFYTVDLNSSLTVP LQSCVKPPYMLVVGNI VIKPDSQITCENC RLLTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTR	340

	LWNSQSSIDQKLANQINDLRQTVIWMGDRMLSLHRFQLQCDWNTSDFCIT PQIYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTE AIAGVADGLANLNPVTWVKTI GSTT IINL ILILVCLFCLLLVCRCTQQLRR DSDHRE RAMMTMAVLSKRKGGNVGKSKRDQIVTVSV;	
	NPSEMQRKAPPRRRRHRNRAPLTHKMNMVTSEEQMKLPSTKKAEPPTWAQ LKKLTQLATKYLENTKVTQTPE SMLLAALMIVSMVVS LPMPAGAAAANYTY WAYVPFPPLIRAVTWMDNPIEVYVND SVWVPGPIDDRCPAKPEEEGMMINI SIGYRYPPICLGRAPGCLMPAVQNLVEVPTVSPISRFTYH MVSGMSLRPR VNYLQDFS YQ RSLKFRPKGKPCPKEIPKESKNT EVLVWEECVANS AVILQN NEFGT IIDWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKL QSFPYWEWGEKGI STPRPKIVSPVSGPEHPELWRLTVASHHIRIWSGNQTL ETRDRKPFYTVDLNSSLTVPLQSCVKPPYMLVGNIVIKPDSQITITCENCR LLTCIDSTFNWQHRILLVRAREGVWIPVSM DRPWEASPSVHILTEVLKGV LNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHS SVQSVNFVNDWQKNSTR LWNSQSSIDQKLANQINDLRQTVIWMGDRMLSLHRFQLQCDWNTSDFCITP QIYNESEHHWDMVRRHLQGREDNLTLDISKLEQIFEASKAHLNLVPGTEA IAGVADGLANLNPVTWVKTI GSTT IINL ILILVCLFCLLLVCRCTQQLRRD SDHRE RAMMTMAVLSKRKGGNVGKSKRDQIVTVSV	341

PLASMA MEMBRANE RECRUITMENT ELEMENT

[0193] In some aspects, the lipid delivery particle provided herein comprises a plasma membrane recruitment element. The lipid delivery particle disclosed herein can comprise a membrane encapsulating a protein core, which can comprise a plasma membrane recruitment element. The plasma membrane recruitment element can localize itself to the membrane of the lipid delivery particles. The plasma membrane recruitment element can be utilized to recruit a component (e.g., a payload) to the membrane of the lipid delivery particles via forming a chimeric protein of the plasma membrane recruitment element and a component to be localized to the membrane or other mechanisms of attachment. In some cases, the membrane can encapsulate a protein core. In some cases, at least a portion of the plasma membrane recruitment element forms the basic structure of the lipid delivery particle, such as a portion of the protein core inside the lipid delivery particle. In some cases, at least a portion of the plasma membrane recruitment element binds to the membrane of the lipid delivery particle from the inside.

[0194] The plasma membrane recruitment element can play a role in the assembly of the lipid delivery particle, such as packing various components (e.g., a payload) into the lipid delivery particles. The plasma membrane recruitment element can direct budding of the lipid delivery particles from a producer cell. In some cases, expressing plasma membrane recruitment element alone or together with an envelope protein disclosed herein in a producer cell can lead to formation of the lipid delivery particle.

[0195] In some cases, the plasma membrane recruitment element has a viral origin. For instance, the plasma membrane recruitment element comprises a retroviral gag protein, e.g., a retroviral polyprotein that comprises one or more of a matrix (MA) polypeptide, an RNA-binding phosphoprotein polypeptide, a capsid (CA) polypeptide, or a nucleocapsid (NC) polypeptide. The

plasma membrane recruitment element can comprise HIV gag or a mutant thereof. The plasma membrane recruitment element can comprise a gag from murine leukemia virus (MLV) or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a gag from Moloney murine leukemia virus (MMLV) or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Respiratory syncytial virus (RSV) M or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Human Papillomavirus (HPV) L1 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise HPV L2 protein or a mutant thereof. The plasma membrane recruitment element can comprise Hepatitis B virus (HBV) core protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Hepatitis C virus (HCV) core protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise hepatitis E virus (HeV) M protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Chikungunya virus (CHIKV) C-E3-E2-6k-E1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise RSV NP or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Human metapneumovirus (HMPV) M or a biologically active mutant thereof. The plasma membrane can comprise a glycoprotein from a flavivirus. The flavivirus can comprise Chikungunya virus, Zika virus, Dengue virus, or West Nile virus. The plasma membrane recruitment element can comprise Zika virus (ZIKV) C or a biologically active mutant thereof. The plasma membrane recruitment element can comprise ZIKV prM/M or a mutant thereof. The plasma membrane recruitment element can comprise Dengue virus (DENV) C-prM or a biologically active mutant thereof. The plasma membrane recruitment element can comprise West Nile Virus (WNV) prME protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise WNV CprME protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Filovirus VP40 or Z protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Baculovirus P1 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Rotavirus VP7 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Rotavirus VP2 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Rotavirus VP6 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Porcine Circovirus Type 2 (PCV2) capsid or a biologically active mutant thereof. The plasma membrane recruitment element can comprise baculovirus VP2 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise baculovirus VP5 protein or a biologically active mutant thereof. The plasma membrane

recruitment element can comprise baculovirus VP3 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise or baculovirus VP7 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Ebola nucleocapsid or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Parovirus VP1 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Parovirus VP2 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Newcastle disease virus (NDV) M protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Human polyomavirus 2 (JCPyV) VP1 protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise Human parainfluenza virus type 3 (HPIV3) M protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise HPIV3N protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise or Mumps virus (MuV) M proteins or a biologically active mutant thereof. The plasma membrane recruitment element can comprise SARS M protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise SARS E protein or a biologically active mutant thereof. The plasma membrane recruitment element can comprise SARS N protein or a biologically active mutant thereof.

[0196] In some cases, the plasma membrane recruitment element is a mammalian protein or part thereof. For example, the plasma membrane recruitment element can include a pleckstrin homology (PH) domain or a transmembrane domain of a mammalian protein, such as a mouse protein or a human protein. In some cases, the plasma membrane recruitment element has a human origin. Utilizing the plasma membrane recruitment element of a human origin in the lipid delivery particle can give rise to reduced immunogenicity for administration to a human subject. The plasma membrane recruitment element can include a gag from human endogenous retrovirus, such as Human Endogenous Retrovirus K (e.g., HERV-K113, HERV-K101, HERV-K102, HERV-K104, HERV-K107, HERV-K108, HERV-K109, HERV-K115, HERV-K11p22, and HERV-K12q13) and Human Endogenous Retrovirus-W or a mutant thereof. The plasma membrane recruitment element can include a hGAGKcon or a mutant thereof. The plasma membrane recruitment element can include an endogenous gag of a mammal (e.g., human) from retrotransposons (e.g., Arc from vertebrate lineage of Ty3/gypsy retrotransposon), which are also ancestral to retroviruses. In some cases, the plasma membrane recruitment element comprises a portion from human Arc. In some embodiments, the plasma membrane recruitment element may be linked to a retroviral gag protein. In some embodiments, the plasma membrane recruitment element may comprise a retroviral gag protein.

[0197] The plasma membrane recruitment element can include a pleckstrin homology (PH) domain from a human protein or a biologically active mutant thereof. The PH domains can play a role in protein–membrane interactions via binding to phosphatidylinositol phosphate (PIP), for example PIP2 or PIP3, or other lipids or proteins within the membrane of the lipid delivery particles. PH domains with different sequences can have varied affinities and selectivity when binding different PIPs. The plasma membrane recruitment element can include a PH domain of human phospholipase C δ 1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human Akt1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a mutant PH domain of human Akt1 with E17K substitution or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human 3-phosphoinositide-dependent protein kinase 1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human Daap1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of mouse Grp1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human Grp1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of Grp1 (e.g., mouse Grp1) or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human OSBP or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human Btk1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human FAPP1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human CERT or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human PKD or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human PHLPP1 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human SWAP70 or a biologically active mutant thereof. The plasma membrane recruitment element can comprise a PH domain of human MAPKAP1 or a biologically active mutant thereof.

[0198] The plasma membrane recruitment element can also include a membrane protein (e.g., a human membrane protein), a transmembrane domain thereof, or a biologically active mutant thereof. For example, the transmembrane domain of a human protein can be a tetraspanin. In some cases, the plasma membrane recruitment element comprises a transmembrane domain of human CD9 or a biologically active mutant thereof. In some cases, the plasma membrane recruitment element comprises a transmembrane domain of human CD47 or a biologically active mutant thereof. In some cases, the plasma membrane recruitment element comprises a transmembrane

domain of human CD63 or a biologically active mutant thereof. In some cases, the plasma membrane recruitment element comprises a transmembrane domain of human CD81. In some cases, the plasma membrane recruitment element comprises a portion from human Arc.

[0199] The plasma membrane recruitment element can comprise a retroviral gag or a biologically active mutant thereof. The mutant of a retroviral gag can include only a portion of the retroviral gag. The plasma membrane recruitment element can include a gag of an alpha retrovirus. The plasma membrane recruitment element can include a gag of a beta retrovirus or mutant thereof. The plasma membrane recruitment element can include a gag of a gamma retrovirus or mutant thereof. The plasma membrane recruitment element can include a gag of a delta retrovirus or mutant thereof. The plasma membrane recruitment element can include a gag of an epsilon retrovirus or mutant thereof. The plasma membrane recruitment element can include a gag of a spumavirus or mutant thereof. The retroviral gag can include a gag of HIV (e.g., HIV-1), a gag of murine leukemia virus (MLV), a gag of Moloney murine leukemia virus (MMLV), a gag of Simian immunodeficiency virus (SIV), a gag of Rous sarcoma virus (RSV), a gag of human T-cell leukemia virus type-1 (HTLV), or a gag of bovine leukemia virus (BLV), or mutants thereof. The plasma membrane recruitment element can include a gag of HIV (e.g., HIV-1) or a biologically active mutant thereof. The plasma membrane recruitment element can include a gag of MLV or a biologically active mutant thereof. The plasma membrane recruitment element can include a gag of RSV or a biologically active mutant thereof. The plasma membrane recruitment element can include a gag of Friend murine leukemia virus (FMLV) or mutant thereof.

[0200] In some cases, the plasma membrane recruitment element includes those described in **Table 4** with a further truncation on the N-terminus. For example, for those amino acid sequences start with a N-terminal methionine, the N-terminal methionine can be absent. In some cases, the plasma membrane recruitment element includes those described in **Table 4** with a further truncation on the C-terminus. In some cases, the plasma membrane recruitment element includes those described in **Table 4** with one amino acid substitution. In some cases, the plasma membrane recruitment element includes those described in **Table 4** with two or more amino acid substitutions. In some cases, the plasma membrane recruitment element includes those described in **Table 4** and a heterologous peptide sequence fused to the N-terminal or C-terminal.

[0201] In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 50% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane

recruitment element comprises an amino acid sequence that has at least about 60% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 70% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 75% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 80% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 85% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 90% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 95% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 96% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 97% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 98% sequence identity to a sequence set forth in **Table 4**. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 99% sequence identity to a sequence set forth in **Table 4**.

[0202] In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 50% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 60% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 70% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 75% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment

element comprises an amino acid sequence that has at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 80% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 85% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 90% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 95% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 96% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 97% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 98% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389. In some cases, the plasma membrane recruitment element comprises an amino acid sequence that has at least about 99% sequence identity to a sequence set forth in any one of SEQ ID NOs: 342-389.

Table 4. Exemplary plasma membrane recruitment elements and their sequences

Plasma membrane recruitment element	Amino Acid Sequences	SEQ ID NO:
hGAGKcon*	MGQTKSKIISKYASYLSFIKILLKRGGVKVSTKNLIKLFQII EQFCPWF PEQGTLDLKDWKRIKELKQAGRKGNIIP LTVWNDWAI IKAALFPQTE EDSVSVSDAPGSCIIDCNENTRKKSQKETEGHCEYVAEPVMAQSTQNV DYNQLQEVITYPETLKLEKGKPELVGPSESKPRGTSPLPAGQVPVTLQPQ KQVKENKTQPPVAYQYWPPAELQYRPPPESEQGYPGMPAPQGRAPYPQ PPTRRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTLEPMPGEG AQEGEPPTVEARYKSFSIKMLKDMKEGVKQYGPNSPYMRTL LDSIAHGH RLIPYDWEILAKSLSLPSQFLQFKTWIDGVQEQVRRNRAANPPVNIDA DQLLGIGQNWSTISQQALMQNEAIEQVRAICLRAWEEKIQDPGSTCPSFN TVRQGSKEPYPDFVARLQDVAQKSI ADEKARKVIVELMAYENANPEQCS AIKPLKGKVPAGSDVIS EYVKACDGIGGAMHKAMLMAQAITGVVLGGQV RTFGGKCYNCGQIGHLKKNCPVLNKQNITIQATTTGREPPDLCPCKKG KHWSQCRSKFDKNGQPLSGNEQRGQPQAPQQTGAFFIQPFVPQGFQGG QPPLSQVFQGISQLPQYNNCP PPQAAVQQ	342
	GQTKSKIISKYASYLSFIKILLKRGGVKVSTKNLIKLFQII EQFCPWFP EQGTLDLKDWKRIKELKQAGRKGNIIP LTVWNDWAI IKAALFPQTEE DSVSVSDAPGSCIIDCNENTRKKSQKETEGHCEYVAEPVMAQSTQNV	343

	YNQLQEVIY PETLKLEGKGPELVGPSES KPRGT SPLPAGQVPVTLQPQK QVKENKTQPPVAYQYWP PAELQYRPPESQYGYPGMP PAPQGRAPYPQP PTRRLNPTAPPSRQGSSELHEI IDKS RKEGDTEAWQFPVTLE PMPPGEGA QEGEPPTVEARYKSFSI KMLKDMKEGVKQYGPNSPYMRTL LDSIAHGHR LIPYDWEILAKSS LSPSQFLQFKTWIDGVQEQVRNRAANPPVNI DAD QLLGIGQNWSTISQQALMQNEAIEQVRAICLRAWEKI QDPGSTCPSFNT VRQGSKEPYPDFVARLQDVAQKSIADEKARKVIVELMAYENANPECQSA IKPLKGKVPAGSDVISEYVKACDGI GGAMHKAMLMAQAITGVVLGGQVR TFGGKCYNCGQIGHLKKNCVPLNKQNITI QATTTGREPPDLCPRCKKGK HWASQCRSKFDKNGQPLSGNEQRGQPQAPQQTGAFFI QPFVPQGFQGGQ PPLSQVFQGISQL PQYNPCPP PQAAVQQ	
Human Arc	MELDHRTSGGLHAYPGPRGGQVAKPNVILQIGKCR AEMLEHVRRTHRHL LAEVSKQVERELKGLHRSVGKLESNLDGYVPTS DSQRWKKS IKACLCRC QETIANLERWVKREMHVWREVFYRLERWADRLESTGGKYPVGSESARHT VSVGVGGPESYCHEADGYDYTVSPYAITPPPAAGELPGQEPAEAQQYQP WVPGEDGQPS PGVDTQI FEDPREFLSHLEEYLRQVGGSEYWLSQIQNH MNGPAKKWWEFKQGSVKNWVEFKKEFLQYSEGTLSREAIQRELDLPQKQ GEPLDQFLWRKRDLYQTL YVDADEEEIIQYVVGTLQPKLKRFLRHP LPK TLEQLIQRGMEVQDDLEQAAEPAGPHLPVEDEAETLT PAPNSESVASDR TQPE	344
	ELDHRTSGGLHAYPGPRGGQVAKPNVILQIGKCR AEMLEHVRRTHRHL AEVSKQVERELKGLHRSVGKLESNLDGYVPTS DSQRWKKS IKACLCRC ETIANLERWVKREMHVWREVFYRLERWADRLESTGGKYPVGSESARHTV SVGVGGPESYCHEADGYDYTVSPYAITPPPAAGELPGQEPAEAQQYQPW VPGEDGQPS PGVDTQI FEDPREFLSHLEEYLRQVGGSEYWLSQIQNH NGPAKKWWEFKQGSVKNWVEFKKEFLQYSEGTLSREAIQRELDLPQKQ EPLDQFLWRKRDLYQTL YVDADEEEIIQYVVGTLQPKLKRFLRHP LPK LEQLIQRGMEVQDDLEQAAEPAGPHLPVEDEAETLT PAPNSESVASDR TQPE	345
Pleckstrin homology domain of human phospholipase Cδ1 (hPLCδ1)	MDSGRDFLT LHGLQDDEDLQALLKGSQLLKVKS SSWRRERFYKLQEDCK TIWQESRKVMRTPESQLFSIEDIQEVRMGHRTEGLEKFARDVPEDRCFS IVFKDQRNTLDLIA PSPADAQHWWLGLHKI IHHS GSMDQRQKLQHWIHS CLRKADKNKDNKMSFKE LQNF LKELNIQ	346
	DSGRDFLT LHGLQDDEDLQALLKGSQLLKVKS SSWRRERFYKLQEDCKT IWQESRKVMRTPESQLFSIEDIQEVRMGHRTEGLEKFARDVPEDRCFSI VFKDQRNTLDLIA PSPADAQHWWLGLHKI IHHS GSMDQRQKLQHWIHS LRKADKNKDNKMSFKE LQNF LKELNIQ	347
Pleckstrin homology domain of human Akt1 (hAkt1)	MSDVAIVKEGWLH KRGEYIKTWRPRYFLLKNDGTFIGYKER PQDVDQRE APLNNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREE WTTAIQTVADGLKKQEE EEMDFRSGSPSDNSGAEMEVS LAKPKHRVTM NEFEYLKLLGKGT FGKVDPPV	348
	SDVAIVKEGWLH KRGEYIKTWRPRYFLLKNDGTFIGYKER PQDVDQREA PLNNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREEW TTAIQTVADGLKKQEE EEMDFRSGSPSDNSGAEMEVS LAKPKHRVTMN EFEYLKLLGKGT FGKVDPPV	349
	MSDVAIVKEGWLH KRGEYIKTWRPRYFLLKNDGTFIGYKER PQDVDQRE APLNNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREE WTTAIQTVADGLKKQEE EEMDFRSGSPSDNSGAEMEVS LAKPKHRVTM NEFEYLKLLGKGT FGKVDPPV	350
	SDVAIVKEGWLH KRGEYIKTWRPRYFLLKNDGTFIGYKER PQDVDQREA PLNNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREEW TTAIQTVADGLKKQEE EEMDFRSGSPSDNSGAEMEVS LAKPKHRVTMN EFEYLKLLGKGT FGKVDPPV	351
Mutant Pleckstrin homology	MSDVAIVKEGWLH KRGEYIKTWRPRYFLLKNDGTFIGYKER PQDVDQRE APLNNFSVAQCQLMKTERPRPNTFIIRCLQWTTVIERTFHVETPEEREE WTTAIQTVADGLKKQEE EEMDFRSGSPSDNSGAEMEVS LAKPKHRVTM NEFEYLKLLGKGT FGKVDPPV	352

domain of human Akt1 with E17K substitution (hAkt1E17K)	SDVAIVKEGWLHKRGKYIKTWRPRYFLLKNDGTFIGYKERPQDVDQREAPLNNFSVAQCQLMKTERPRNTFIIIRCLQWTTVIERTFHVETPEEREWEWTTAIQTVDGLKKQEEEEEMDFRSGSPSDNSGAEEMEVSLAKPKHRVTMNEFEYLKLLGKGTFGKVDPPV	353
Pleckstrin homology domain of human 3-phosphoinositide-dependent protein kinase 1 (hPDK1)	MKMGPVDKRKGLFARRRQLLLTEGPHLYYVDPVNKVLKGEIPWSQELRPEAKNFKTFFVHTPNRTYYLMDPSGNAHKWCRKIQEVWRQRYQSH	354
	KMGPDVDRKGLFARRRQLLLTEGPHLYYVDPVNKVLKGEIPWSQELRPEAKNFKTFFVHTPNRTYYLMDPSGNAHKWCRKIQEVWRQRYQSH	355
Human CD9	MSPVKGGTKCIKYLLFGFNFI FWLAGIAVLAIGLWLRFDSTQKSI FEQE TNNNNSSFYTGVIILIGAGALMMLVGFLGCCGAVQESQCMLGLFFGFLLVIFAIEIAAAIWGYSHKDEVIKEVQEFYKDTYNKLKTKDEPQRET LKAIHYALNCCGLAGGVEQFISDICPKKDVLETFTVKSCPDAIKEVFDNKFHIIGAVGIGIAVVMIFGMI FSMILCCAIRRNREMV	356
	SPVKGGTKCIKYLLFGFNFI FWLAGIAVLAIGLWLRFDSTQKSI FEQET NNNNS SFYTGVIILIGAGALMMLVGFLGCCGAVQESQCMLGLFFGFLLVIFAIEIAAAIWGYSHKDEVIKEVQEFYKDTYNKLKTKDEPQRET LKAHYALNCCGLAGGVEQFISDICPKKDVLETFTVKSCPDAIKEVFDNKFHIIGAVGIGIAVVMIFGMI FSMILCCAIRRNREMV	357
Human CD47	MWPLVAALLLGSACCGSAQLLFNKT KSVEFTFCNDTVVIPC FVTNMEAQNTTEVYVWKWKFKGRDIYTFDGLNKS TVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTELTREGETIIELKYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTGILILLHYVVFSTAIGLTSFVIAILLVIQVIAYILAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFVE	358
	WPLVAALLLGSACCGSAQLLFNKT KSVEFTFCNDTVVIPC FVTNMEAQN TTEVYVWKWKFKGRDIYTFDGLNKS TVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTELTREGETIIELKYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTGILILLHYVVFSTAIGLTSFVIAILLVIQVIAYILAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFVE	359
Human CD63	MAVEGGMKCVKFLLYVLLLAFCACAVGLIAGVGVAQLVLSQTIIQGATPGSLLPVVIIAVGVFLFLVAFVGGCGACKENYCLMITFAIFLSLIMLVEVAAAIAGYVFRDKVMSEFNNNFRQQMENYPKNNHTASILDRMQADFKCCGAANYTDWEKIPSMSKNRVPDSCCINVTVGCGINFNEKAHKEGCVEKIGGWLKRKNVLVVAAAALGIAFVEVLGIVFACCLVKSIRSGYEV	360
	AVEGGMKCVKFLLYVLLLAFCACAVGLIAGVGVAQLVLSQTIIQGATPGSLLPVVIIAVGVFLFLVAFVGGCGACKENYCLMITFAIFLSLIMLVEVAAAIAGYVFRDKVMSEFNNNFRQQMENYPKNNHTASILDRMQADFKCCGAANYTDWEKIPSMSKNRVPDSCCINVTVGCGINFNEKAHKEGCVEKIGGWLKRKNVLVVAAAALGIAFVEVLGIVFACCLVKSIRSGYEV	361
Human CD81	MGVEGCTKCIKYLLFVFNFFVFWLAGGVILGVALWLRHDPQTNNLLYLELGDKPAPNTFYVGIYILIAVGAVMMFVGFLGCYGAIQESQCLLGTFFFTCLVILFACEVAAGIWGFVNKDQIAKDVKQFYDQALQQAVVDDDDANNAKAVVKTFHETLDCCGSSTLTALTTSVLKNNLCPSGSNIISNLFKEDCHQKIDD LFSGLYLIIGIAAIVVAVIMIFEMILSMVLCCGIRNSSVY	362
	GVEGCTKCIKYLLFVFNFFVFWLAGGVILGVALWLRHDPQTNNLLYLELGDKPAPNTFYVGIYILIAVGAVMMFVGFLGCYGAIQESQCLLGTFFFTCLVILFACEVAAGIWGFVNKDQIAKDVKQFYDQALQQAVVDDDDANNAKAVVKTFHETLDCCGSSTLTALTTSVLKNNLCPSGSNIISNLFKEDCHQKIDDLFSGKLYLIIGIAAIVVAVIMIFEMILSMVLCCGIRNSSVY	363

Pleckstrin homology domain of Human Dapp1	MQTGRTEDDLVP TAPSLGTKEGYLT KQGGLVKTWKTRWFTLHRNELKYF KDQMSPEPIRILDLTECSAVQFDYSQERVNCFCLVFPFRTFYLC AKTGV EADEWIKILRWKLSQIRKQLNQEGGTIR	364
	QTGRTEDDLVP TAPSLGTKEGYLT KQGGLVKTWKTRWFTLHRNELKYFK DQMSPEPIRILDLTECSAVQFDYSQERVNCFCLVFPFRTFYLC AKTGVE ADEWIKILRWKLSQIRKQLNQEGGTIR	365
Pleckstrin homology domain of Mouse Grp1	MPFKI PEDDGNDLTHTFFNP DREGWLLKLGG RVK TWKRRWFILTDNCLY YFEYTTDKEPRGII PLENLSIREVEDPRKPNC FELYNPSHKGQVIKACK TEADGRVVEGNHVYRISAPS PEEKEEWMKSIKASISRDPFYDMLATRK RRIANKK	366
	PFKIPEDDGNDLTHTFFNP DREGWLLKLGG RVK TWKRRWFILTDNCLY YFEYTTDKEPRGII PLENLSIREVEDPRKPNC FELYNPSHKGQVIKACKT EADGRVVEGNHVYRISAPS PEEKEEWMKSIKASISRDPFYDMLATRK RRIANKK	367
Pleckstrin homology domain of Human Grp1	MNP DREGWLLKLGG RVK TWKRRWFILTDNCLY YFEYTTDKEPRGII PLENLSIREVEDPRKPNC FELYNPSHKGQVIKACKT EADGRVVEGNHVYRISAPS PEEKEEWMKSIKASISRDPFYDMLATRK RRIANKK	368
	NPDREGWLLKLGG RVK TWKRRWFILTDNCLY YFEYTTDKEPRGII PLENLSIREVEDPRKPNC FELYNPSHKGQVIKACKT EADGRVVEGNHVYRISAPS PEEKEEWMKSIKASISRDPFYDMLATRK RRIANKK	369
Pleckstrin homology domain of Human OSBP	MGS GSAREGWLFKWTNYIKGYQRRWFVLSNGLLSY YRSKAEMRHTCRGT INLATANITVEDSCNFIISNGGAQTYHLKASSEVERQRWVTALELAKAK AVK	370
	GSGSAREGWLFKWTNYIKGYQRRWFVLSNGLLSY YRSKAEMRHTCRGTI NLATANITVEDSCNFIISNGGAQTYHLKASSEVERQRWVTALELAKAKA VK	371
	SGSAREGWLFKWTNYIKGYQRRWFVLSNGLLSY YRSKAEMRHTCRGTIN LATANITVEDSCNFIISNGGAQTYHLKASSEVERQRWVTALELAKAKAV K	372
Pleckstrin homology domain of Human Btk1	MAAVILESIFL KRSQQKKKTSPLNFKKRLFLLTVHKLSYYEYDFERGRR GSKKGSIDVEKITCVETVVPEKNPPPERQIPRGEESSEMEQISIIERF PYPFQVVYDEGPLYVFS PTEELRKRWIHQ LKNVIRYNSDLVQKYHPCFW IDGQYLCCSQ TAKNAMGCQILENRNGSLKP	373
	AAVILESIFL KRSQQKKKTSPLNFKKRLFLLTVHKLSYYEYDFERGRRG SKKGSIDVEKITCVETVVPEKNPPPERQIPRGEESSEMEQISIIERFP YPFQVVYDEGPLYVFS PTEELRKRWIHQ LKNVIRYNSDLVQKYHPCFWI DGQYLCCSQ TAKNAMGCQILENRNGSLKP	374
Pleckstrin homology domain of Human FAPP1	MEGVLYKWTNYLTGWQPRWFVLDNGILSY YDSQDDVC KSGKSGSIKMAVC EIKVHSADNTRMELIIPGEQHFYMKAVNAAERQRWLVALGSSKACLTDT	375
	EGVLYKWTNYLTGWQPRWFVLDNGILSY YDSQDDVC KSGKSGSIKMAVCE IKVHSADNTRMELIIPGEQHFYMKAVNAAERQRWLVALGSSKACLTDT	376
Pleckstrin homology domain of Human CERT	MSDNQSWNSSGSEEDPETESGP PVERCGVLSKWTNYIHGWQDRWVVLKN NALSYYKSEDETEYGCRGSICLSKAVITPHDFDECRFDISVND SVWYLRAQDPDHRQQWIDAI EQHKTESGYG	377
	SDNQSWNSSGSEEDPETESGP PVERCGVLSKWTNYIHGWQDRWVVLKNN ALSYYKSEDETEYGCRGSICLSKAVITPHDFDECRFDISVND SVWYLRAQDPDHRQQWIDAI EQHKTESGYG	378
	MPVERCGVLSKWTNYIHGWQDRWVVLKNNALSYYKSEDETEYGCRGSIC LSKAVITPHDFDECRFDISVND SVWYLRAQDPDHRQQWIDAI EQHK T	379

	PVERCGVLSKWTNYIHGWQDRWVVLKNNALSYYKSEDETEYGCGRGSICL SKAVITPHDFDECREFDISVNDVWYLRAQDPDHRQQWIDAI EQHKT	380
Pleckstrin homology domain of Human PKD	MGTVMKEGWMVHYTSKDTLRKRHYWRLDSKCITL FQNDTGS RYYKEIPL SEILSLEPVKTSALIPNGANPHCFEITTANVVYVGENVVNPSSPSPNN SVLTSGVGADVARMWEIAIQHALM	381
	GTVMKEGWMVHYTSKDTLRKRHYWRLDSKCITL FQNDTGS RYYKEIPLS EILSLEPVKTSALIPNGANPHCFEITTANVVYVGENVVNPSSPSPNNS VLTSGVGADVARMWEIAIQHALM	382
	TVMKEGWMVHYTSKDTLRKRHYWRLDSKCITL FQNDTGS RYYKEIPLSE ILSLEPVKTSALIPNGANPHCFEITTANVVYVGENVVNPSSPSPNNSV LTSGVGADVARMWEIAIQHALM	383
	MRIQLSGMYNVRKKGKMLPVNRWTRRQVILCGTCLIVSSVKDSLTKGMH VLPLIGGKVEEVKKHQHCLAFSSSGPQSQTYYICFDTFTEYLRWLRQVS KVAS	384
Pleckstrin homology domain of Human PHLPP1	RIQLSGMYNVRKKGKMLPVNRWTRRQVILCGTCLIVSSVKDSLTKGMHV LPLIGGKVEEVKKHQHCLAFSSSGPQSQTYYICFDTFTEYLRWLRQVSK VAS	385
	MDVLKQGYMMKKGHRRKNWTERWFVLKPNII SYVSEDLKDKKGDILLD ENCCVESLPDKDGKKCLFLVKCFDKTFEISASDKKKKQEWIQAIIHSTIH	386
Pleckstrin homology domain of Human SWAP70	DVLKQGYMMKKGHRRKNWTERWFVLKPNII SYVSEDLKDKKGDILLDE NCCVESLPDKDGKKCLFLVKCFDKTFEISASDKKKKQEWIQAIIHSTIH	387
Pleckstrin homology domain of Human MAPKAP1	MDMLS SHHYKSFKVSMIHR LRFTTDVQLGISGDKVEIDPVTNQASTKF WIKQKPISIDSDLLCACDLAE EKSPSHAI FKLTYLSNH DYKHL YFESDA ATVNEIVLKVN YILES	388
	DMLSSH HYKSFKVSMIHR LRFTTDVQLGISGDKVEIDPVTNQASTKFW IKQKPISIDSDLLCACDLAE EKSPSHAI FKLTYLSNH DYKHL YFESDAA TVNEIVLKVN YILES	389

*hGAGKcon is a consensus sequence from ten proviral GAG sequences. The GAG sequences used to derive this consensus GAG sequence are from the following HERVs: HERV-K113, HERV-K101, HERV-K102, HERV-K104, HERV-K107, HERV-K108, HERV-K109, HERV-K115, HERV-K11p22, and HERV-K12q13.

[0203] In some cases, the lipid delivery particle disclosed herein comprises a protein core that is composed of at least a structural protein of a viral origin, for instance, a retroviral gag protein. In some of these cases, the lipid delivery particle comprises a retroviral gag-pro-pol polyprotein, *e.g.*, a gag-pro-pol poly protein from HIV, MMLV, or FMLV, which can help assemble a protein core of the lipid delivery particle. In some of these cases, some of the gag-pro-pol polyprotein is cleaved, *e.g.*, by pro (protease) present freely or in the gag-pro-pol polyprotein. Without wishing to be bound by any particular theory, the cleavage by pro can be inefficient, and the resultant cleavage products can include gag polyprotein, gag-pro polyprotein, free pro, and free pol (polymerase). In some cases, a retroviral gag polyprotein can be further cleaved into MA, CA, NC, and other small fragments, if any. In some other cases, the lipid delivery particle comprises a retroviral gag-pro polyprotein without the pol component, and the gag-pro polyprotein can help form a protein core of the lipid delivery particle. The gag-pro can also be cleaved by pro, in some cases, inefficiently, into separate gag and pro proteins. In some cases, there can be different plasma membrane recruitment elements in a lipid delivery particle. For instance, a gag-pro or gag-

pro-pol polyprotein from one species of virus (*e.g.*, a retrovirus, *e.g.*, a HIV) can help assemble form a protein core of the lipid delivery particle, while a chimeric protein in the lipid delivery particle, discussed *infra*, can comprise a payload fused with a gag protein from a different species of virus (*e.g.*, an MMLV), or from a HERV, or a PH domain or transmembrane domain of a human protein (*e.g.*, a PH domain Akt1 with E17K substitution).

PAYLOAD

[0204] A payload in a lipid delivery particle of the present disclosure can comprise a protein, a polypeptide, a nucleic acid (*e.g.*, DNA or RNA), or any combinations thereof.

[0205] The payload can be a part of the chimeric protein disclosed herein or can comprise a part of the chimeric protein disclosed herein. Alternatively or additionally, the payload can include an entity in the lipid delivery particle separate from the chimeric protein disclosed herein. For instance, in some cases, the payload is a protein or polypeptide coupled to a plasma membrane recruitment element. In some cases, the payload comprises a first moiety (*e.g.*, a nucleic acid-binding protein) that is fused to a plasma membrane recruitment element, and further comprises a second moiety that is coupled to the first moiety via covalent or non-covalent interaction. For instance, the first moiety can be a nucleic acid binding protein that is fused with the plasma membrane recruitment element, and the second moiety can be a nucleic acid molecule that binds to the nucleic acid binding protein.

[0206] In some cases, a payload is directly packaged within the lipid delivery particles and delivered into a target cell in its free form. In some cases, a payload can be fused to a plasma membrane recruitment element (*e.g.*, pleckstrin homology domain) and form a chimeric protein as part of the lipid delivery particles, and then delivered into the target cell. In some cases, the plasma membrane recruitment element (*e.g.*, pleckstrin homology domain) forms at least part of a protein core of the lipid delivery particle. In some embodiments, the payload in its free form or as part of a chimeric protein is within the inside cavity of the protein core of the lipid delivery particles disclosed herein. In some cases, the payload in its free form derives from a cleavage of the chimeric protein comprising the payload.

[0207] In some cases, a lipid delivery particle can deliver more than one payload. Each of the payloads can independently comprise nucleic acid-binding moiety, a nucleic acid-modifying moiety, a fusion protein, or a nucleic acid, or any combinations thereof.

[0208] In some embodiments, the plasma membrane recruitment element and the payload are coupled via any suitable method. Covalent coupling between the plasma membrane recruitment element and a payload peptide can include inteins that can form peptide bonds, direct protein-protein chimeras generated from a single reading frame. In some cases, nucleic acids base pairing

to other nucleic acids via hydrogen bonding interactions (*e.g.*, DNA/RNA, DNA/DNA, or RNA/RNA hybrids), protein-protein binding, or protein-nucleic acid molecule binding can be involved for the coupling between the plasma membrane recruitment element and the payload. Examples of protein-nucleic acid molecule binding include an RNA binding protein (RBP) and an RBP binding sequence (*e.g.*, an RNA) that binds to the RBP. In some embodiments, each of the plasma membrane recruitment element and the payload is fused to a heterologous sequence, and the two heterologous sequences dimerize or multimerize with or without the need for a chemical compound to induce the protein-protein binding, such as a single-stranded nucleic acid sequence or protein dimerization domains). In some embodiments, each of the plasma membrane recruitment element and the payload is fused to one member of a pair of binding partners (*e.g.*, antibody and its target antigen). In some embodiments, the plasma membrane recruitment element is fused to an RBP, and the payload is fused to a RBP binding sequence. Examples of suitable protein domains or nucleic acid molecules for forming the non-covalent connections include single chain variable fragments, nanobodies, affibodies, DmrA/DmrB/DmrC, FKBP/FRB, dDZFs, Leucine zippers, proteins that bind to DNA and/or RNA, optogenetic protein domains that can dimerize or multimerize in the presence of certain light wavelengths, proteins with quaternary structural interactions, and/or naturally reconstituting split proteins. Examples of RBPs and their RBP binding sequences that can be used include a sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in **Table 5**. Examples of RBPs and their RBP binding sequences that can be used include a sequence having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a sequence set forth in any one of SEQ ID NOs: 390-425. In some cases, the RBP comprises an amino acid sequence as set forth in **Table 5**. In some cases, the RBP comprises any one of the sequences set forth in **Table 5**. In some cases, the RBP comprises one or more of the sequences set forth in **Table 5**. In some cases, the RBP comprises more than one of the sequences set forth in **Table 5**. In some cases, the RBP comprises multiple sequences set forth in **Table 5**. In some cases, the RBP binding sequence comprises an amino acid sequence as set forth in **Table 5**. In some cases, the RBP binding sequence comprises any one of the sequences set forth in **Table 5**. In some cases, the RBP binding sequence comprises one or more of the sequences set forth in **Table 5**. In some cases, the RBP binding sequence comprises more than one of the sequences set forth in **Table 5**. In some cases, the RBP binding sequence comprises multiple sequences set forth in **Table 5**.

Table 5. Exemplary RNA binding proteins (RBP) and corresponding RBP binding sequences

RBP	RBP Amino Acid Sequence	SEQ ID NO:	RBP binding sequence	SEQ ID NO:
MS2 (N55 K)	MASNFTQFVLVDNNGGTGDVTVA PSNFANGIAEWISSNSR SQAYKVTCSVRQS SAQKRKYTIKVEVPKVATQTVGGVEL PVAAWRSYLNME LTIP I FATNSDCE LIVKAMQGLLKDGN PIPSAIAANS GIY	390	GGGAGCACAU GAGGAUCACC CAUGUGCGAC UCCACAGUC ACUGGGGAGU CUUCCC	414
	ASNFTQFVLVDNNGGTGDVTVA PSNFANGIAEWISSNSR SQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELP VAAWRSYLNME LTIP I FATNSDCE LIVKAMQGLLKDGNP IPSAIAANS GIY	391		
Com	MGKSI RCKNCNKL LFKADSFDHIEI RCPRCKRH I IMLNA CEHPT EKHC GKREKITHS DETVRY	392	CUGAAUGCCU GCGAGCAUC	415
	GKSI RCKNCNKL LFKADSFDHIEI RCPRCKRH I IMLNAC EHPTE KHCGKREKITHS DETVRY	393		
PP7	MAKTI VLA VGEATRTL TEIQSTADRQIFEEKVGPLVGRRL RLTA SLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPK VRYTQVW SHDVTIVANSTEASRKS LYDLTKSLVATSQVE DLVVNLVPLGRSLE	394	GGAGCAGACG AUAUGGCGUC GCUCCGGCC	416
	AKTI VLA VGEATRTL TEIQSTADRQIFEEKVGPLVGRRL RLTA SLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPKV RYTQVW SHDVTIVANSTEASRKS LYDLTKSLVATSQVED LVVNLVPLGRSLE	395		
TAT BP6.7	MAVPETRPNHTIY INNLSKI KKDELKKS LYAIFSQFGQ ILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFY DKPMRIQYAKTDKRIPAKMKGT FV	396	GGCCAGAUCU GAGCCUGGGA GCUCUCUGGC C	417
	AVPETRPNHTIY INNLSKI KKDELKKS LYAIFSQFGQI LDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYD KPMRIQYAKTDKRIPAKMKGT FV	397		
SLBP	MADFETDES VLMRRQKQINYGKNTIAYDRIYI KEVPRHLR QPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFW D	398	CAAACAAACA AACCAAAGGC UCUUCUCAGA GCCACCCA	418
	ADFETDES VLMRRQKQINYGKNTIAYDRIYI KEVPRHLR QPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFW D	399		
L7Ae	MYVRFEPEDMQNEALS LLEKVRESGKVKKGTNETTKAV ERGLAKLVYIAEDVDPPEIVAH LPL LCEEKNVPYIYVKS KNDLGRAVGIEVPCASAAI INEGELRKELGSLVEKIKGL QK	400	GGGCGUGAUC CGAAAGGUGA CCC	419
	YVRFEPEDMQNEALS LLEKVRESGKVKKGTNETTKAVE RGLAKLVYIAEDVDPPEIVAH LPL LCEEKNVPYIYVKS KNDLGRAVGIEVPCASAAI INEGELRKELGSLVEKIKGL QK	401		
Csy4 (H29 A)	MGDHYLDIRLRPDPEFP PAQLMSVLF GKLAQALVAQGGD RIGVSFPDLDESRLGERLR I HASADDLRALLARPWLE GLRDHLQFGEPAVVPHTPYRQVSRVQA KSNPERLRRRL MRRHDLSEEEARKRIPDTVARALDLPFVT LRSQSTGQHFR LFIRHGPLQVTAE EGGFTCYGLSKGGFVPWF	402	GUUCACUGCC GUAUAGGCAG CUAAGAAA	420
	GDHYLDIRLRPDPEFP PAQLMSVLF GKLAQALVAQGGDR IGVSFPDLDESRLGERLR I HASADDLRALLARPWLEG LRDHLQFGEPAVVPHTPYRQVSRVQA KSNPERLRRRLM RRRHDLSEEEARKRIPDTVARALDLPFVT LRSQSTGQHFR LFIRHGPLQVTAE EGGFTCYGLSKGGFVPWF	403		
Lamb da N	MDAQT RRRERRAEKQAQWKAANPLL VGVS AKPVNRPILS LNRKPKSRVESALNPIDLTVLAEYHKQIESNLQRIERKN QRTWYSKPGERGITCSGRQKIKGKSIPLI	404	GGAUCCGGGC CCUGAAGAAG GGCCCUUCC UUU	421
	DAQTRRRERRAEKQAQWKAANPLL VGVS AKPVNRPILS LNRKPKSRVESALNPIDLTVLAEYHKQIESNLQRIERKNQ RTWYSKPGERGITCSGRQKIKGKSIPLI	405		

Lamb da N (non anti- termin ating)	MDAQT RRRERRAE ERAQWKAAN	406	GCCCUGAAAA AGGGC	422
	DAQTRRRERRAE ERAQWKAAN	407		
HTLV -1 Rex	MPKTRRRPRRSQRKRP	408	UAGGC GACGG UACGCAAGUA CUCUUGCGCC GGCCUA	423
	PKTRRRPRRSQRKRP	409		
HIV-1 REV	MTRQARRNRRRRWRERQR	410	GGCUGGACUC GUACUUCGGU ACUGGAGAAA CAGCC	424
	TRQARRNRRRRWRERQR	411		
UIA (U1hp II)	MHTIY INNLNEKIKKDELKKS LYAIF SQFGQILDILVSR SLKMRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQY ART	412	AGCUUAUCCA UUGCACUCCG GAUGAGCU	425
	HTIYINNLNEKIKKDELKKS LYAIF SQFGQILDILVSR LKMRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYA RT	413		

Nucleic acid binding domains and nucleic acid modifying domains

[0209] In some cases, the payload comprises a nucleic acid-binding moiety, a nucleic acid-modifying moiety, a fusion protein, or a nucleic acid. In some cases, the payload comprises a nucleic acid-binding domain, e.g., a DNA-binding protein domain or polypeptide or an RNA-binding domain or polypeptide e.g., an RNA-binding protein (RBP). A nucleic acid-binding moiety can be capable of binding a nucleic acid. A nucleic acid-binding domain can bind to a nucleic acid in a nonspecific or a site-specific manner.

[0210] In some cases, the nucleic acid-binding moiety binds to a nucleic acid in a site-specific manner. For example, a nucleic acid-binding moiety can comprise an aptamer binding domain that selectively binds to a specific target. In some cases, a nucleic acid-binding moiety recognizes a specific recognition sequence in the target nucleic acid. In some cases, a nucleic acid-binding moiety comprises an aptamer binding domain. In some cases, a nucleic acid binding moiety selectively binds to a sequence or a structural element in a nucleic acid molecule. In some cases, an RNA-binding domain selectively binds to a specific sequence motif in an RNA molecule. In some cases, a nucleic acid-binding moiety selectively binds to a structural element in a nucleic acid molecule. For example, a nucleic acid-binding domain can bind to a stem-loop in a nucleic acid molecule.

[0211] In some cases, a nucleic acid-binding moiety is or comprises a guidable polypeptide domain, a transcriptional regulatory domain, or a nucleic acid-modifying domain. A guidable polypeptide domain can be capable of binding to a polynucleotide (e.g. an RNA guide) that can direct the guidable polypeptide domain a target site. In some cases, the guidable polypeptide domain forms a complex with the RNA guide and recognizes the target sequence through DNA-

RNA base pairing. In some cases, a nucleic-acid binding moiety is or comprises a transcriptional regulatory domain. In other cases, a nucleic-binding moiety can help recruit a transcriptional repressor or activator to a target site. In some cases, a nucleic acid-binding moiety is or comprises a nucleic acid-modifying moiety. In some cases, the present disclosure uses nucleic acid-binding moieties to recruit a nucleic acid-modifying moiety to a target site. In some cases, a nucleic-acid binding moiety comprises catalytic activity. In other cases, a nucleic acid-binding moiety is catalytically inactive. In some cases, a nucleic-acid binding moiety comprising catalytic activity is modified to have a reduced level of activity compared to its wild-type counterpart.

[0212] In some cases, the payload in the present disclosure comprises a nucleic acid modifying domain. A nucleic acid-modifying domain can comprise a polypeptide domain, a nucleic acid or a combination thereof (e.g., a ribonucleoprotein complex). A nucleic acid-modifying domain can be capable of modifying nucleic acid, such as cleaving double-stranded nucleic acid; nicking a single-stranded nucleic acid; introducing a mutation, deletion, or insertion in a nucleic acid; methylating or demethylating a nucleic acid, or altering the structure of DNA (e.g., changing chromatin structure through modifying histones). For example, a nucleic acid modifying domain can comprise a nuclease domain, a nickase domain, a deaminase domain, a polymerase, reverse transcriptase domain, a recombinase domain, a transposase domain, or an epigenetic modifying domain. A nuclease domain can be capable of cleaving phosphodiester bonds between nucleotides in nucleic acids. A nuclease domain can comprise an exonuclease (e.g., a nuclease capable of cleaving nucleic acids from the ends) or an endonuclease (e.g., a nuclease capable of cleaving nucleic acids in the middle). In some cases, a nucleic acid modifying effector or nucleic acid binding domain is a nickase, which can be capable of cleaving a single-strand in a double-stranded DNA. Nucleic acid modifying domains can be useful for gene editing, or for regulating, activating or inhibiting gene expression.

Guidable polypeptide domain

[0213] In some cases, the payload in the present disclosure comprises a guidable polypeptide domain (e.g., a CRISPR-Cas protein domain). In some cases, a guidable polypeptide domain is capable of binding to a polynucleotide (e.g., a RNA guide) that directs it to a target site. In some cases, the guidable polypeptide domain forms a complex with the polynucleotide and recognizes the target sequence through DNA-RNA base pairing. In some cases, a guidable polypeptide domain is a CRISPR/CRISPR-associated (Cas) domain. A CRISPR domain can be a natural or an engineered domain. A Cas protein or domain can be derived from a CRISPR system or share structural and/or functional similarities to a protein involved in a CRISPR system. A CRISPR system is a system encoding DNA sequence arrays known as clustered regularly interspaced short palindromic repeats (CRISPRs), which can be found in microbial genomes or phage genomes. In

some cases, CRISPR systems comprise genes encoding CRISPR-associated (Cas) proteins and/or small RNA guide molecules (e.g., crRNA or tracrRNA) that assemble with the CRISPR domain. In some cases, the CRISPR-Cas domain forms a complex with one or more RNA guide molecules to form an effector ribonucleoprotein complex. The effector ribonucleoprotein complex can recognize a target sequence through sequence specific DNA-RNA base pairing with a spacer sequence in the RNA guide. In some cases, target recognition activates one or more nuclease domains (e.g., a RuvC domain or HNH domain) in the CRISPR domain to make a double-stranded cut at the target DNA. A CRISPR-Cas domain complexed with an RNA guide can be capable of inactivating target gene through a gene knockout. In some cases, the CRISPR domain is used to enable gene insertion and/or deletion, which can inactivate, modify, or restore the gene's function.

[0214] In some cases, the guidable polypeptide domain is any suitable nuclease, e.g., a CRISPR-associated (Cas) protein or a Cas nuclease which functions in a non-naturally occurring CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated) system. In bacteria, this system can provide adaptive immunity against foreign DNA (Barrangou, R., *et al*, "CRISPR provides acquired resistance against viruses in prokaryotes," *Science* (2007) 315: 1709-1712; Makarova, K.S., *et al*, "Evolution and classification of the CRISPR-Cas systems," *Nat Rev Microbiol* (2011) 9:467-477; Garneau, J. E., *et al*, "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA," *Nature* (2010) 468:67-71; Sapranaukas, R., *et al*, "The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*," *Nucleic Acids Res* (2011) 39: 9275-9282).

[0215] Suitable nucleases include CRISPR-associated (Cas) proteins or Cas nucleases including type I CRISPR-associated (Cas) polypeptides, type II CRISPR-associated (Cas) polypeptides (e.g., Cas9 or Cas14), type III CRISPR-associated (Cas) polypeptides, type IV CRISPR-associated (Cas) polypeptides, type V CRISPR-associated (Cas) polypeptides (e.g., Cpf1/Cas12a, C2c1, or C2c3), and type VI CRISPR-associated (Cas) polypeptides (e.g., C2c2/Cas13a, Cas13b, Cas13c, Cas13d).

[0216] One or more components of a CRISPR/Cas system (e.g., modified and/or unmodified) delivered by the lipid delivery particles disclosed herein can be utilized as a genome engineering tool in a wide variety of organisms including diverse mammals, animals, plants, and yeast. A CRISPR/Cas system can comprise a guide nucleic acid such as a guide RNA (gRNA) complexed with a Cas protein for targeted regulation of gene expression and/or activity or nucleic acid editing. An RNA-guided Cas protein (e.g., a Cas nuclease such as a Cas9 nuclease) can specifically bind a target polynucleotide (e.g., DNA) in a sequence-dependent manner. The Cas protein, if possessing nuclease activity, can cleave the DNA (Gasiunas, G., *et al*, "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria,"

Proc Natl Acad Sci USA (2012) 109:E2579-E286; Jinek, M., *et al*, “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity,” Science (2012) 337:816-821; Sternberg, S. H., *et al*, “DNA interrogation by the CRISPR RNA-guided endonuclease Cas9,” Nature (2014) 507:62; Deltcheva, E., *et al*, “CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III,” Nature (2011) 471:602-607), and has been widely used for programmable genome editing in a variety of organisms and model systems (Cong, L., *et al*, “Multiplex genome engineering using CRISPR Cas systems,” Science (2013) 339:819-823; Jiang, W., *et al*, “RNA-guided editing of bacterial genomes using CRISPR-Cas systems,” Nat. Biotechnol. (2013) 31:233-239; Sander, J. D. & Joung, J. K, “CRISPR-Cas systems for editing, regulating and targeting genomes,” Nature Biotechnol. (2014) 32:347-355).

[0217] In some cases, the Cas protein is mutated and/or modified to yield a nuclease deficient protein or a protein with decreased nuclease activity relative to a wild-type Cas protein. A nuclease deficient protein can retain the ability to bind DNA but can lack or have reduced nucleic acid cleavage activity. A protein encoded by a donor sequence comprises a Cas nuclease (*e.g.*, retaining wild-type nuclease activity, having reduced nuclease activity, and/or lacking nuclease activity) can function in a CRISPR/Cas system to regulate the level and/or activity of a target gene or protein (*e.g.*, decrease, increase, or elimination). The Cas protein can bind to a target polynucleotide and prevent transcription by physical obstruction or edit a nucleic acid sequence to yield non-functional gene products. In some cases, the Cas protein cleaves both strands of DNA. In some cases, the Cas protein cleaves one strand of DNA.

[0218] In some embodiments, the nuclease is a Cas protein that forms a complex with a guide nucleic acid, such as a guide RNA (gRNA). In some embodiments, the donor sequence disclosed herein encodes a Cas protein that forms a complex with a single guide nucleic acid, such as a single guide RNA (sgRNA). In some embodiments, the donor sequence disclosed herein encodes a Cas protein that forms a complex with two separate RNA molecules of a dual guide nucleic acid (dgRNA). In some embodiments, the donor sequence in the lipid delivery particles disclosed herein comprises or encodes an RNA-binding protein (RBP) optionally complexed with a guide nucleic acid, such as a guide RNA (*e.g.*, sgRNA, dgRNA), which is able to form a complex with a Cas protein. In some embodiments, the gRNA comprises a scaffolding sequence that tethers the gRNA to the Cas protein. In some embodiments, the gRNA comprises a scaffolding sequence and a spacer sequence that directs the Cas protein to a specific locus. In some embodiments, the scaffolding sequence is configured to bind to the positively charged grooves in the Cas9 protein. In some embodiments, the scaffolding sequence is configured to bind to the Cas protein in the payload. In some cases, Cas undergoes a conformational change when the gRNA binds to the target locus. In some cases, the conformational change in Cas shifts the molecule from an inactive,

non-DNA binding conformation into an active DNA-binding conformation. In some cases, the Cas protein undergoes a conformational change if the spacer sequence has sufficient homology to the sequence at the target locus. In some embodiments, gRNAs can be modified. Exemplary modifications to the gRNA are provided in United States Patent Number 11,479,767 B2, United States Patent Application Publication Number US2020/0339980 A1, and United States Patent Application Publication Number US2021/0079389 A1, each of which is incorporated herein by reference in its entirety.

[0219] One or more components of any suitable CRISPR/Cas system can be delivered by the lipid delivery particle described in the present disclosure. A CRISPR/Cas system can be referred to using a variety of naming systems. Exemplary naming systems are provided in Makarova, K.S. *et al*, “An updated evolutionary classification of CRISPR-Cas systems,” *Nat Rev Microbiol* (2015) 13:722-736 and Shmakov, S. *et al*, “Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems,” *Mol Cell* (2015) 60:1-13. A CRISPR/Cas system can be a type I, a type II, a type III, a type IV, a type V, a type VI system, or any other suitable CRISPR/Cas system. A CRISPR/Cas system as used herein can be a Class 1, Class 2, or any other suitably classified CRISPR/Cas system. Class 1 or Class 2 determination can be based upon the genes encoding the effector module. Class 1 systems generally have a multi-subunit crRNA-effector complex, whereas Class 2 systems generally have a single protein, such as Cas9, Cpf1, C2c1, C2c2, C2c3, or a crRNA-effector complex. A Class 1 CRISPR/Cas system can use a complex of multiple Cas proteins to effect regulation. A Class 1 CRISPR/Cas system can comprise, for example, type I (*e.g.*, I, IA, IB, IC, ID, IE, IF, IU), type III (*e.g.*, III, IIIA, IIIB, IIIC, IIID), and type IV (*e.g.*, IV, IVA, IVB) CRISPR/Cas type. A Class 2 CRISPR/Cas system can use a single large Cas protein to effect regulation. A Class 2 CRISPR/Cas systems can comprise, for example, type II (*e.g.*, II, IIA, IIB) and type V CRISPR/Cas type. CRISPR systems can be complementary to each other, and/or can lend functional units in trans to facilitate CRISPR locus targeting. Examples of Cas proteins that can be used as part of the CRISPR systems described herein include c2c1, Cas13a (formerly C2c2), Cas13b, Cas13c, Cas13d, c2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 or Csx12), Cas10, Cas10d, Cas14, Cas10, Cas10d, CasF, CasG, CasH, Cas12a (formerly Cpf1), Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CasX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, and Cul966, and homologs or modified versions thereof. Examples of mutant Cas9 proteins or Cas9 variants include SpG, SpCas9-NG, Cas9-NRNH, SpG, SpRY, Cas9-VQR, Cas9-EQR, SaCas9-KKH, Nme2Cas9, eNme2-C, eNme2-C.NR, eNme2-T.1, eNme2-T.2, SpRY, eSpCas9(1.1), SpCas9-HF1, nSpCas9,

eSpCas9, Sniper-Cas9, HypaCas9, evoCas9, Cas9TX, HscCas9-v1.2, superFi-Cas9, efSaCas9, SaCas9-HF, Cas9-HF, Cas mini, SaCas9, SpCas9(H840A), dSpCas9, SpCas9(N863A), SpCas9(D839A), SpCas9(H983A), as well as others described in Chuang CK *et al.*, *Int J Mol Sci* 2021 Sep 13;22(18):9872 and Li, T. et al., *Sig Transduct Target Ther* 8, 36 (2023), each of which is incorporated herein by reference in its entirety.

[0220] A CRISPR system can comprise single subunit or multi-subunit effectors. In some cases, a CRISPR system is a Class 1 CRISPR system. A Class 1 CRISPR system can be a type I, type III, or a type IV system. A Class 1 type I CRISPR system can comprise a multi-subunit effector. In some cases, a Class 1 type I CRISPR system comprises a protein or domain in the Cascade-Cas3 protein complex. A Class 1 type I CRISPR system can comprise a Cas6, Cas7, Cas5, Cas11, Cas8, or Cas3 domain. A Class 1 type III CRISPR system can comprise a multi-subunit effector. In some cases, a Class 1 type III CRISPR system comprises a Csm complex or a Cmr complex. In some cases, a Class 1 type III CRISPR system comprises a Cas6, a Cas7 (Csm3 or Cmr4), a Cas7-related (Csm5, Cmr1, or Cmr6), a Cas5 (e.g., Csm4 or Cmr5), a Cas11 (e.g., Csm2 or Cmr3), or a Cas10 (e.g., Csm1 or Cmr2) domain. A Class 1 type IV CRISPR system can comprise a Cas6, a Cas7, a Cas5, a Cas11, a Cas8 (e.g., Csf1), or a DinG or CysH domain. In some cases, a CRISPR system comprises Cmr1, Cmr3, Cmr4, Cmr5, or Cmr6. In some cases, a CRISPR system comprises Csb1, Csb2, or Csb3. A CRISPR system can comprise Csf1, Csf2, Csf3, or Csf4. A CRISPR system can comprise Csn2, Csm2, Csm3, Csm4, Csm5, or Csm6. A CRISPR system can comprise Csc1 or Csc2. A CRISPR system can comprise Cas1, Cas1B, Cas2, or Cas4. A CRISPR system can comprise Csy1, Csy2, or Csy3. A CRISPR system can comprise Cse1 or Cse2. A CRISPR system can comprise Csn2. A CRISPR system can comprise CsaX, Csx1, Csx3, Csx10, Csx14, Csx15, Csx16, or Csx17. In some cases, a CRISPR system comprises a modified version of any one of the foregoing Cas proteins. In some cases, a modified version of the foregoing Cas protein comprises a nickase mutation. In some cases, the nickase mutation corresponds to the D10A mutation of the wild type Cas9 protein. In some cases, the nickase mutation corresponds to the H840A mutation of the wild type Cas9 protein. In some cases, the nickase mutation occurs in the RuvC domain of the wild type Cas9 protein. In some cases, the nickase mutation occurs in the HNH domain of the wild type Cas9 protein. In some cases the RuvC domain can be mutated to prevent cleavage of the non-target DNA strand. In some cases the HNH domain can be mutated to prevent cleavage of the target DNA strand. In some cases, a modified version of the foregoing Cas protein comprises one or more mutations that disrupt cleavage activity. In some cases, a Cas protein with disrupted cleavage activity is catalytically inactive or catalytically dead. In some cases, the catalytically dead mutations occur in the RuvC domain and the HNH domain of the wild

type Cas9 protein. In some cases, the catalytically inactive mutations correspond to the D10A mutation and the H840A mutation of the wild type Cas9 protein

[0221] In some cases, a CRISPR system is a Class 2 CRISPR system. A Class 2 CRISPR system can be a Class 2 type II CRISPR system, a Class 2 type V CRISPR system, or a Class 2 type VI CRISPR system. A Class 2 type II CRISPR system can comprise a Cas9 domain (also known as Csn1 and Csx12). A Cas9 domain can be a SpyCas9, a GeoCas9, a SauCas9, a KhuCas9, a AinCas9, an FmaCas9, a SgaCas9, a ScCas9, a SauriCas9 domain. A Cas9 domain can be a hyperactive Cas9 domain. A Class 2 type V CRISPR system can comprise a Cas12 domain. A Cas12 domain can be a Cas12a, a Cas12b, a Cas12c, a Cas12d, a Cas12e, a Cas12f, a Cas12g, a Cas12h, a Cas12i, a Cas12j, a Cas12k, a Cas12l, or a Cas12m domain. A Class 2 type VI CRISPR system can comprise a Cas13 domain.

[0222] In some cases, a CRISPR system comprises a circularly permuted Cas9.

[0223] In some cases, a CRISPR system comprises CjCas9, Cas13a, Cas13b, Cas13c, or Cas13d. In some cases, a CRISPR system comprises Cas14, xCas9, or SpCas9-NG.

[0224] In some cases, a CRISPR-Cas domain comprises one or more subdomains. For example, a Cas9 domain can comprise a Rec1, a Rec2, a Rec3, a RuvC, an HNH, or a Wedge/PAM-interacting domain. A Cas12 domain can comprise a Rec1, Rec2, a crRNA oligonucleotide binding domain (OBD), a Nuc domain, a PAM-interacting (PI) domain, or a RuvC domain. In some cases, the RuvC domain comprises nuclease activity. In some cases, the HNH domain comprises nuclease activity. The PAM-interacting domain can bind to a protospacer adjacent motif (PAM) sequence that is next to a target sequence in a target nucleic acid molecule. PAM recognition can help activate a nuclease domain to make a cut at the target sequence. In some cases, a CRISPR protein or domain is an engineered or mutated variant of a protein involved in a CRISPR system. An engineered or mutated CRISPR domain can comprise a truncation, a deletion of a part of one or more domains or subdomains, or a mutation of an active site (e.g., a RuvC active site or HNH active site). In some cases, a CRISPR domain with a mutation of one or more active sites is catalytically inactive (e.g., dCas9). In some cases, a CRISPR domain with one or more mutated active sites comprises less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nuclease activity of its wildtype counterpart. For example, a dCas9 can result from the point mutations D10A in the RuvC domain and the point mutation H840A in the HNH domain. In other cases, a mutation can result in a CRISPR nickase. A nickase can generate nick or a single-stranded cut. A nickase can generate a nick in the strand complementary to the RNA guide (e.g., the targeting strand) or in the strand on the non-targeting strand. For example, a RuvC mutation D10A in a Cas9 domain can produce a Cas9 nickase domain that nicks the targeting strand. An

HNH mutation H840A in a Cas9 domain can produce a Cas9 nickase domain that nicks the non-targeting strand.

[0225] A Cas protein can comprise one or more domains. Examples of domains include, guide nucleic acid recognition and/or binding domain, nuclease domains (*e.g.*, DNase or RNase domains, RuvC, HNH), DNA binding domain, RNA binding domain, helicase domains, protein-protein interaction domains, and dimerization domains. A guide nucleic acid recognition and/or binding domain can interact with a guide nucleic acid. A nuclease domain can comprise catalytic activity for nucleic acid cleavage. A nuclease domain can lack catalytic activity to prevent nucleic acid cleavage. A Cas protein can be a chimeric Cas protein that is fused to other proteins or polypeptides. A Cas protein can be a chimera of various Cas proteins, for example, comprising domains from different Cas proteins.

[0226] In some cases, a CRISPR system comprises an Argonaute (Ago) domain.

[0227] Another example of a Cas protein that can be used as part of the prime editor includes Cas14. A Cas14 protein or polypeptide (also termed as “CasZ” protein or polypeptide) can bind and/or modify (*e.g.*, cleave, nick, methylate, demethylate, etc.) a target nucleic acid and/or a polypeptide associated with target nucleic acid (*e.g.*, methylation or acetylation of a histone tail) (*e.g.*, in some cases the CasZ protein includes a chimeric partner with an activity, and in some cases the CasZ protein provides nuclease activity). In some cases, the Cas14 protein or polypeptide is a naturally occurring protein (*e.g.*, naturally occurs in prokaryotic cells) (*e.g.*, a CasZ protein). In other cases, the Cas14 protein or polypeptide not a naturally occurring polypeptide (*e.g.*, the Cas14 protein is a variant Cas14 protein, a chimeric protein, and the like). A Cas14 protein includes 3 partial RuvC domains (RuvC-I, RuvC-II, and RuvC-III, also referred to herein as subdomains) that are not contiguous with respect to the primary amino acid sequence of the Cas14 protein but form a RuvC domain once the protein is produced and folds. A naturally occurring Cas14 protein functions as an endonuclease that catalyzes cleavage at a specific sequence in a targeted nucleic acid (*e.g.*, a double stranded DNA (dsDNA)). The sequence specificity is provided by the associated guide RNA, which hybridizes to a target sequence within the target DNA. The naturally occurring Cas14 guide RNA is a crRNA, where the crRNA includes (i) a guide sequence that hybridizes to a target sequence in the target DNA and (ii) a protein binding segment that binds to the Cas14 protein. Examples of Cas14 proteins include those described U.S. Patent Publication Nos. US20200172886 and US20210214697, Harrington LB *et al.*, *Science*. 2018 Nov 16;362(6416):839-842; Aquino-Jarquín G. *Nanomedicine*. 2019 Jun;18:428-431; each of which is incorporated herein by reference in its entirety. In some cases, the donor sequence disclosed herein encodes Cas14 polypeptide or a nucleic acid molecule encoding Cas14 polypeptide. In some cases, the donor sequence disclosed herein encodes Cas14a polypeptide. In

some cases, the donor sequence disclosed herein encodes Cas14b polypeptide. In some cases, the donor sequence disclosed herein encodes Cas14c polypeptide.

[0228] A Cas protein can be from any suitable organism. Examples include *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Staphylococcus aureus*, *Nocardiopsis dassonvillei*, *Streptomyces pristinae spiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycolides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanothece* sp., *Microcystis aeruginosa*, *Pseudomonas aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccii*, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Fingoldia magna*, *Natronaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochrochium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrogla mobilis*, *Thermosiphon africanus*, *Acaryochloris marina*, *Leptotrichia shahii*, *Leptotrichia wadeii*, *Leptotrichia wadeii* F0279, *Rhodobacter capsulatus* SB1003, *Rhodobacter capsulatus* R121, *Rhodobacter capsulatus* DE442, *Lachnospiraceae bacterium* NK4A179, *Lachnospiraceae bacterium* MA2020, *Clostridium aminophilum* DSM 10710, *Paludibacter propioniceigenes* WB4, *Carnobacterium gallinarum* DMS4847, *Carnobacterium gallinarum* DSM4847, and *Francisella novicida*. In some aspects, the organism is *Streptococcus pyogenes* (*S. pyogenes*). In some aspects, the organism is *Staphylococcus aureus* (*S. aureus*). In some aspects, the organism is *Streptococcus thermophilus* (*S. thermophilus*).

[0229] A Cas protein can be derived from a variety of bacterial species including *Veillonella atypical*, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Listeria seeligeri*, *Listeria weihenstephanensis* FSL R90317, *Listeria weihenstephanensis* FSL M60635, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus kitaharae*, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Fingoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subsp. *Torquens*, *Ilyobacter polytropus*, *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus*

cellulolyticus, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratifractor salsuginis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subsp. *Succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum rubrum*, *Candidatus Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subsp. *Jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subsp. *Multocida*, *Sutterella wadsworthensis*, *proteobacterium*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes*, and *Francisella novicida*.

[0230] A Cas protein as disclosed herein can be a wildtype or a modified form of a Cas protein. A Cas protein can be an active variant, inactive variant, or fragment of a wild type or modified Cas protein. A Cas protein can comprise an amino acid change such as a deletion, insertion, substitution, variant, mutation, fusion, chimera, or any combination thereof relative to a wild-type version of the Cas protein. A Cas protein can be a polypeptide with at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity or sequence similarity to a wild type exemplary Cas protein. A Cas protein can be a polypeptide with at most about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% sequence identity and/or sequence similarity to a wild type exemplary Cas protein. Variants or fragments can comprise at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity or sequence similarity to a wild type or modified Cas protein or a portion thereof. Variants or fragments can be targeted to a nucleic acid locus in complex with a guide nucleic acid while lacking nucleic acid cleavage activity.

[0231] A Cas protein can comprise one or more nuclease domains, such as DNase domains. For example, a Cas9 protein can comprise a RuvC-like nuclease domain and/or an HNH-like nuclease domain. The RuvC and HNH domains can each cut a different strand of double-stranded DNA to make a double-stranded break in the DNA. A Cas protein can comprise only one nuclease domain (*e.g.*, Cpf1 comprises RuvC domain but lacks HNH domain).

[0232] A Cas protein can comprise an amino acid sequence having at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity or sequence similarity to a nuclease domain (*e.g.*, RuvC domain, HNH domain) of a wild-type Cas protein.

[0233] A Cas protein can be modified to optimize regulation of gene expression. A Cas protein can be modified to increase or decrease nucleic acid binding affinity, nucleic acid binding specificity, and/or enzymatic activity. Cas proteins can also be modified to change any other activity or property of the protein, such as stability. For example, one or more nuclease domains of the Cas protein can be modified, deleted, or inactivated, or a Cas protein can be truncated to remove domains that are not essential for the function of the protein or to optimize (*e.g.*, enhance or reduce) the activity of the Cas protein for regulating gene expression.

[0234] In some embodiments, the prime editor delivered by the lipid delivery particles of the present disclosure contain a nuclease-null DNA binding protein derived from a DNA nuclease that can induce transcriptional activation or repression of a target DNA sequence. In some embodiments, the donor sequence encodes a nuclease-null RNA binding protein derived from an RNA nuclease that can induce transcriptional activation or repression of a target RNA sequence. For example, a donor sequence can encode a Cas protein which lacks cleavage activity.

[0235] A Cas protein can be a chimeric protein. For example, a Cas protein can be fused to a heterologous functional domain. A heterologous functional domain can comprise a cleavage domain, an epigenetic modification domain, a transcriptional activation domain, or a transcriptional repressor domain. A Cas protein can also be fused to a heterologous polypeptide providing increased or decreased stability. The fused domain or heterologous polypeptide can be located at the N-terminus, the C-terminus, or internally within the Cas protein.

[0236] The regulation of genes can be of any gene of interest. It is contemplated that genetic homologues of a gene described herein are covered. For example, a gene can exhibit a certain identity and/or homology to genes disclosed herein. Therefore, it is contemplated that a gene that exhibits or exhibits about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% homology (at the nucleic acid or protein level) can be modified. It is also contemplated that a gene that exhibits or exhibits about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity (at the nucleic acid or protein level) can be modified.

[0237] A Cas protein can be provided in any form. For example, a Cas protein can be provided in the form of a protein, such as a Cas protein alone or complexed with a guide nucleic acid. A Cas protein can be provided in the form of a nucleic acid encoding the Cas protein, such as an RNA (*e.g.*, messenger RNA (mRNA)) or DNA.

[0238] The nucleic acid encoding the Cas protein that is part of the prime editor can be codon optimized for efficient translation into protein in a particular cell or organism.

[0239] In some embodiments, a Cas protein is a dead Cas protein. A dead Cas protein can be a protein that lacks nucleic acid cleavage activity.

[0240] A Cas protein can comprise a modified form of a wild type Cas protein. The modified form of the wild type Cas protein can comprise an amino acid change (*e.g.*, deletion, insertion, or substitution) that reduces the nucleic acid-cleaving activity of the Cas protein. For example, the modified form of the Cas protein can have less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nucleic acid-cleaving activity of the wild-type Cas protein (*e.g.*, Cas9 from *S. pyogenes*). The modified form of Cas protein can have no substantial nucleic acid-cleaving activity. When a Cas protein is a modified form that has no substantial nucleic acid-cleaving activity, it can be referred to as enzymatically inactive and/or “dead” (abbreviated by “d”). A dead Cas protein (*e.g.*, dCas, dCas9) can bind to a target polynucleotide but may not cleave the target polynucleotide. In some aspects, a dead Cas protein is a dead Cas9 protein.

[0241] A dCas9 polypeptide can associate with a guide nucleic acid molecule (*e.g.*, PEGRNA) to activate or repress transcription of target DNA. Guide nucleic acid molecules can be introduced into cells expressing the engineered chimeric receptor polypeptide. In some cases, such cells contain one or more different guide nucleic acid molecules that target the same nucleic acid. In other cases, the guide nucleic acid molecules target different nucleic acids in the cell. The nucleic acids targeted by the guide nucleic acid molecule can be any that are expressed in a cell such as an immune cell. The nucleic acids targeted can be a gene involved in immune cell regulation. In some embodiments, the nucleic acid is associated with cancer. The nucleic acid associated with cancer can be a cell cycle gene, cell response gene, apoptosis gene, or phagocytosis gene. The recombinant guide nucleic acid molecule can be recognized by a CRISPR protein, a nuclease-null CRISPR protein, variants thereof, derivatives thereof, or fragments thereof.

[0242] Enzymatically inactive can refer to a polypeptide that can bind to a nucleic acid sequence in a polynucleotide in a sequence-specific manner, but may not cleave a target polynucleotide. An enzymatically inactive site-directed polypeptide can comprise an enzymatically inactive domain (*e.g.*, nuclease domain). Enzymatically inactive can refer to no activity. Enzymatically inactive can refer to substantially no activity. Enzymatically inactive can refer to essentially no activity. Enzymatically inactive can refer to an activity less than 1%, less than 2%, less than 3%, less than 4%, less than 5%, less than 6%, less than 7%, less than 8%, less than 9%, or less than 10% activity compared to a wild-type exemplary activity (*e.g.*, nucleic acid cleaving activity, wild-type Cas9 activity).

[0243] One or a plurality of the nuclease domains (*e.g.*, RuvC, HNH) of a Cas protein can be deleted or mutated so that they are no longer functional or comprise reduced nuclease activity

(*e.g.*, deactivated or dead Cas, *i.e.*, “dCas”). For example, in a Cas protein comprising at least two nuclease domains (*e.g.*, Cas9), if one of the nuclease domains is deleted or mutated, the resulting Cas protein, known as a nickase, can generate a single-strand break at a CRISPR RNA (crRNA) recognition sequence within a double-stranded DNA but not a double-strand break. Such a nickase can cleave the complementary strand or the non-complementary strand, but may not cleave both. If all of the nuclease domains of a Cas protein (*e.g.*, both RuvC and HNH nuclease domains in a Cas9 protein; RuvC nuclease domain in a Cpf1 protein) are deleted or mutated, the resulting Cas protein can have a reduced or no ability to cleave both strands of a double-stranded DNA. An example of a mutation that can convert a Cas9 protein into a nickase is a D10A (aspartate to alanine at position 10 of Cas9) mutation in the RuvC domain of Cas9 from *S. pyogenes*. H939A (histidine to alanine at amino acid position 839) or H840A (histidine to alanine at amino acid position 840) in the HNH domain of Cas9 from *S. pyogenes* can convert the Cas9 into a nickase. An example of a mutation that can convert a Cas9 protein into a dead Cas9 is a D10A (aspartate to alanine at position 10 of Cas9) mutation in the RuvC domain and H939A (histidine to alanine at amino acid position 839) or H840A (histidine to alanine at amino acid position 840) in the HNH domain of Cas9 from *S. pyogenes*.

[0244] A dead Cas protein can comprise one or more mutations relative to a wild-type version of the protein. The mutation can result in less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nucleic acid-cleaving activity in one or more of the plurality of nucleic acid-cleaving domains of the wild-type Cas protein. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the complementary strand of the target nucleic acid but reducing its ability to cleave the non-complementary strand of the target nucleic acid. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the non-complementary strand of the target nucleic acid but reducing its ability to cleave the complementary strand of the target nucleic acid. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains lacking the ability to cleave the complementary strand and the non-complementary strand of the target nucleic acid. The residues to be mutated in a nuclease domain can correspond to one or more catalytic residues of the nuclease. For example, residues in the wild type exemplary *S. pyogenes* Cas9 polypeptide such as Asp10, His840, Asn854 and Asn856 can be mutated to inactivate one or more of the plurality of nucleic acid-cleaving domains (*e.g.*, nuclease domains). The residues to be mutated in a nuclease domain of a Cas protein can correspond to residues Asp10, His840, Asn854 and Asn856 in the wild type *S. pyogenes* Cas9 polypeptide, for example, as determined by sequence and/or structural alignment.

[0245] As examples, residues D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, and/or A987 (or the corresponding mutations of any of the Cas proteins) can be mutated. For example, *e.g.*, D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A, and/or D986A. Mutations other than alanine substitutions can be suitable.

[0246] A D10A mutation can be combined with one or more of H840A, N854A, or N856A mutations to produce a Cas9 protein substantially lacking DNA cleavage activity (*e.g.*, a dead Cas9 protein). A H840A mutation can be combined with one or more of D10A, N854A, or N856A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. A N854A mutation can be combined with one or more of H840A, D10A, or N856A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity. A N856A mutation can be combined with one or more of H840A, N854A, or D10A mutations to produce a site-directed polypeptide substantially lacking DNA cleavage activity.

[0247] In some cases, a dCas9 can be fused to other proteins. In some cases, dCas9 can be fused to SunTag, KRAB, VPS4, P3000, VPR, VP64, V64-p65-Rta, VP160, VP192, HDAC1, DNMT3A, TET1, SPH, KRAB-MeCP2, epigenetic regulators, or other proteins. In some cases, a dCas9 fusion comprises a ZIM3 KRAB-Cas9 fusion. In some cases, a Cas9 fusion can be a paired dCas9 system. In some cases, the dCas9 can be part of a SAM system or REDMAP system. Examples of Cas9 variants and fusion proteins can be found in Li, T. et al., *Sig Transduct Target Ther* **8**, 36 (2023), which is incorporated in its entirety.

[0248] In some embodiments, a Cas protein is a Class 2 Cas protein. In some embodiments, a Cas protein is a type II Cas protein. In some embodiments, the Cas protein is a Cas9 protein, a modified version of a Cas9 protein, or derived from a Cas9 protein. For example, a Cas9 protein lacking cleavage activity. In some embodiments, the Cas9 protein is a Cas9 protein from *S. pyogenes* (*e.g.*, SwissProt accession number Q99ZW2). In some embodiments, the Cas9 protein is a Cas9 from *S. aureus* (*e.g.*, SwissProt accession number J7RUA5). In some embodiments, the Cas9 protein is a modified version of a Cas9 protein from *S. pyogenes* or *S. Aureus*. In some embodiments, the Cas9 protein is derived from a Cas9 protein from *S. pyogenes* or *S. Aureus*. For example, a *S. pyogenes* or *S. Aureus* Cas9 protein lacking cleavage activity.

[0249] Cas9 can generally refer to a polypeptide with at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% sequence identity and/or sequence similarity to a wild type exemplary Cas9 polypeptide (*e.g.*, Cas9 from *S. pyogenes*). Cas9 can refer to a polypeptide with at most about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% sequence identity and/or sequence similarity to a wild type exemplary Cas9 polypeptide (*e.g.*, from *S. pyogenes*). Cas9 can refer to the wildtype or a modified form of the Cas9 protein that can comprise an amino

acid change such as a deletion, insertion, substitution, variant, mutation, fusion, chimera, or any combination thereof.

[0250] In some embodiments, a guidable polypeptidedomain is a Cas9 or variant thereof. In some embodiments, the Cas9 or variant thereof is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or a variant thereof. In some embodiments, a guidable polypeptide domain is Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, or Argonaute (Ago domain), any of which optionally has a nickase activity. In some embodiments, a guidable polypeptide domain comprises an amino acid sequence at least 80%, 85%, 90%, 95%, or 99% identical to any one of sequences listed in **Table 17** below. In some embodiments, a guidable polypeptide domain comprises an amino acid sequence at least 80%, 85%, 90%, 95%, or 99% identical to any one of sequences set forth in **SEQ ID NOs: 426-446**. In some cases, a guidable polypeptide domain is a Cas9 H840A nickase. In some cases, a guidable polypeptide domain is Cas9 D10A nickase. Cas9-H840A. In some cases, a guidable polypeptide domain is a Cas12a/b nickase.

Table 17. Exemplary guidable polypeptide domain sequences

Description	SEQ ID NO:	Sequence
AsCas12 a (Previously known as Cpfl) <i>Acidaminococcus</i> <i>sp.</i> (strain BV3L6) UniProtKB U2UMQ6	426	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFI EEDKARNDHYKELKP IIDRIYKTYADQCLQLVQLDWNLSAAIDSYRKEKTEETRNLALIEEQATYRN AIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHEN ALLRSFDKFTTYFSGFYENRKNVFS AEDI STAI PHRI VQDNFPKF KENCHIF TRLITAVPSLREHFENVKKAIGIFVSTSI EEVFSFPFFYNQLLTQTQIDL YNQ LLGGISREAGTEKIKGLNEVLNLAI QKNDETAHIIASLPHRFIPLFKQILSD RNTLSFILEEFKSDEEVIQSFCKYKTLRNENVLETA EALFNE LNSIDLTHI FISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQ RSLKHE DINLQEIISAAGKELSEAFKQKTSEILSHAHAA LDQPLPTTLKKQEEKEILK SQLDSLLGLYHLLDWFVAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNY ATKKPYSVKFKLNFQMP TLAGSGWDVNKEKNNGAILFVKNGLYYL GIMP KQK GRYKALSFEPTKTSSEGFDKMYDYFPDAAKMI PKCSTQLKAVTAHFQTHHT PILLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKG YREALCKW IDFTRDFLSKYTKTTSIDLSSLRPS SQYKDLGEYYAELNPLLYHISFQRIAE KEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIK LNGQAE LFYRPKS RMKRMARHLGEKMLNKKLKDQKTP IPDTLYQELYDYVNH RLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKFFHV PITLNYQAANS PSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRS LNTIQ QFDYQKKLDNREKERVAA RQAWSVVGTIKDLKQGYLSQVIHEIVDLMIH YQA VVLENLNF GFKS KRTGIAEKAVYQQFEKMLIDKLNCLVLKDYP AEKVG GVL NPYQLTDQFTSFAKMGTSQGF LFYVPAPYTSKIDPLTGFVDPFVWKT IKNHE SRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFORGLPGFMPAWDIVFEKNE TQFDAKGTFPIAGKRIVPIENHRFTGRYRDLYPANELIALLEEKGIVFRDG SNILPKLLENDSDHAIDTMVALIRS VLQMRNSNAATGEDYINS PVRDLNGVC FDSRFQNP EWPMADANGAYHIALKGQLLLNHLKESKDLKLQNGI SNQDWLA YIQELRN
AsCas12a nickase (e.g., R1226A)	427	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFI EEDKARNDHYKELKP IIDRIYKTYADQCLQLVQLDWNLSAAIDSYRKEKTEETRNLALIEEQATYRN AIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHEN ALLRSFDKFTTYFSGFYENRKNVFS AEDI STAI PHRI VQDNFPKF KENCHIF TRLITAVPSLREHFENVKKAIGIFVSTSI EEVFSFPFFYNQLLTQTQIDL YNQ

		LLGGI SREAGTEKIKGLNEVLNLAI QKNDETAH I IAS LPHRFIPLFKQI LSD RNTLS FILEEFKS DEEVIQSFCKYKTLRNENVLETAEALFNELNSIDLTHI FISHKKLET ISSALCDHWDTLRNALYERRISELTGKI TKSAAKEKVQRSLKHE DINLQEIISAAGKELSEAFKQKTSEILSHAHAAALDQPLPTTLKKQEEKEILK SQLDS LLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMEPSLS FYNKARNY ATKKPYSVKEFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPQKQ GRYKALSFEPTSEKTFSEFDMYYDYFPDAAKMI PKCSTQLKAVTAHFQHTT PILLSNNFIEPLEITKEIYDLNNPEKEPKKFQATAYAKKTGDQKGYREALCKW IDFTRDFLSKYTKTTSIDLSSLRPS SQYKDLGEYYAELNPLLYHISFQRIAE KEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIK LNGQAELFYRPKS RMKMAHRLGEEKMLNKKLDQKTP IPDTLYQELDYDVNH RLSHDLSDEARALLPNVITKEVSHEI IKDRRFTSDKFFHV PITLNYQAANS PSKFNQVRNAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSINTIQ QFDYQKKLDNREKERVAAQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQA VVLENLNF GFKS KRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGLV NPYQLTDQFTSFAKMGTSQGF LFYVPAPYTSKI DPLTG FVDPFVWKT IKNHE SRKHFLLEGDFDLHYDVKTGDFILHFKMNRNLSFQRLPGFMPAWDIVFEKNE TQFDAKGTPIAGKRIVPVIENTHRFTGRYRDLYPANELIALLEKGIVFRDG SNILPKLLENDSDHAIDTMVALIRSVLQMANSNAATGEDYINSPVRDLNGVC FDSRFQNPPEWPMADANGAYHIALKGQLLNLHLKESKDLKLQNGISNQDWLA YIQELRN
BhCas12b <i>B. hisashii</i> 1108 AA 130.4 kDa	428	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLI RQEA IYEHHEQ DPKNPKKVS KAEIQAELWDFVLKMQKCNS FTHEVDKDEVFNILRELYEELVP SSVEKKGEANQLSNKFLYPLVDPNS QSGKGTAS SGRKPRWYNLKIAGDPSWE EEKKKWEEDKKKDLAKILGKLAEGLIPLFIPTDSNEPIVKEIKWMEKSR NQSVRRDLKDMFIQALERFLSWESWNLKVKEEY EKVEKEYKTEERIKEDIQ ALKALEQYKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPS EKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLATFCEI DKKKKDAKQQTFTLADPINHPLWVRFEERSGSNLNKYRILTEQLHTEKLLK KLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTY KDESIKFPLKGTLLGARVQFDRDHLRRYPHKVESGNVGRIFYNMTVNIEPTE SPVSKSLKIHRRDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMS IDLQORQAAAASIFEVVDQKPDIEGKLFPIKGTELYAVHRASFN IKLPGET LVKSREVLRLKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTWKWISQ ENSDVPLVYQDEL IQIRELMYKPYKDWVAFKQLHKKRLEVEIGKEVKHWRKS LSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRF AIDQ LNHLNALKE DRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQII LFEDLSN YNPYEERSRFENS KLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAK TGSPGIRCSVVTKEKLQDNRFKLNQREGRLTLDKIAVLKEGDLYPDKGGEK FISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVKYCKAYQVDGQTVYIP ESKDQKQKIIIEFEGEGYFILKDGVEWVNAGKLKIKKGSSKQSSSELVDSDI LKDSFDLASELKGEKMLYRDP SGNVFVPSDKWMAAGVFFGKLERILISKLTN QYSISTIEDSSKQSM
LbaCas12a <i>L. bacterium</i> 1228 AA 143.9 kDa	429	MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLEVEDEKRAEDYKGVKK LLDRY YLSFINDVLHSIKLKNLNNYISLFRKKT RTEKENKELENLEINLRKE IAKAFKGNEGYKSLFKKDI IETILPEFLDDKDEIALVNSFNGFTTAFTGFFD NRENMFSEEAKSTIAFR CINENLTRYISNMDIFEKVDAIFDKHEVQEI KEK ILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYIN LYNQKTKQKLPKFKPLYKQVLS DRESLSFYGEGYTSDEEVLEVFRNTLNKNS EIFFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIRDKWNA EYDDIHLKKKAVVTEKYEDDRKSFKKIGSFSLEQLQEYADADLSVVEKLKE IIIQKVDEIYKVYGSSEKLFDAADFVLEKSLKKNDAVVAIMKDLLDSVKS FEN YIKAFFGEGKETNRDES FYGDFVLAYDILLVKDHIYDAIRNYVTQKPYSKDK FKLYFQNPQFMGGWDKDKETDYRATILRYGSKY YLAIMDKKYAKCLQKIDKD DVNGNYEKI NYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKG DMFNLDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEEQ GYKVSFESASKKEVDKLVVEEGKLYMFQIYNKDFSDKS HGTPNLHTMYFKLLF DENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDPNPKTTTSL YDVYKDKRFSEDQYELHIPIA INKCPKNI FKINTEVRVLLKHDDNPYVIGID RGERNLLYIVVVDGKGNIVEQYSLNEIINNFGIRIKTDYHSLLDKKEKERF EARQNWTSIENIKELKAGYISQVVKICE LVEKYDAVIALEDLNSGFKNSRV KVEKQVYQKFEKMLIDKLNMYVDKKS NPCATGGALKGYQITNKFESFKSMST

		QNGFI FYIPAWLT SKID PSTG FVNL LKTKYTSI ADSKKFIS SFDRIMYVPEE DLFEFALDYKNFSRTDADYIKKWKL YSYGNRIRIFRN PKKNNVFDWEEVCLT SAYKELFNKYGINYQQGDIRALLCEQSDKAFYS SFMAIMSLMLQMRNSITGR TDVDFLISPVKNS DGI FYDSRNYEAQENA ILPKNADANGAYNIARKVLWAIG QFKKAEDEKLDKVKIAI SNKEWLEYAQTSVKH
Cas9 nickase (Met minus) <i>Streptococcus</i> <i>pyogenes</i> Q99ZW2 Cas9 with R863A	430	DKKYS IGLD IGTNSVGWAVIT DEYKVPSKKFKVLGNT DRHS IKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEESF LVEEDKKHERHPI FGNI VDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAK AII SARLSKSRRL ENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAE AKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLG ELHAILRRQEDFY PFLKDNREKIEKILTFRIPYVVG PLARGNSRFAWMTRKS EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLS GEQKKAIVD LLLFKTNRKVTVKQLKEDYFKKIE CFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLT FEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK TILDFLKS DGFANRNF MQLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAG SPAIIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDV DHI V PQS FLKDDSIDNKVLT RSDKNAGKSDNPSEEVVKMKMKNYWRQ LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD RMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVG TALIKKYPKLES EFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSN IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVN IVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAGKYKEVKKDLIIK LPKYSLFEL ENGRKRLASAGELQKGNE LALPSKYVNFLYLASHYEKLGKSP EDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVL SAYNKH RDKP IREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVL DATLIHQSI GLYETRIDLSQLGGD
dead Cas9 or dCas9 <i>Streptococcus</i> <i>pyogenes</i> Q99ZW2 Cas9 with D10X and H810X Where "X" is any amino acid	431	MDKKYSIGLXIGTNSVGWAVIT DEYKVPSKKFKVLGNTDRHS IKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPI FGNI VDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRL ENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHL GELHAILRRQEDFY PFLKDNREKIEKILTFRIPYVVG PLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVY YNELTKVKYVTEGMRKPAFLS GEQKKAIVD LLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLA GSPAIIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR LSDYDV DHI V PQS FLKDDSIDNKVLT RSDKNRGKSDNPSEEVVKMKMKNYWR QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAHDAYL LNAVVG TALIKKYPKLES EFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSN NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVN NIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAGKYKEVKKDLIIK KLPKYSLFEL ENGRKRLASAGELQKGNE LALPSKYVNFLYLASHYEKLGKSP PEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVL SAYNKH RDKP PIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD
LfCas9	432	MKEYHIGLDIGTSSIGWAVTDSQFKLMRIKGKTAIGVRLFEEGKTA AERRTF RTTRRLKRRKWRHLHYLDEIFAPHLQEV DENFLRRLKQSNIH PEDPTKNQAF IGKLLFPDL LKKNERGYPTLI KMRDEL PVEQRAHYPVMNIYKLREAMINEDR

<i>Lactobacillus fermentum</i> wild type GenBank: SNX31424.11		QFDLREVVYLA VHHIVKYRGHFLNNA SVDKFKVGRIDF DKS FNV LNEAYEELQ NGEGS FTIE PSKVEKIG QLLLDTKMRKLD RQKAVAKL LEVKVADKEET KRNK QIATAMSKLVLGY KADFATVAMANGNEWKIDLS SETS EDEI EKFREELS DAQ NDILTEITS LFSQIMLNEIVPNGMS ISESMMDRYW THERQLAEVKEYLATQP ASARKEFDQVYNKYIGQAPKERGF DLEKGLKKI LSKKENWKEIDE LLKAGDF LPKQRTSANGVIPHQMHOQELDR IIEKQAKYYPWLATENPATGERDRHQAKY ELDQLVSFRIPYYVGPLVTPEVQKATSGAKFAWAKRKEDGEITPWNLWDKID RAESAEAFIKRMTVKDTYLLNEDVL PANS LLYQKYNVLNELLNNVRVNGRRLS VGIKQDIYTELFKKKKT VKAS DVAS LVMAKTRGVNKP SVEGLSDPKKFN SNL ATYLDLKSIVGDKVDDNRYQT DLENI IEWRSVFEDGEIFADKLTEVEWL TDE QRSALVKKRYKGWGRLSKKLLTGIVDENGQRIIDLMWNTDQNFKEIVDQPVF KEQIDQLNQKAITNDGMTLRERVESVLDDAYTS POKKAIWQVVRVVEDIVK AVGNAPKSI SIEFARNEGNGKEITRSRRTQLQKLFEDQAHEL VKDTS LT EEL EKAPDLSDRYFYFTQGKDMYTGPINFDEISTKYDIDHILPQS FVKDNSL DNRVLT SRKENNKKSDQVPAKLYAAKMKP YWNQLLQGLITQRKFENLT KDV DQNIKYRSLGFVKRQLVETRQVIKLTANI LGSMYQEAGTEI IETRAGLT KQL REEFDLPKRVENDYHHA VDAYLTTFAGQYLNRRYPKLRSFVYGEYMKFKH GSDLKLRNFNFHELMEGDKS QGKVVDQQTGELITTRDEVAKSFDRLNMGKY MLVSKEVHDRSDQLYGATIVTAKESGKLTSPIEIKKNRLVDLYGAYTNGTSA FMTI IKFTGNKPKYKVI GIPTTSAASLKRAGKPGSES YNQLHRI IKS NPKV KKGFEIVVP HVSYGQLIVDGDCKFTLASPTVQH PATQLVLS KKSLETIS SGY KILKDKPAIANERLIRVFDEVVGQMNRYFTIFDQRSNRQKVADARDKFLSLP TESKYEGAKKVQVGKTEVITNLLMGLHANATQGD LKVLGLATFGFFQSTTGL SLED TMIVYQSPTGLFERRICLKD I
SaCas9 <i>Staphylococcus aureus</i> wild type GenBank: AYD60528.1	433	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLOEIFS NEMAKVDDSF FHRLEES FLVEEDKKHERHP IFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVKLF IQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENL I AQLPGEKKNGLFGNLIALS LGLTPNFKSNFDLAE DAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHA ILRRQEDFY PFLKDNREKIEKILTFRIPYYVG PLARGNSRFAWMTRK SEETITPWN FEEVVDKGASAQS FIERMTNFDKNLPNEKVLPHKS LLYEYFTV YNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGT YHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILD FLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGD SLHEHIANLA GSPAIIKKGILQTVKVVD ELVKVMGRHKPENIVI EMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR LSDYD VDHIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWR QLLNAKLITQRKFDNLT KAERGGLS ELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAH DAY LNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYS NIMNFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGS PEDNEQQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVLSAYNKHRDK PIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD
LcCas9 <i>Lactobacillus crispatus</i> NCBI Reference Sequence: WP_1334780 44.1	434	MKIKNYNLA LTPSTSAVGHVEVDDDLNILEPVHHQKAIGVAKFGEGETAEAR RLARSARRTKRRANRINH YFNEIMKPEIDKVDPLMFDRIKQAGLSPLDERK EFRTVIFDRPNIASYYHNQFP TIWHLQKYLMITDEKADIRLIY WALHSL LKH RGHFFNTTPMSQFKPGKLN LKDDMLALDDYNDLEGLS FAVANSPEIEKV IKD RSMHKKEKIAELKKLIVNDVPDKDLAKRNNKIITQIVNAIMGNSFHLNFI FD MDLDKLTSAKWSFKLDDPELDTKFDAISGSMTDNQIGIFETLQKIYSAISLL DILNGSSNVDAKNALYDKHKRDLNLYFKFLNTLPDEIAKT LKAGYTLYIGN RKKDL LAARKLLKVNVAKNFS QDDFYKLINKELKSIDKQGLQTRFSEKVGEL VAQNNFLPVQRSSDNVFI PYQLNAITFNKILENQGKYD FFLVKPNPAKKDRK NAPYELS QLMQFTIPYYVGPLVTPEEQVKS GIPKTSRFAWMVRKDNGAITPW NFYDKVDIEATADKFIKRSIAKDSYLLSELVLPKHSLLYEKYEVEFNELSNVS

Wild type		LDGKKLSGGVKQI LFNVEVFKTKNVNTSRILKALAKHNIPGSKITGLSNPEE FTSSLQTYNAWKKYFPNQIDNFAYQQDLEKMI EWSTVFEDHKILAKKLDEIE WLDDDQKKFVANTRLRGWGRLSKRLLTGLKDNYGKSIMQRLETTKANFQQIV YKPEFREQIDKISQAAAKNQSLEDILANSYTSPSNRKAIRKTMSSVVDEYIKL NHGKEPDKIFLMFQRSEQEKGKQTEARSKQLNRILSQLKADKSANKLFSKQL ADEFSNAIKKSKYKLNDKQYFYFQQLGRDALTGEVIDYDELYKYTVLHIIPR SKLTDDSQNNKVLTKYKIVDGSVALKFGNSYS DALGMPKAFWTELNRLKLI PKGKLLNLTDFSTLNKYQRDGYIARQLVETQQIVKLLATIMQSRFKHTKII EVRNSQVANIRYQFDYFRIKNLNEY YRGFDAYLAAVVGTYLYKVYPKARRLF VYGQYLKPKKTNQENQDMHLDSEKKSQGFNFWNLLY GKQDQIFVNGTDVIA FNRKDLITKMTVYNYKSQKISLAIDYHNGAMFKATLFPRNDRDTAKTRKLI PKKKDYDTDIYGGYTSNVDGYMLLAEI IKRDGNKQYGFYGVPSRLVSELDTL KKTRYTEYEEKLKEI IKPELGVDLKKIKKIKILKNKVPFNQVIIDKGSKFFI TSTSYRWNYRQLILSAESQQT LMDLVVDFDFS NHKARKDARKNADERLIKVY EELIYQVKNYMPMFVELHRCYEKLVD AQTKFSLKISDKAMVLNQILILLHS NATSPVLEKLG YHTRFTLGKKHNLISENAVLTQSITGLKENHVS IKQML
StCas9 <i>Streptococcus thermophilus</i> UniProtKB/S wiss-Prot: G3ECR1.2 Wild type	435	MLFNKCIISINLDFSNKEKCMTKPYSIGLDIGTNSVGWAVITDNYKVP SKK MKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLLKRTARRRYTRRRNRILYLQ EISTEMATLDDAFFQRLDDSLVFPDDKRDSPYPIFGNLVEEKVYHDEFPTI YHLRKYLDSTKKADLRVLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQD FLDTYNALFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNGIFS EFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFL KAKKLYDAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNIS LKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDR EDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPPFLAKNKERIEKILTFR IPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFD LYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVRL YFKDKRKVTDKDIIEYLAHIYGYDGIELKGIEKQFNSSLSTYHDLNIIINDK EFLDDSSNEAIEEIIHTLTI FEDREMIKQRLSKFENIFDKSVLKKLSRRHY TGWGKLSAKLINGIRDEKSGNTILDYLDGDISNRNFMQLIHDDALSFKKKI QKAQIIGDEDKGNIKEVVKS L PGSPAIIKKGILQSIKIVDELVKVMGGRKPES IVVEMARENQYTNQGKSN SQRLKRLEKSLKELGSKILKENIPAKLSKIDNN ALQNDRLYLYYLQNGKDMYTGD DLDIDRLSNYDIDHIIPQAFLKDNSIDNKV LVSSASNRGKSDDFPSLEVVKRKT FWYQLLKS KLISQRKFDNLT KAERGGL LPEDKAGFTQRQLVETRQITKHVARLLDEKFNNKKDENNRVAVRTVKIITLKS TLVSQFRKDFELYKVREINDFHHAH DAYLNAVIASALLKKY PKLEPEFVYGD YPKYN SFREKRSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGE SVWNKESDLATVRRVLSYPQVNVVKVVEEQNHGLDRGPKGLFNANLSSKPK PNSNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQ GISILDRINYRKDKLNFLEKGYKDIELIELPKYSLFELS DGSRRMLASIL STNNKRGEI HKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEE LFYYILEFNENYVGAKKNGKLLNSA FQSWQNHS IDELCSSFIGPTGSERKGL FELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLA KLGE
PdCas9 <i>Pedicoccus damnosus</i> NCBI Reference Sequence: WP_0629132 73.1 Wild type	436	MTNEKYSIGLDIGTSSIGFAVVNDNNRVI RVKGKNAIGVRLFDEGKAAADRR SFRTRRSFRTRRRRLSRRWRLLKLREIFDAYITPVDEAFFIRL KESNLSP KDSKKQYSGDILFNDRS DKDFYEKYPTIYHLRNALMT EHRKFDVREIYLAIH HIMKFRGHFLNATPANNFKVGRNLNLEEKFEELNDIYQRVFPDESIEFRTDNL EQIKEVLLDNKRSRADRQRTLVS DIYQSS EKDIEKR NKAVATEILKAS LGN KAKLNVITNVEVDKEAAKEWSITFDSESI DDDLAKIEGQMTDDGHEIIEVLR SLYSGITLSAIVPENHTLSQSMVAKYDLHKDHLKLFFKKLINGMTDTKKAKNL RAAYDGYIDGVKGKVLPOEDFYKQVQVNLDDSAEANEIQTYIDQDIFMPKQR TKANGSIPHQLQQQELDQIENQKAYYPWLAELNPNPDKKRQQLAKYKLDL VTRVFPYYVGPMITAKDQKNQSGAEFAWMIRKEPGNITPWNFDQKVDRMATA NQFIKRMTTDTYLLGEDVLP AQSLLYQKFEVLNELLKIRI DHKPSIEQKQ QIFNDLFKQFKNVTIKHLQDYLVSQGQYSKRPLIEGLADEKRFNSSLSTYS LCGIFGAKLVEENDRQEDLEKII EWSTIFEDKKIYRAKLNLDLTWLTDQKEK LATKRYQGWGRLSRKLLVGLKNSEHRNIMDILWITNENFMQIQAEPDFAKLV TDANKGMLEKTDSQDVINDLYTSPQNKKAIRQILLVVHDIQNAMHGOAPAKI HVEFARGEERNPRRSVQRQRQVEAA YEKVSNELVSAKVRQEFKEA INNKRDF KDRLFLYFMQGGIDIYTGKQLNIDQLSSYQIDHILPQAFVKDDSLTNRVLTN ENQVKADSVPIDIFGKKMLSVWGRMKDQGLISKGYRNLTMNPENISAHTEN

		GFINRQLVETRQVIKLA VNILADEYGDSTQIISVKADLSHQMREDFELLKNR DVNDYHHAFFDAYLAAFI GNYLLKRY PKLESYFVYGD FKKFT QKET KMRRFNF IYDLKHCDQVVNKETGE ILWT KDED IKYI RHLFAYKK ILVS HEVREKRGALY NQTIYKAKDDKGS GQES KKLIRIKDDKETKIYGGYSGKSLAYMTI VQIT KKN KVSYRVIGI PTLALARLNKLENDSTENNGELYKIIKPQFTHYKVDKNGEII ETDDDFKIVVSKVRFQQLI DDAGQF FMLASDTYKNNAAQQLV ISNNALKA INN TNITDCPRDDLRLDNRLDS AFDE IVKKMDKY FSAY DANN FREK IRNS NLI FYQLPVEDQWENNKITELGKRTVLT RILQGLHANATT DMS IFKI KTFPGQL RQRSGISLS ENAQLIYQSPTGLFERRVQLNKIK
EcCas9 <i>Enterococcus cecorum</i> NCBI Reference Sequence: WP_0473385 01.1 Wild type	437	RRKQRIQILQELLGEEVLKTD PGFFHRMKESRYVVEDKRTL DGKQVELPYAL FVDKDYTDKEYYKQFPT INHL IVYLMTTS DTPDIRLVYLALHYMKNRGNFL HSGDINNVDKINDILEQLDNVLETFLDGWNLLKLSYVEDIKNIYNRDLGRGE RKKAFVNTLGAKT KA EKAFCS LISGGSTNLAEL FDDS SLKEIETPKIEFASS SLEDKIDGI QEAL EDRFAVIEAAKRLYDWKTLT DILGDSSSLAEARVNS YQM HHEQLLELKS LVKEYLDRKVFQEVFVSLNVANNYPAY IGH TKINGKKKELEV KRTKRNDFY SYVKKQVI EPIKKKVS DEAVLT KLSEIESLIEVDKY LPLQVNS DNGVI PYQVKLNE LTRI FDNLENRI PVLRENDRKIIKT FKFRIPIYVGS LNG VVKNGKCTNWMVRKEEGKIYPWNFEDKVDLEASAEQF IRRMTNKCTYLVNED VLPKYSLLYSKYLVLSELNNLRIDGRPLDVKIKQDIYENVFKNRKVTLKKI KKYLLKEGI ITDDDELS GLADDVKS SLTAYRDFKEKLGHLDLSEA QMENIIL NITLFGDDKLLKRLAALYP FIDDKSLNRIAT LNYRDWGRLSERFLSGITS VDQETGELRTIIQCMYETQANLMQL LAEP YHFVEAIEKENPKVDLESISYRI VNDLYVSPA VKRQIWQT LLVI KDIKQVMKHDPERIFI EMAREKQESKKT KSR KQVLS EVYKKAKE YEHL FEKLNSLT EEQLRSKKIYLY FTQLGKCMYSGEPI FENLV SANSNYDI DHIY PQSKTIDDS FNNIVLVKKS LNAYKSNHY PIDKNIR DNEKVKT LWNTLVSKGL ITKEKYERLIRSTPFS DEELAGFIARQLVETRQST KAVAE ILSNWFPE SEIVYSKAKNVSNFRQDFEILKVRELNDCHHAHDAY LNI VVGNA YHTKFTNS PYRF IKNKANQE YNLRLKLLQKVNKIESNGVVAWVGQSEN NPGTIATVKKVIRNTVLISRMVKEVDGQLFDLTLMKKKGKGVPI KSSDERL TDISKYGGYNKAT GAYFTFVKSKRGKVRSFEYVPLHLSKQFENNNE LKE YIEKDRGLTDVEI LIPKVLINSLFRYNGSLVRITGRGDTRL LLVHEQPLYVS NSFVQQLKSVSSY KLKXSENDNAKLTKTATEKLSNIDELYDGLLRKLDLPIY SYWFSSIKE YLVE SRTKYIKLSIEEKALVIFEILHLFQSDAQVPLKILGLS TKPSRIRIQKNLKD TDKMSIIHQSPSGIFEHEIELTSL
AhCas9 <i>Anaerostipes hadrus</i> NCBI Reference Sequence: WP_0449242 78.1 Wild type	438	MQNGFLGITVSSEQVGWAVTN PKYELERASRKDLWGVRLFDKAETAEDRRMF RTNRRLNQRKKNRIHYLRDIFHEEVNQKDPNFFQQLDES NFCEDDR TVEFNF DTNLYKNQFPTVYHLRKYLMETKDKPDIRLVYLA FSKFMKNRGHFLYKGNLG EVMDFENSMKGFCESLEKFNI DFPTLSDEQVKEVRDI LCDHKIAKTVKKKNI ITITKVKSKTAKAWIGLFCGC SVPVKVLFQDI DEEIVTDPEKIS FEDASYDD YIANIEKGVGIYYEAI VSAKMLFDWS ILNEILGDHQL LSDAMIAEY NKHHDD LKRLQKIIKGTGSR ELYQDIFINDVSGNYVCYVGHAKTMS SADQKQFYT FLK NRLKNVNGI SSED AEWI DTEI KNGTLLPKQTKRDN SVIPHQLQLREFELILD NMQEMY PFLKENREKLLKIFNFVIPYYVG PLKGVVRKGESTNWMVPKKG DVI HPWNFDEMVDKEASAEC FISMTGNCSYL FNEKVL PKNSLLYET FEVLNELN PLKINGEPI SVELKQRI YEQL FLTGKKVT KKS LTKYLIKNGYDKDIELSGID NEFHSNLKSHIDFEDYDNLSDEEVEQIILRITVFEDKQLLDYLNRE FVKLS EDERKQICSLSYKGWGNLSEMLLNGITVTDSNGVEVS VMDMLWNTNLNMQI LSKKYGYKAEIEHYNKEHEKT IYNREDLMDYLNIPPAQRRKVNQLITIVKSL KKTYGVPNKIFFKISREHQDDPKRT SSRKEQLKYLYKSLKSEDEKHLMKELD ELNDHELSDNKVLYFLQKGRCIYS GKKNLSRLRKS NYQNDIDYIYPLSAV NDRSMNNKVL TGIQENRADKYTFPVDSEIQKKMKGFWMELVLQGFMTKEY FRLSRENDFSKSELVS FIEREISDNQQSGRMIA SVLQYYFPESKIVFVKEKL ISSFKRDFHLISSYGHNLQA AKDAYITIVVGNVYHTKFTMDPAIYFKNHKR KDYDLNRLFLENI SRDQIAWESGPYGSIQTVRKEYAQNHIAVTKRVVEVKG GLFKQMPLKKGHEYP LKTNDPRFGNIAQYGGYTNVTGSYFVLVESMEKGKK RISLEYVPVYLHERLEDDPGHKLLKEYLV DHRKLNHP KILLAKVRKSNL LKI DGFYYRLNGRSGNALIILTNAVELIMDDWQTKTANKISGYMKRAIDKKARVY QNEFHIQELEQLYDFYLDKLNKGVYKNRKNNQAELIHNEKEQFME LKTEDQC VLLTEIKKLFVCS PMQADLT LIGGSKHTGMIAMSSNVTKADFAVIAEDPLGL RNKVIYSHKGEK
KvCas9 <i>Kandleria</i>	439	MSQNNNKIYNI GLDIGDASVGWAVVDEHYNLLKRHGKHMWGSRLFTQANTAV ERRSSRSTRRRYNKRREIRL LREIMEDMVLVDVPTFFIRLANVSFLDQEDK

<i>vitulina</i> NCBI Reference Sequence: WP_0315899 69.1 Wild type		KDYLKENYHSNYNLFIDKDFNDKTY YDKY PTIYHLRKKHLCEskekedprliY LALHHIVKY RGNFLYEGQKFSMDVSNIEDK MIDVLRQFNEI NLFEYVEDRKK IDEVLNVLKEPLS KKHKA EKA FALFDTTKDNKAAYKE LCAALAGNKFNVTKM LKEAELHDEDEKDISFKFS DATFDDAFVEKQPL LGDCVEFI DLLHDIYSWVE LQNILGSAHTSEPSISAAMIQRYEDHKND LKLLKDVI RKYL PKKY FEVFRDE KSKKNYCNYNHPSKT PVDE FYKY IKKL IEKI DDPDVKT I LNKI ELES FML KQNSRTNGAVPYQMQLDELNKILEN QSVY YSDLKDNE DKIRSILT FRIPYYF GPLNITKDRQFDWIIKKEGKENERI LPWNANEI VDVDKTADEFI KMRNFCT YFPDEPVMAKNSLTVSKYEVLNEINKLRINDHL IKRDMKDKMLHT LFM DHKS ISANAMKKWLKVNQYFSNTDDIKIEGFQKENACSTSLTPWIDFTKIFGKINE SNYDFIEKI IYDVTVFEDKKI LRRRLKKEYDLDEEKI KKLKLKYSGWSRLS KKLLSGIKTKYKDSTRT PETVLEVMERTNMNLMQVINDEKL GFKKTIDDANS TSVSGKFSAEVQELAGS PAIKRGIWQALLIVDE IKKIMKHEPAHVYIEFARN EDEKERKDS FVNQMLKL YKDYDFEDET EKEANKHLKGEDAKSKIRSERL KLY YTQMGKCMYTGKSLDIDRLDTYQVDHIVPQSLLKDDSIDNKVLVLSSENQRK LDDLVI PSSIRNKMYGFWEKL FNNKIIISP KKFYSLIKTEFNEKDQERFINRQ IVETRQITKHVAQIIDNHYENTKVVTVRADLSHQFREYHI YKNRDINDFHH AHDAYIATILGTYIGHRFESLDAKYIYGEYKRI FRNQKNKGKEMKKNNDGFI LNSMRNIYADKDTGEIVWDPNYIDRIKKCFYKDCFTKKLEENNGTFFNVNT VLPNDTNSDKDNTLATVPVNKYRSNVNKYGGFSGVNS FIVAIKGGKKKGKKV IEVNKLTGIPLMYKNADEEIKINYLKQAE DLEEVQIGKEILKNQLIEKDGGGL YYIVAPTEI INAKQLILNESQTKLVCEIYKAMKYKNYDNLDSEKI IDLYRL INKMELYYPEYRKQLVKKFEDRYEQLKVISIEEKCNI IKQI LATLHCNS SIG KIMYSDFKISTTIGRLNGRTISLDDISFIAESPTGMYSKKYKL
EfCas9 <i>Enterococcus faecalis</i> NCBI Reference Sequence: WP_0166310.44.1 Wild type	440	MRLFEEGHTAEDRRLKRTARRRISRRRNRLRYLQAFFEEAMTDLDENFFARL QESFLVPEDKKWHRHPI FAKLEDEVAYHETYPTIYHLRKKLADSSSEQADLRL IYLALAHIVKYRGHFLIEGKLSTENTSVKDDQFQFMVIYNQTFVNGESRLVS APLPESVLIEEELTEKASRTKKSEKVLQQFPQEKANGLFGQFLKLMVGNKAD FKKVFGLEEEAKITYASESYEEDLEGILAKVGDEYSDVFLAAKNVYDAVELS TILADSDKKSHAKLSSSMIVRFTEHQEDLKKFKRFIRENCPEYDNLFKNEQ KDGYAGYIAHAGKVSQKFYQYVKKIIQDIAGA EYFLEKIAQENFLRKQRTF DNGVI PHQIHLAE LQAI IHRQAAYYPFLKENQE KIEQLVTFRIPYVGP LSK GDASTFAWLKRQSEEP RPWNLQETV DLDQSATAFIERMTNFD TYLPSEKVL PKHSLLYEKF MVFNELTKISYTD DRGIKANFSGKEKEKIFDYLFKTRRKVKK KDIIQFYRNEYNT EIVT LSGLEEDQFNAS FSTYQDLLKCGLTRAE LDHP DNA EKLEDIIKILTI FEDRQRI RTQLST FKGQFSAEVLKKLERKH YTGWGRLSKK LINGIYDKE SGKT ILDYLVKDDGVS KHYNRNFMQLINDSQLSFKNAIQKAQS SEHEETLSETVNELAGS PAIKKGIYQSLKIVDELVAIMGYAPKRI VVEMARE NQTTSTGKRRSIQRLKIVEKAMAEI GSNLLKEQPTTNEQLRDTRLFLYYMQN GKDMYTGD ELSLHRLSHYDIDHIIPQSF MKDDSLDNLVLVGSTENRGKSDDV PSKEVVKDMKAYWEKLYAAGLISQRKFQRLTKGEQGGTLTEDKAHFIQRQLV ETRQITKNVAGILDQRYNAKS KEKKVQIITLKA SLTSQFRSIFGLYK VREVN DYHHGQDAY LNCVVATTLLKVYPNLAPEFVYGEY PKFQTFKENKATAKAIY TNLLRFFTEDEPRFTKDGEILWSNSYLKTIKKE LNYHQMNIVKKVEVQKGGF SKESI KPKGPSNKLI PVKNGLDPOKYGGFDSPVVAYTVLFTHEKGKKPLIKQ EILGITIMEKTRFEQNPILFLEEKGF LRPVLMKLPKYTLYEFPEGRRRLLA SAKEAQKGNQMVLP EHLTLTYHAKQCLLPNQSESLAYVEQHQP EFOEILER VVDFAEVHTLAKSKVQQIVKL FEANQTADVKEIAASF IQLMQFNAMGAPSTF KFFQKDIERARYTSIKEIFDATTIYQSPTGLYETRRKVVD
<i>Staphylococcus aureus</i> Cas9	441	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGAR RRLKRRRRRHRIQRVKKL LFDYNLLTDHSELSGINPYEARVKGLS QKLSEEEF SAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNS KALEEKYVAELQLER LKKDGEVRGSINRFKTS DYVKEAKQLLKVQKAYHQLDQSFI DTYI DLLETRR TY YEGPGE GSPFGWKDI KEWY EMLMGHCTYFPEELRSVKYAYNADLYNALND LNNLVITRDENEKLEYEKFQI IENVFKQKKKPTLKQIAKE ILVNEEDI KGY RVTSTGKPEFTNLKVYHDIKDITARKEI IENAE LLDQIAKILTIYQSSEDIQ EELTNLNLSELTQEEIEQISNLKGYTGTHNL SLKAINLILDELWHTNDNQIAI FNRLKLVPKKVDLSQKKEIPTTLVDDFILSPVVKRSFIQS I KVINAI IKKYG LPNDII IELAREKNSKDAQKMINEMQKRNRQTNERIEEII RTTGKENAKYLI EKIKLHDMQEGKCLYSLEAIPLEDL LNNP FNYEVDHIIPRSVSFDNSFNNKV LVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILN LAKGKGRI SKTKKEY

		LLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSIN GGFTSFLRRKWKFKKERNKGYKHAEDALI IANADFI FKEWKKLDKAKKVME NQMFEEKQAESMPEIETE QEYKEIFITPHQIKH IKDFKDYKYSHRVDKKNR ELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHD PQTYQKLKLIMEQYGDENPLYKYEEETGNYLT KYSKKDNGPVIKKIKY YGN KLNAHLDITDDYPNSRNKVVKLSLKPYPYRFVYLDNGVYKFVTVKNLDVI KKE NYEEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVI GVNNDLL NRIEVNMIDITYREYLENMNDKRPPRI IKTIASKTQS IKKYSTDI LGNLYEV KSKKH PQII KKG
<i>Geobacillus thermodenitri ficans Cas9</i>	442	MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRL ARSARRRLRRRKHRLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEA LDRKLNDELARILLHLAKRRGFRSNRKSERTNKENSTMLKHIEENQSILSS YRTVAEMVVKDPKFSLHKRNKEDNYTNTVARDDLEREIKLI FAKQREYGNIV CTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFEPKEKRAPKATYTFQSFT VWEHINKLRLVSPGGIRALTDERRLIYKQAFHKNKITFHDVRTLNLNLPDDT RFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVYKGGAASFRPIDFD TFGYALTMFKDDTDIRSYLRNEYEQNGKRMENTADKVYDEELIEELLNLSFS KFGHL SLKALRNI LPYMEQGEVYSTACERAGYFTGPKKKQKTVL LPNIPPI ANPVVMRALTQARKVVNAI IKKYGSPVSIHIELARELSQSFDERRKMQKEQE GNRKKNETAIRQLVEYGLTLNPTGLDIVKFKLWSEQNGKCAYSLOPIEIERL LEPGYTEVDHVI PYSRSLDDS YTNKVLVLTKENREKGNRTPAEYLGLSERW QQFETFVLTNKQFSKKKRDRLLRLHYDENEENE FKNRNLNDTRYISRFLANF IREHLKFADSDDKQKVYTVNGRITAHLSRWNFNKNREESNLHHAVDAAIVA CTTPSDIARVTAIFYQRREQNKELSKKTDPPQFPQWPHFADELQARLSKNPKE SIKALNLGNYDNEKLES LQPVFVSRMPKRSITGAHQETLRRYIGIDERSGK IQTVVKKKLSEIQLDKTHGFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQE PLYKPKNGELGPIIRTIKIIDTTNQVIPLNDGKTVA YNSNIVRVDVFEKDG KYYCVPIYTIDMMKGILPNKAI EPNKPYSWKEMTEDYTFRFSLYPNDLIRI EFPREKTIKTAVGEEIKIKDLFAYYQTI DSSNGGLSLVSHDNNFSLRSIGSR TLKRFEKYQVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL
<i>FnCas9 Fusobacterium nucleatum NCBI Reference Sequence: WP_0607989 84.1</i>	443	MKKQKFSDY YLGF DIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFEEAKTAA ERRVQRNSRRRLKRRKWRLNLEEI FSNEILKIDSNFFRRLKESSLWLEDKS SKEKFTLFNDNDYKDYDFYKQYPTIFHLRNELIKNPEKKDIRLVYLAHSIF KSRGHFLFEGQNLKEIKNFETLYNNLIAFLEDNGINKIIDKNNIEKLEKIVC DSKKG LKDKKEKFKEIFNSDKQLVAIFKLSVGS SVSLNDLFDTDEYKKGEVE KEKISFREQIYEDDKPI YYSILGEKIELLDIAKTFYDFMVLNNILADSQYIS EAKVKLYEEHKKDLKNLYIIRKYNKGNYDKLFKDKNENNY SAYI GLNKEKS KKEVI EKSRLKIDDLIKNIKGYLPKVEEIEEKDKAIFNKILNKIELKTI LPK QRISDNGTLPYQI HEAELEKI LENQSKYYDFLNYEENGII TKDKLMTFKFR IPYYVGPLNSYHKDKGGSWIVRKEEGKILPWNFEQKVDIEKSAEEFIKRM NKCTYLNGEDVIPKDTFLYSEYVILNELNKVQVNDEFLEENKRKIIDELFK ENKKVSEKKFKEYLLVKQIVDGTIELKGVKDSFNSNYISYIRFKDIFGEKLN LDIYKEISEKSILWKCLYGDCKIFEKKIKNEYGDILTDEIKKINTFKFNN WGRLEKLLTGIEFINLETGECYSSVMDALRRTNYNLMELSSSKFTLQESIN NENKEMNEASYRDLIEESYVPSLKR AIFQTLKIYEEIRKITGRVPKVFIE MARGGDESMKNKIPARQEQ LKLYDSCGNDIANFSIDIKEMKNSLISYDNN SLRQKKLYLYYLQFGKCMYTGREIDLRLQNNNDTYDIDHI YPRS KVIKDDS FDNLVLVLKNENA EKSNEYPVKKEIQEKMKS FWRFLKEKNFISDEKYKRLTG KDDFELRGFMARQLVNVROTTEVKGKILQQIEPEIKIVYSKAEIASSFREMF DFIKVRELNDTHHAKDAYLNI VAGNVYNTKFTEKPYRYLQEI KENYDVKKIY NYDIKNAWDKENSLEIVKKNMEKNTVNITRFI KEKKGQLFDLNP I KKGETS EIIISIKPKVYNGKDDKLNKEYGYKSLNPAYFLYVEHKEKNKRIKSFERVNL VDVNNIKDEKSLVKYLI ENKKLVEPRVIKKVYKQVI LINDYPYSIVTLDSN KLMDFENLKLFL ENKYEKILKNVI KFLEDNQGKSEENYKF IYLLKKDRYEK NETLESVKDRYNLEFNEMYDKFLEKLD SKDYKNYMNKKYQELL DVKEKFIK LNLFDKAFT LKSF LDFLNRKTMADFSKVG LTKYLGKIQKIS SNVLSKNELYL LEESVTGLFVKKIKL
<i>SpCas9 Streptococcus pyogenes wild type</i>	444	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYT RRRKNRICY LQEI FSNEMAKVDDSFHRL EES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENL I AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE

Encoded product of SWBC2D7W0 14		DAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDI LRVNTEI TKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHA ILRRQEDFYPFLLKDNREKIEKILTFRIPYYVG PLARGNSRFAWMTRK SEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTV YNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEIS GVEDRFNASLGT YHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAH LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILD FLKSDGFANRNFQMQLI HDDS LTFKEDIQKAQVSGQDLSLHEHIANLA GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR LSDYDVDHIVPQS FLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKMKNYWR QLLNAKLITQRKFDNLT KAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAH DAY LNAVVG TALIKKY PKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS NIMNFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQV NIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGSKKLSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKH RDK PIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDSPKKRKVS SDYKDHDGDYKD HDI DYKDDDDKAAG
SpCas9 <i>Streptococcus pyogenes</i> M1 SwissProt Accession No. Q99ZW2 Wildtype	445	MDKKY SIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHS IKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEI FSNEMAKVDDSF FHRLEES FLVEEDKKHERHP IFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENL I AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDI LRVNTEI TKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHA ILRRQEDFYPFLLKDNREKIEKILTFRIPYYVG PLARGNSRFAWMTRK SEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTV YNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEIS GVEDRFNASLGT YHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAH LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILD FLKSDGFANRNFQMQLI HDDS LTFKEDIQKAQVSGQDLSLHEHIANLA GSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGS QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR LSDYDVDHIVPQS FLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKMKNYWR QLLNAKLITQRKFDNLT KAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAH DAY LNAVVG TALIKKY PKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS NIMNFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQV NIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGSKKLSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKH RDK PIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD
SpCas9 <i>Streptococcus pyogenes</i> MGAS1882 wildtype NC_017053.1	446	MDKKY SIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHS IKKNLIGALL FGSGETAETRLKRTARRRYTRRKNRICYLQEI FSNEMAKVDDSF FHRLEES FLVEEDKKHERHP IFGNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQIYNQLFEENPINASRVDA KAILSARLSKSRRLENL I AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDI LRVNTEI TKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHA ILRRQEDFYPFLLKDNREKIEKILTFRIPYYVG PLARGNSRFAWMTRK SEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTV YNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEIS GVEDRFNASLGA YHDL LKIIKDKDFLDNEENEDILEDIVLTLT

	LFEDRGMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG KTILDFLKS DGFANRNFQMQLI HDDS LTFKEDIQKAQVSGQGHSLEQIANLA GSPAI KKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERM KRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVDHIVPQSFIKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKMKMKNYWRQ LLNAKLITQRKFDNLTKAERGGGLSELDKAGFIKRQLVETROITKHVAQILDS RMNTKYDENDKLI REVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSN IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVN IVKKEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIK LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSF EDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVL SAYNKH RDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SIT GLYETRIDLSQLGGD
--	---

Polymerase

[0251] In some cases, the payload in the present disclosure comprises a polymerase (e.g., reverse transcriptase). A polymerase can comprise a natural or an engineered domain. A polymerase can be capable of synthesizing nucleic acids. A polymerase can be a DNA polymerase or an RNA polymerase. In some cases, a polymerase is a reverse transcriptase. A reverse transcriptase can synthesize DNA from deoxyribonucleotides. In some cases, a reverse transcriptase adds deoxyribonucleotides to the 3' end of a nucleic acid primer to synthesize DNA. In some cases, a reverse transcriptase uses an RNA template and uses base-pairing interactions to synthesize a DNA strand that is complementary to the RNA template. The reverse transcriptase domain can be a reverse transcriptase from any organism, phage, virus, or an engineered or mutated variant. The reverse transcriptase domain can be a reverse transcriptase derived from or sharing structural or sequencing similarity to a reverse transcriptase in a CRISPR system. The reverse transcriptase can be an M-MLV or HIV reverse transcriptase. The reverse transcriptase can be a human LINE-1 reverse transcriptase or a group II intron reverse transcriptase. The reverse transcriptase can be a human endogenous retrovirus reverse transcriptase.

Transposase Domain

[0252] In some cases, a nucleic-acid modifying effector or a nucleic acid-binding moiety comprises a transposase domain. A transposase domain can be a natural or an engineered domain. A transposase domain can be capable of aiding the translocation of a transposable element, a nucleic acid sequence that can change its position within a genome. In some cases, a transposase domain comprises a TnsA, a TnsB, a TnsC, or a TnsD domain. In some cases, a transposase domain comprises a TniQ domain. In some cases, a transposase domain is derived from or shares sequence or structural similarity with a transposase in a CRISPR system (e.g., a CRISPR-associated transposase). In some cases, a transposase domain is derived from or share sequence or structural similarity with a transposase domain from a type I CRISPR-associated transposon

(CAST) system. In some cases, transposase domain is derived from or share sequence or structural similarity with a transposase domain from a type V CRISPR-associated transposon (CAST) system. A transposase domain can be capable of binding to a guidable polypeptide domain. In some cases, a transposase domain is coupled to a guidable polypeptide domain. In some cases, a transposase domain is capable of binding to a type I CRISPR-Cas domain (e.g., a Cascade domain, a Cas8 domain, or a Cas5 domain). In some cases, a transposase domain is capable of binding to a type V CRISPR-Cas domain (e.g., a Cas12 domain). In some cases, a transposase domain is capable of mediating targeted insertion of a nucleic acid into a target nucleic acid. In some cases, a transposase domain is capable of mediating targeted insertion of a nucleic acid that is at least 5 kb, at least 6 kb, at least 7 kb, at least 8kb, at least 9kb, at least 10kb, at least 11kb, at least 12kb, at least 13kb, at least 14kb, or at least 15 kb into a target nucleic acid.

Transcriptional Regulatory Element

[0253] In some cases, the payload comprises a transcriptional regulatory domain. A transcriptional regulatory domain can be a natural or an engineered domain. A transcriptional regulatory domain can be capable of regulating, activating, or inhibiting gene expression. For example, a transcriptional repressor can silence gene expression by binding to the promoter of a gene. A transcriptional activator can bind to enhancers or regulatory elements to activate expression of a gene. A transcriptional regulatory domain can comprise a transcription factor. A transcriptional regulatory domain can comprise a transcriptional activation domain or a transcriptional repression domain. For example, a transcriptional activation domain can be or comprise a CAP domain, a VP64 domain, a p65 domain, an Rta domain, a synergistic activation mediator (SAM) domain, a SunTag domain, a VPR domain, a DNA demethylase domain, a histone methyltransferase domain, a histone acetyltransferase domain, or a histone demethylase domain. A transcriptional repression domain can be or comprise a dCas9 domain, a KRAB domain, a Sin3 interacting domain (SID), or a MePC2 domain, a DNA methyltransferase domain, a histone deacetylase domain, a histone methyltransferase domain, or a histone demethylase domain. In some cases, a transcriptional regulatory domain comprises an epigenetic modifying effector domain. For example, an epigenetic modifying effector can be a DNA methyltransferase, a DNA demethylase, a histone methyltransferase, a histone demethylase, a histone acetyltransferase, or a histone deacetylase domain. A DNA methyltransferase domain can be capable of methylating a nucleic acid. A DNA demethylase domain can be capable of demethylating a nucleic acid. A histone methyltransferase domain can be capable of methylating a histone. A histone demethylase domain can be capable of demethylating a histone. A histone acetyltransferase domain can be capable of adding an acetyl group to a histone. A histone deacetylase domain can be capable of removing an acetyl group from a histone.

Zinc Finger Domain

[0254] In some embodiments, the payload comprises a zinc finger domain. A zinc finger domain can be a natural or an engineered domain. A zinc finger domain can bind to a specific DNA sequence in a target nucleic acid. A zinc finger domain can comprise from 1 to 10, from 2 to 10, from 3 to 10, from 4 to 10, from 5 to 10, from 6 to 10, from 7 to 10, from 8 to 10, from 9 to 10 zinc fingers, from 1 to 8, from 2 to 8, from 3 to 8, from 4 to 8, from 5 to 8, from 6 to 8, from 7 to 8, from 8 to 8, from 9 to 8 zinc fingers. In some cases, a zinc finger domain comprises a two-handed zinc finger domain. A two-handed zinc finger domain can comprise two clusters of zinc finger domains that are separated by intervening amino acids. A two-handed zinc finger domain can bind to two noncontiguous target DNA sequences. In some cases, the spacing between the two noncontiguous target sequences comprises from 1 to 15, from 1 to 12, from 1 to 10, from 1 to 8, or from 1 to 5 nucleotides. For example, a two-handed type of zinc finger binding protein can be SIP1. A cluster of zinc finger domains in a two-handed zinc finger domain can be capable of binding to a unique target nucleic acid sequence.

TALE Domain

[0255] In some embodiments, the payload comprises a TALE domain. A TALE domain can be a natural or an engineered domain. A TALE domain can bind to a specific DNA sequence. A TALE domain can comprise one or more effector domains. A TALE effector domain can comprise a central repeat domain comprising tandem repeats. A tandem repeat can comprise repeat variable residues (RVD). One or more RVDs can detect a specific DNA base. Different TALE effector domains may have a different number of repeats and a different order of their repeats. The C-terminal repeat is usually shorter in length (e.g., about 20 amino acids). Sequential repeats and their RVDs can recognize sequential DNA bases.

[0256] A TALE domain described herein can be derived from a TALE effector from a bacterial species. The TALE domain can be engineered to target a given nucleic acid sequence based on their DNA base specificities. The TALE domain can be engineered to remove or add a TALE effector domain. In some cases, the TALE domain corresponds to a perfect match to a nucleic acid target sequence. In some cases, the TALE domain of an epigenetic effector corresponds to one or more mismatches to a target base in the target nucleic acid.

Fusion Protein

[0257] In some cases, the payload in the present disclosure comprises a fusion protein. A fusion protein can comprise two or more polypeptide domains of any of the polypeptide domains described elsewhere herein. A fusion protein can be a natural or an engineered fusion protein. In some cases, the two or more polypeptide domains are coupled together. The two or more polypeptide domains can be coupled together directly or coupled together indirectly. For example,

a first polypeptide domain can be coupled directly to a second polypeptide domain. Alternatively, the first polypeptide domain can be coupled indirectly to the second polypeptide domain by coupling with a third polypeptide domain that is coupled directly to the second polypeptide domain. In some cases, a first polypeptide domain is coupled to the N-terminus of a second polypeptide domain. In some cases, a first polypeptide domain is coupled to the C-terminus of a second polypeptide domain. In some cases, a first polypeptide domain is coupled to an internal component of a second polypeptide domain. In some cases, the two or more polypeptide domains are covalently linked. In some cases, the two or more polypeptide domains are noncovalently linked. In some cases, the two or more polypeptide domains are coupled together by a linker. For example, a linker may be a peptide linker. A linker can be a rigid linker, which helps maintain a fixed distance between the polypeptide domains that it links. A linker can be a flexible linker, which can allow some flexibility in movement of one polypeptide domain relative to the other polypeptide domain that it is linked to. In some cases, a linker is a cleavable linker. For example, a cleavable linker can comprise a disulfide bond. Alternatively, a cleavable linker can be an enzymatic cleavable linker, e.g., a linker comprising a protease cleavage site.

[0258] The present disclosure provides fusion proteins comprising a guidable polypeptide domain (e.g., a CRISPR domain). A fusion protein comprising a guidable polypeptide domain (e.g., a CRISPR domain) can comprise one or more of a FokI domain, a deaminase domain, a reverse transcriptase domain, an RNA binding domain, a transcriptional regulatory domain, a plasma membrane recruitment domain, a transmembrane domain, a signaling domain, a receptor domain, a packaging domain, or a targeting domain.

[0259] In some cases, the present disclosure provides a fusion protein comprising a guidable polypeptide domain (e.g., a CRISPR domain) coupled to a deaminase domain. A guidable polypeptide domain (e.g., a CRISPR domain) coupled to a deaminase domain can be used for base editing. A base editor can be capable of editing a nucleic acid sequence in a target nucleic acid molecule. A base editor can be capable of enabling the generation of base conversions or point mutations in a target nucleic acid. For example, a cytosine base editor can comprise a guidable polypeptide domain (e.g., a CRISPR domain) and a cytidine deaminase domain. An adenine base editor can comprise CRISPR domain and an adenosine deaminase domain. In some cases, a base editor enables the conversion of C to G, A to I, or C to U. A cytosine base editor can be capable of enabling the conversion of a C-G base pair to a T-A base pair. A glycosylase base editor can be capable of enabling the conversion of a G-C base pair to a C-G base pair or a G-T base pair. An adenine base editor can be capable of enabling the conversion of an A-T base pair to a G-C base pair. In some cases, a base editor comprises a catalytically inactive guidable polypeptide domain (e.g., a CRISPR domain) (e.g., dCas9, dCas12a, or dCas13b). In other cases, the base

editor comprises a guidable polypeptide nickase domain (e.g., nCas9). In some cases, the base editor enables a base pair conversion without introducing a double-stranded break. In some cases, the base editor enables base pair conversions in a target window. In some cases, the base editor comprises a targeting window of from 1 to 20 bases, from 1 to 19 bases, from 1 to 18 bases, from 1 to 17 bases, from 1 to 16 bases, from 1 to 15 bases, from 1 to 14 bases, from 1 to 13 bases, from 1 to 12 bases, from 1 to 11 bases, from 1 to 10 bases, from 1 to 9 bases, from 1 to 8 bases, from 1 to 7 bases, from 1 to 6 bases, from 1 to 5 bases, from 1 to 4 bases, from 1 to 3 bases, or from 1 to 2 bases. In some cases, a base editor has a targeting window of from 3 to 10 bases, from 3 to 9 bases, from 3 to 8 bases, from 3 to 7 bases, from 3 to 6 bases, from 3 to 5 bases, or from 3 to 4 bases. In some cases, the guidable polypeptide domain (e.g., a CRISPR domain) is coupled to the N-terminus of a deaminase domain. In some cases, the guidable polypeptide domain (e.g., a CRISPR domain) is coupled to the C-terminus of a deaminase domain. In some cases, the guidable polypeptide domain (e.g., a CRISPR domain) is coupled to an internal component of a deaminase domain.

[0260] In some cases, the fusion protein comprises a guidable polypeptide domain (e.g., a CRISPR domain) coupled to a reverse transcriptase domain. A guidable polypeptide domain (e.g., a CRISPR domain) coupled to a reverse transcriptase can be capable of enabling prime editing. A prime editor can be capable of editing a nucleic acid sequence in a target nucleic acid molecule. A prime editor can be capable of mediating insertion or deletion of a nucleic acid sequence in a target nucleic acid molecule. In some cases, the prime editor enables a sequence insertion or sequence deletion without introducing a double-stranded break. In some cases, the prime editor introduces a nick at the target site. The prime editor can enable insertion of a template sequence in a target nucleic acid molecule. The template sequence can comprise the desired edit. In some cases, a prime editor reverse transcribes a template sequence to synthesize a complementary strand. In some cases, the synthesized complementary strand is inserted in the target nucleic acid molecule. In some cases, the prime editor uses a primer to carry out reverse transcription. The prime editor can install nucleotides to the 3' end of a primer strand. In some cases, a primer strand is generated by nicking the target nucleic acid molecule. In some cases, nicking a strand of the target nucleic acid molecule produces a flap with a 3' OH group. In some cases, the prime editor uses the flap with the 3' OH group as the primer to carry out reverse transcription. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to the N-terminus of a reverse transcriptase domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to the C-terminus of a reverse transcriptase domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to an internal component of a reverse transcriptase domain.

[0261] In some cases, the fusion protein comprises a guidable polypeptide domain (e.g., a CRISPR domain) coupled to a transcriptional regulatory domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to a transcriptional regulatory domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to a transcriptional activation domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to a transcriptional repression domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to a transcriptional regulatory domain. A guidable polypeptide domain (e.g., a CRISPR domain) coupled to a transcriptional regulatory domain can be capable of enabling CRISPR interference (CRISPRi) or CRISPR activation (CRISPRa). In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) can be coupled to a transcriptional regulatory domain such as P3000 or DNMT3. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to the C-terminus of a transcriptional regulatory domain. In some cases, a guidable polypeptide domain (e.g., a CRISPR domain) is coupled to an internal component of a transcriptional regulatory domain. Any of the payloads described herein can further comprise a plasma membrane recruitment domain, transmembrane domain, a signaling domain, a receptor domain, a packaging domain, or a targeting domain. Any of payloads described herein can comprise or be engineered to comprise a protein tag, a peptide tag, or small molecule tag. For example, a payload can comprise a small nuclear localization signal (NLS), a nuclear export signal (NES), a cell penetrating peptide (CPP), a mitochondria penetrating peptide (MPP), a solubility tag, or a fluorescent tag.

Base Editor

[0262] In some cases, the payload to be delivered by the lipid containing particles of the present disclosure comprises a nucleobase editor (also termed as “base editor”) or one or more components of a nucleobase editing (also termed as “base editing”) complex.

[0263] The term “base editor (BE),” or “nucleobase editor (NBE),” as used herein, can refer to an agent comprising a polypeptide that is capable of making a modification to a base (e.g., A, T, C, G, or U) within a nucleic acid sequence (e.g., DNA or RNA). In some embodiments, the base editor is capable of deaminating a base within a nucleic acid. In some embodiments, the base editor is capable of deaminating a base within a DNA molecule. In some embodiments, the base editor is capable of deaminating an adenosine (A) in DNA. In some embodiments, the base editor is capable of deaminating a cytosine (C) in DNA. In some embodiments, the base editor is capable of converting a guanine (G) in DNA through a glycosylase.

Deaminase Domain

[0264] In some cases, the payload in the present disclosure comprises a deaminase domain. The deaminase domain can be a natural or an engineered domain. A deaminase domain can be capable

of carrying out deamination reactions in DNA. A deaminase domain can be capable of enabling the generation of base conversions or point mutations in a target nucleic acid. For example, a deaminase domain can be a cytidine deaminase domain or an adenosine deaminase domain. A cytidine deaminase domain can be capable of converting cytosine to uracil. A cytidine deaminase domain can be capable of enabling the conversion of a C-G base pair to a T-A base pair. For example, a cytidine deaminase can be or comprise a APOBEC1 cytidine deaminase. An adenosine deaminase domain can be capable of converting an adenosine to hypoxanthine. An adenosine deaminase domain can be capable of converting an adenosine to an inosine. An adenosine deaminase can comprise TadA or a TadA mutant. In some embodiments, TadA comprises a monomer. In some embodiments, TadA comprises a heterodimer comprising a wildtype TadA and a mutated Tad A. In some embodiments, TadA comprises a homodimer comprising two wildtype TadA domains or two mutated TadA domains. An adenosine deaminase domain can be capable of enabling the conversion of an A-T base pair to a G-C base pair. A deaminase domain can be a mutated variant. In some cases, a deaminase domain enables the conversion of C to G, A to I, or C to U.

[0265] In some cases, the payload in the present disclosure comprises a glycosylase domain. The glycosylase domain can be a natural or an engineered domain. A glycosylase-based guanine base editor can be designed to remove G, and the AP site generated is repaired by translesion synthesis and/or DNA replication, leading to G-to-C or G-to-T conversion. A glycosylase domain can be capable of enabling the generation of base conversions or point mutations in a target nucleic acid. For example, a glycosylase domain can be a guanine glycosylase domain. Examples of glycosylase base edits can be found in Sun N, et al., *Mol Ther.* 2022 Jul 6;30(7):2452-2463 and Huawei Tong, et al., *National Science Review*, Volume 10, Issue 8, August 2023, each of which is incorporated in its entirety herein.

[0266] In some cases, the base editor disclosed herein comprises a deaminase or a functional domain thereof ("deaminase domain") that catalyzes deamination reaction.

[0267] The term "deaminase" or "deaminase domain," as used herein, refers to a protein or enzyme that catalyzes a deamination reaction. In some embodiments, the deaminase or deaminase domain is an adenosine deaminase, catalyzing the deamination of adenosine, converting it to the nucleoside hypoxanthine. In some embodiments, the deaminase or deaminase domain is a cytidine deaminase, catalyzing the hydrolytic deamination of cytidine or deoxycytidine to uridine or deoxyuridine, respectively. In some embodiments, the deaminase or deaminase domain is a cytidine deaminase domain, catalyzing the hydrolytic deamination of cytosine to uracil. In some embodiments, the deaminase or deaminase domain is a naturally-occurring deaminase from an organism, such as a human, chimpanzee, gorilla, monkey, cow, dog, rat, or mouse. In some

embodiments, the deaminase or deaminase domain is a variant of a naturally -occurring deaminase from an organism, that does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally -occurring deaminase from an organism.

[0268] As used herein, an “adenosine deaminase” is an enzyme that catalyzes the deamination of adenosine, converting it to the nucleoside hypoxanthine. Under standard Watson-Crick hydrogen bond pairing, an adenosine base hydrogen bonds to a thymine base (or a uracil in case of RNA). When adenine is converted to hypoxanthine, the hypoxanthine undergoes hydrogen bond pairing with cytosine. Thus, a conversion of “A” to hypoxanthine by adenosine deaminase will cause the insertion of “C” instead of a “T” during cellular repair and/or replication processes. Since the cytosine “C” pairs with guanine “G”, the adenosine deaminase in coordination with DNA replication causes the conversion of an A•T pairing to a C•G pairing in the double-stranded DNA molecule.

[0269] In some embodiments, the base editor is a chimeric protein comprising a nucleic acid programmable R/DNA binding protein (napR/DNAbp) fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase) domain. The term “nucleic acid programmable D/RNA binding protein (napR/DNAbp)” refers to any protein that can associate (*e.g.*, form a complex) with one or more nucleic acid molecules (*i.e.*, which can broadly be referred to as a “napR/DNAbp-programming nucleic acid molecule” and includes, for example, guide RNA in the case of Cas systems) which direct or otherwise program the protein to localize to a specific target nucleotide sequence (*e.g.*, a gene locus of a genome, or an RNA molecule) that is complementary to the one or more nucleic acid molecules (or a portion or region thereof) associated with the protein, thereby causing the protein to bind to the nucleotide sequence at the specific target site. This term napR/DNAbp embraces CRISPR Cas9 proteins, as well as Cas9 equivalents, homologs, orthologs, or paralogs, whether naturally occurring or non-naturally occurring (*e.g.*, engineered or recombinant), and can include a Cas9 equivalent from any type of CRISPR system (*e.g.*, type II, V, VI), including Cpf1 (a type-V CRISPR-Cas systems), C2c1 (a type V CRISPR-Cas system), C2c2 (a type VI CRISPR-Cas system) and C2c3 (a type V CRISPR-Cas system). Further Cas-equivalents are described in Makarova et al., “C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector,” *Science* 2016; 353(6299), the contents of which are incorporated herein by reference. However, the nucleic acid programmable R/DNA binding protein (napR/DNAbp) that can be used in connection with this disclosure are not limited to CRISPR-Cas systems. The present disclosure embraces any such

programmable protein, such as the Argonaute protein from *Natronobacterium gregoryi* (NgAgo) which can also be used for DNA-guided genome editing. NgAgo-guide DNA system does not require a PAM sequence or guide RNA molecules, which means genome editing can be performed simply by the expression of generic NgAgo protein and introduction of synthetic oligonucleotides on any genomic sequence. See Gao F, Shen X Z, Jiang F, Wu Y, Han C. DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute. *Nat Biotechnol* 2016; 34(7):768-73, which is incorporated herein by reference.

[0270] In some cases, the napR/DNAbp is derived from a nuclease disclosed herein, such as, Cas9 (*e.g.*, dCas9 and nCas9), CasX, CasY, Cas14, Cpf1, C2cl, C2c2, C2c3, Argonaute protein, or a variant thereof. In some embodiments, the base editor comprises a Cas9 (*e.g.*, dCas9 and nCas9), CasX, CasY, Cpf1, C2cl, C2c2, C2c3, or Argonaute protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a Cas9 nickase (nCas9) fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a CasX protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a nuclease-inactive Cas9 (dCas9) fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a CasY protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a Cas14 protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a Cpf1 protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a C2cl protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a C2c2 protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises a C2c3 protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase). In some embodiments, the base editor comprises an Argonaute protein fused to a deaminase (*e.g.*, cytidine deaminase or adenosine deaminase).

[0271] In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenosine. In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenosine in a deoxyadenosine residue of DNA. The adenosine deaminase can be derived from any suitable organism (*e.g.*, *E. coli*). In some embodiments, the adenosine deaminase is a naturally-occurring adenosine deaminase that includes one or more mutations corresponding to any of the mutations provided herein (*e.g.*, mutations in ecTadA). One of skill in the art will be able to identify the corresponding residue in any homologous protein and in the respective encoding nucleic acid by methods well known in the art, *e.g.*, by sequence alignment

and determination of homologous residues. Accordingly, one of skill in the art would be able to generate mutations in any naturally-occurring adenosine deaminase (*e.g.*, having homology to ecTadA) that corresponds to any of the mutations described herein, *e.g.*, any of the mutations identified in ecTadA. In some embodiments, the adenosine deaminase is from a prokaryote. In some embodiments, the adenosine deaminase is from a bacterium. In some embodiments, the adenosine deaminase is from *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Shewanella putrefaciens*, *Haemophilus influenzae*, *Caulobacter crescentus*, or *Bacillus subtilis*. In some embodiments, the adenosine deaminase is from *E. coli*.

[0272] In some cases, the deaminase domain of the base editor disclosed herein is derived from a cytidine deaminase. In some cases, the cytidine deaminase domain is derived from the apolipoprotein B mRNA-editing complex (APOBEC) family deaminase, such as APOBEC1 deaminase, APOBEC2 deaminase, APOBEC3A deaminase, APOBEC3B deaminase, APOBEC3C deaminase, APOBEC3D deaminase, APOBEC3F deaminase, APOBEC3G deaminase, or APOBEC3H deaminase. In some cases, the cytidine deaminase is a modification of an APOBEC family deaminase. In some cases, the cytidine deaminase is an evolved derivative of an APOBEC family deaminase.

[0273] In some embodiments, the base editor comprises BE1, BE2, BE3, BE4, BE4max, or another base editor variant. In some embodiments, the base editor comprises BE4max (R33A) AUGI-hUNG complex (CGBE1).

[0274] In some embodiments, the base editor is fused to, or further comprises as part of a chimeric protein, an inhibitor of base excision repair, for example, a uracil glycosylase inhibitor (UGI) domain. In some embodiments, the base editor is fused to one, two, or three UGI domains. In some embodiments, the base editor is fused to one or more UGI domains. In some embodiments, a UGI domain reduces off target effects, specifically the conversion of C to G or C to A.

[0275] In some cases, the base editor disclosed herein is a chimeric protein that comprises a structure such as, NH₂-[deaminase domain]-[napR/DNAbp]-[UGI domain]-COOH; NH₂-[deaminase domain]-[napR/DNAbp]-[UGI]-[UGI]-COOH; NH₂-[deaminase domain]-[napR/DNAbp]-[UGI]-COOH; NH₂-[UGI]-[deaminase domain]-[napR/DNAbp]-COOH; NH₂-[deaminase domain]-[UGI]-[napR/DNAbp]-COOH; NH₂-[napR/DNAbp]-[UGI]-[deaminase domain]-COOH; or NH₂-[napR/DNAbp]-[deaminase domain]-[UGI]-COOH; wherein each instance of “-” comprises an optional linker.

[0276] In some cases, the base editor is fused to, or further comprises as part of a chimeric protein, a uracil binding protein (UBP). The term “uracil binding protein” or “UBP,” as used herein, refers to a protein that is capable of binding to uracil. In some embodiments, the uracil binding protein is a uracil modifying enzyme. In some embodiments, the uracil binding protein is a uracil base

excision enzyme. In some embodiments, the uracil binding protein is a uracil DNA glycosylase (UDG). In some embodiments, a uracil binding protein binds uracil with an affinity that is at least 1%, 2%, 3%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or at least 95% of the affinity that a wild type UDG (*e.g.*, a human UDG) binds to uracil. The term “base excision enzyme” or “BEE,” as used herein, refers to a protein that is capable of removing a base (*e.g.*, A, T, C, G, or U) from a nucleic acid molecule (*e.g.*, DNA or RNA). In some embodiments, a BEE is capable of removing a cytosine from DNA. In some embodiments, a BEE is capable of removing a thymine from DNA. Exemplary BEEs include, without limitation UDG Tyr147Ala, and UDG Asn204Asp as described in Sang et al., “A Unique Uracil-DNA binding protein of the uracil DNA glycosylase superfamily,” *Nucleic Acids Research*, Vol. 43, No. 17 2015; the entire contents of which are hereby incorporated by reference.

[0277] In some embodiments, the UBP is a uracil modifying enzyme. In some embodiments, the UBP is a uracil base excision enzyme. In some embodiments, the UBP is a uracil DNA glycosylase. In some embodiments, the UBP is any of the uracil binding proteins provided herein. For example, the UBP can be a UDG, a UdgX, a UdgX*, a UdgX_On, or a SMUG1. In some embodiments, the UBP comprises an amino acid sequence that is at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a uracil binding protein, a uracil base excision enzyme or a uracil DNA glycosylase (UDG) enzyme.

[0278] In some cases, the base editor is fused to, or comprises as a part of the chimeric protein, a nucleic acid polymerase domain (NAP). For instance, the nucleic acid polymerase domain is a eukaryotic nucleic acid polymerase domain. In some cases, the nucleic acid polymerase domain is a DNA polymerase domain. In some cases, the nucleic acid polymerase domain has translesion polymerase activity. In some cases, the nucleic acid polymerase domain is a translesion DNA polymerase. In some cases, the nucleic acid polymerase domain is from Rev7, Rev1 complex, polymerase iota, polymerase kappa, and polymerase eta. In some cases, the nucleic acid polymerase domain is selected from the group of eukaryotic polymerases consisting of alpha, beta, gamma, delta, epsilon, gamma, eta, iota, kappa, lambda, mu, and nu.

[0279] In some cases, the base editor disclosed herein is a chimeric protein that comprises a structure such as, NH₂-[deaminase domain]-[napR/DNAbp domain]-[UBP]-[NAP]-COOH; NH₂-[deaminase domain]-[napR/DNAbp]-[NAP]-[UBP]-COOH; NH₂-[deaminase domain]-[NAP]-[napR/DNAbp]-[UBP]-COOH; or NH₂-[NAP]-[deaminase domain]-[napR/DNAbp]-[UBP]-COOH; wherein each instance of “-” comprises an optional linker.

[0280] In some cases, the base editor disclosed herein is complexed with a napR/DNAbp-programming nucleic acid molecule. In some cases, the base editing system disclose herein comprises a base editor and a napR/DNAbp-programming nucleic acid molecule, *e.g.*, the base

editor complexed with the napR/DNAbp-programming nucleic acid molecule. In some cases, the lipid containing particles of the present disclosure deliver a base editing system that comprises both a base editor and a napR/DNAbp-programming nucleic acid molecule, *e.g.*, the base editor complexed with the napR/DNAbp-programming nucleic acid molecule. In some cases, a base editor is delivered separately from the napR/DNAbp-programming nucleic acid molecule through lipid containing particles disclosed herein, or together with other delivery methods, into a cell. **[0281]** The term “napR/DNAbp-programming nucleic acid molecule” or equivalently “guide sequence” refers the one or more nucleic acid molecules which associate with and direct or otherwise program a napR/DNAbp protein to localize to a specific target nucleotide sequence (*e.g.*, a gene locus of a genome) that is complementary to the one or more nucleic acid molecules (or a portion or region thereof) associated with the protein, thereby causing the napR/DNAbp protein to bind to the nucleotide sequence at the specific target site. An example is a guide RNA of a Cas protein of a CRISPR-Cas genome editing system.

[0282] Exemplary configurations, sequences, and mutations thereof for deaminase domains, napR/DNAbp domains, UGI domains, and whole base editor proteins, and exemplary configurations of a base editing system (*e.g.*, comprising both a base editor and a napR/DNAbp-programming nucleic acid molecule) that can be delivered by a lipid containing particle disclosed herein include those described in U.S. Patent Publication Nos. US20170121693, US20180073012, US20180312828, US20180170984, US2020010835, US2020172931, US20210230577, US20210198330, US20210277379, US2020399626, US2021371858, US2021380955, US2021277379, US2021301274, US20220127622, US20220313799, US20230055682, US20230159913, US20230086199, US20220127622, US20220313799, US20230279373, US20230055682, each of which is incorporated herein by reference in its entirety. Exemplary configurations, sequences, and mutations thereof for deaminase domains, napR/DNAbp domains, UGI domains, and whole base editor proteins, that can be delivered by a lipid containing particle disclosed herein also include those described in Komor AC et al. *Nature*. 2016 May 19;533(7603):420-4; Kim YB et al. *Nat Biotechnol*. 2017 Apr;35(4):371-376; Rees HA et al. *Nat Commun*. 2017 Jun 6;8:15790; Newby GA et al. *Mol Ther*. 2021 Nov 3;29(11):3107-3124; Huang TP et al. *Nat Protoc*. 2021 Feb;16(2):1089-1128; Lapinaite A et al. *Science*. 2020 Jul 31;369(6503):566-571; Anzalone AV et al. *Nat Biotechnol*. 2020 Jul;38(7):824-844; Rees HA et al. *Nat Rev Genet*. 2018 Dec;19(12):770-788; Koblan LW et al. *Nat Biotechnol*. 2018 Oct;36(9):843-846; and Gaudelli NM et al. *Nature*. 2017 Nov 23;551(7681):464-471; each of which is incorporated herein by reference in its entirety.

Primer Editor

[0283] In some cases, the lipid delivery particles disclosed herein is capable of delivering a payload, such as a prime editing system, or one or more components thereof, such as a ribonucleoprotein (RNP) complex, into a cell in vitro, ex vivo, or in vivo. In some embodiments, the prime editing system, or one or more components thereof, is within the inside cavity of the protein core of the lipid delivery particles disclosed herein.

[0284] Prime editing system is a 'search-and-replace' genome editing technology by which the genome of living organisms can be modified. The term "prime editing system" or "prime editor (PE)" refers the compositions involved in genome editing using target-primed reverse transcription (TPRT) describe herein, can comprise a nucleic acid-guided polypeptide, *e.g.*, nucleic acid-guided polypeptide, a nucleic acid polymerase, chimeric proteins (*e.g.*, comprising guidable polypeptide domain and reverse transcriptase), guide nucleic acid molecule (*e.g.*, guide RNAs), and complexes comprising fusion proteins and guide RNAs, as well as accessory elements, such as second strand nicking components and 5' endogenous DNA flap removal endonucleases (*e.g.*, FEN1) for helping to drive the prime editing process towards the edited product formation.

[0285] In some embodiments, the prime editing system disclosed herein comprises a ribonucleoprotein (RNP) complex. In some cases, the RNP complex comprises a prime editor and a guide nucleic acid molecule. In some cases, the prime editor is formed between one or more proteins and one or more polynucleotides. The prime editor can comprise a nucleic acid-guided polypeptide. The guidable polypeptide domain can comprise a nucleic acid-guided polypeptide, for example a nuclease (*e.g.*, a Cas protein). For instance, the prime editor can comprise a fusion protein, comprising a nucleic acid programmable R/DNA binding protein (*e.g.*, a nuclease, such as a Cas protein) and a nucleic acid polymerase (*e.g.*, a reverse transcriptase or any suitable DNA polymerase). In some cases, the nucleic acid polymerase is coupled to the nucleic acid-guided polypeptide. In some cases, the guide nucleic acid molecule can comprise a guide nucleic acid molecule, *e.g.*, a guide RNA. In some cases, the prime editor is operably linked to the guide nucleic acid molecule via a linker, forming the RNP complex. In some cases, the prime editor is directly linked to the guide nucleic acid molecule, forming the RNP complex.

[0286] In a specific instance, prime editing system comprises a fusion protein that comprises an engineered Cas9 nickase and a reverse transcriptase, and the fusion protein is paired with an engineered prime editing guide RNA (PEgRNA). In some cases, the PEgRNA can direct Cas9 to a target site within a host cell where the lipid delivery particles are delivered. In some cases, the peg RNA can encode the information for installing the desired edit. In some cases, the prime editing system can function through a multi-step process: 1) the Cas9 domain can bind and nick

the target genomic DNA site, which is specified by a spacer sequence in the PEgRNA; 2) the reverse transcriptase can use the nicked genomic DNA as a primer to initiate synthesis of an edited DNA strand using an engineered extension on the PEgRNA as a template for reverse transcription, which can generate a single-stranded 3' flap containing the edited DNA sequence; 3) cellular DNA repair mechanism can resolve the 3' flap intermediate by the displacement of a 5' flap species that occurs via invasion by the edited 3' flap, excision of the 5' flap containing the original DNA sequence, and ligation of the new 3' flap to incorporate the edited DNA strand, forming a heteroduplex of one edited and one unedited strand; and 4) cellular DNA repair mechanism can replace the unedited strand within the heteroduplex using the edited strand as a template for repair, which completes this editing process. In some embodiments, the prime editing machinery edits a target DNA molecule. In some embodiments, the prime editing machinery edits a target RNA molecule. Examples of targeting RNA molecules using prime editing are described in international patent application WO2021072328 and U.S. Patent Application number US20230357766, each of which is incorporated in its entirety. In other instances, a prime editing system is a multi-flap prime editing system that can simultaneously edit both DNA strands. For example, a dual-flap prime editing system comprises two PEgRNAs, which can be used to target opposite strands of a genomic site and direct the synthesis of two complementary 3' flaps containing edited DNA sequence. The pair of edited DNA strands (3' flaps) does not need to directly compete with 5' flaps in endogenous genomic DNA, as the complementary edited strand is available for hybridization instead. In this instance, both strands of the duplex are synthesized as edited DNA, the dual-flap prime editing system obviates the need for the replacement of the non-edited complementary DNA strand. Instead, cellular DNA repair machinery can only excise the paired 5' flaps (original genomic DNA) and ligate the paired 3' flaps into the locus.

[0287] In some embodiments, a prime editing system can be paired with a separate Cas9 nickase and a separate gRNA that nicks the DNA at a locus that is different than the locus targeted by the PEgRNA. In some embodiments, one or more prime editing systems can be paired, each targeting a different locus. In some cases, pairing of two prime editing systems, each of which targets a different locus on the same chromosome, can install large insertions, deletions, or modifications. In some cases, pairing of two prime editing systems, each of which targets a different locus, can install large structural modifications. In some embodiments, a prime editor can install up to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 modifications. In some embodiments, a prime editor can install 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more modifications. In some embodiments, a prime editor can install up to about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, or about 100 modifications. In some embodiments, a prime editor can

install more than about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, or about 100 modifications. In some cases, a prime editor can install more than 100 modifications. In some embodiments, more than one prime editor can be used to install mutations more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more modifications. In some embodiments, more than one prime editor can be used to install mutations more than about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, or about 100 modifications. In some embodiments, more than one prime editor can be used to install mutations more than about 1kb, 5kb, 10kb, 20kb, 30kb, 40kb, 50kb, 60kb, 70kb, 80kb, 90kb, or more.

[0288] Different variants of prime editors have been developed, such as prime editors (PE) PE1, PE2, PE3, PE4, or PE5, some of which are described in Liu, D. et al., *Nature* 2019, 576, 149-157 and Huang Z, Liu G. *Front Bioeng Biotechnol.* 2023;11:1039315, U.S. Patent Application numbers US20210292769, US20230090221, US2022078655, US20230220374, each of which is hereby incorporated by reference herein in its entirety. In some cases, the prime editor comprises a reverse transcriptase (RT) fused with Cas9 H 840A nickase (Cas9n (H840A)) and a prime-editing guide RNA (pegRNA). In some cases, the RT comprises an RNA-dependent DNA polymerase. In some cases, the RT comprises a protein derived from a retrovirus. In some embodiments, the RT comprises Moloney Murine Leukemia Virus (M-MMLV) RT. In some embodiments, the RT comprises a RT from HFV, LtrA, HERV-Kcon, Tel4c, Marathon, Gst-IIC, MA-INT5, or another RT ortholog. In some cases, the RT is modified, mutated, truncated, or evolved. In some cases, the RT comprises a full length RT protein. In some cases, the RT comprises a truncated RT. In some cases, the RT is fused to the Cas protein. In some cases, the RT is fused to the Cas protein at the N terminus to the Cas protein. In some cases, the RT is fused to the Cas protein at the C terminus of the Cas protein. In some cases, the RT is fused to the Cas protein as an inlaid fusion. In some cases, the RT is untethered to the Cas protein. Examples of prime editing architecture are described in Grünewald, J., et al., *Nat Biotechnol* 41, 337-343 (2023) and Gao Z, et al., *Mol Ther.* 2022; 30(9):2942-2951, each of which is incorporated herein in its entirety.

[0289] In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(wt)] and (b) a PEgRNA. In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)] and (b) a PEgRNA. In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus

structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]; (b) a PEgRNA; and (c) a nicking guide RNA that introduces a nick in the non-edited DNA strand. In some cases, the addition of nicking guide RNA increases the chances of the unedited strand to be repaired rather than the edited strand. In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]; (b) a PEgRNA; and (c) a nicking guide RNA that is designed with a spacer that matches only the edited strand but not the original allele before editing, so that the nicking guide RNA is not introduced until after the desired edit is installed. In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]; (b) a PEgRNA; and (c) evading specific DNA mismatch repair (MMR) protein, such as co-expression of a dominant negative MMR protein, such as MLH1dn (*e.g.*, MLH1 Δ 754-756). In some cases, the prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]; (b) a PEgRNA; (c) a nicking guide RNA that introduces a nick in the non-edited DNA strand; and (d) evading specific DNA mismatch repair (MMR) protein, such as co-expression of a dominant negative MMR protein, such as MLH1dn (*e.g.*, MLH1 Δ 754-756). Evading MMR protein, such as by co-expression of MMR protein MLH1dn can increase efficiency of prime editing, as described in International Publication No., WO2023102538 and *Chen et al.*, Cell Volume 184, Issue 22, 28 October 2021, Pages 5635-5652.e29, each of which is hereby incorporated by reference herein in its entirety. An exemplary sequence for MLH1dn (MLH1 Δ 754-756) is: MSFVAGVIRRLDETVVNRIAAGEVIQRPANA KEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTTITTKTADGKCAYRASYS DGK LKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEY GKILEVVG RYSVHNAGISFSVKKQGETVADVRTL PNASTVDNIRSIFGNAV SRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAA YLPKNTHPFLYLSLEISPQNV DVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDS DVE MVEDDSRKEMTA ACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVPNPQWALA QHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWT EEDGPKEGLAEYIVEFLKKKA EMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESLSKECAMFY SIRQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVF. In some cases, other strategies for evading

MMR protein can be adopted, such as installing silent mutations next to the desired edit or co-expressing an antibody targeting the MMR protein. In some cases, the foregoing prime editor comprises (a) a fusion protein having the following N-terminus to C-terminus structure: [bipartite NLSI-[Cas9(R221K)(N394K)(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)]-[bipartite NLS]-[NLS] instead. In some cases, the components in the foregoing prime editors are packaged in a single lipid delivery particle. In some cases, the components in the foregoing prime editors are packaged in two or more lipid deliver particles that are delivered to the recipient cell simultaneously. In some cases, the components in the foregoing prime editors are packaged in two or more lipid deliver particles that are delivered to the recipient cell sequentially.

[0290] In some cases, the prime editing system can comprise a flap endonuclease (*e.g.*, FEN1 or variant thereof) that is delivered as a part of the lipid delivery particle (*e.g.*, fused to a plasma membrane recruitment element as a chimeric protein). The flap endonuclease can comprise naturally occurring enzymes that process the removal of 5' flaps formed during cellular processes, including DNA replication. The flap endonuclease includes those described in Patel *et al.*, *Nucleic Acids Research*, 2012, 40(10): 4507-4519 and Tsutakawa *et al.*, *Cell*, 2011, 145(2): 198-211, each of which is incorporated herein by reference in its entirety.

[0291] Additional elements that can be delivered as a part of the prime editing system via the lipid delivery particles (*e.g.*, fused to the nucleic acid-guided polypeptide, or fused to plasma membrane recruitment element) described herein include inhibitor of base repair (*e.g.*, proteins that inhibit a nucleic acid repair enzyme, for example, a base excision repair enzyme), uracil glycosylase inhibitor domains (*e.g.*, protein that inhibits a uracil-DNA glycosylase base-excision repair enzyme), epitope tags, and reporter gene sequences, including those described in International Publication No. WO2023205744, which is incorporated herein by reference in its entirety.

Epigenetic Editor

[0292] In some cases, the payload to be delivered by the lipid containing particles of the present disclosure comprises an epigenetic editor or one or more components of an epigenetic editing complex (*e.g.*, comprising an epigenetic editor and a nucleic acid molecule that guides the epigenetic editor to bind and/or modify one or more specific target sequences).

[0293] In some cases, the epigenetic editor or epigenetic editing complex disclosed herein has epigenetic activities, such as, methyltransferase activity, demethylase activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity or glycosylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity,

ribosylation activity, deribosylation activity, myristoylation activity, remodelling activity, protease activity, oxidoreductase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, synthase activity, synthetase activity, or demyristoylation activity. In some cases, the epigenetic editor or epigenetic editing complex disclosed herein has a chromosome modification enzyme, or a functional domain that has the functional activity equivalent to a chromosome modification enzyme, such as a methylase, demethylase, acetylase, deacetylase, deaminase, phosphorylase, dephosphorylase, histone modifying enzyme, or nucleotide modifying enzyme. In some cases, the epigenetic editor or epigenetic editing complex disclosed herein has a histone modifying enzyme, or a functional domain that has the functional activity equivalent to a histone modifying enzyme. In some cases, the epigenetic editor or epigenetic editing complex disclosed herein has a nucleotide modifying enzyme, or a functional domain that has the functional activity equivalent to a nucleotide modifying enzyme.

[0294] In some embodiments, the epigenetic editor or epigenetic editing system comprises a protein domain that represses expression of the target gene. For example, the epigenetic editor or epigenetic editing system can comprise a functional domain derived from a zinc finger repressor protein. Sequences of exemplary functional domains of an epigenetic editor or epigenetic editing system that can reduce or silence target gene expression are provided can be found in PCT/US2021/030643 and Tycko *et al.* (Tycko J, DelRosso N, Hess GT, Aradhana, Banerjee A, Mukund A, Van MV, Ego BK, Yao D, Spees K, Suzuki P, Marinov GK, Kundaje A, Bassik MC, Bintu L. High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell*. 2020 Dec 23;183(7):2020-2035.e16. doi: 10.1016/j.cell.2020.11.024. Epub 2020 Dec 15. PMID: 33326746; PMCID: PMC8178797.), each of which is incorporated here by reference in its entirety.

[0295] In some embodiments, the epigenetic editor or epigenetic editing system makes an epigenetic modification at a target gene that activates expression of the target gene. In some embodiments, the epigenetic editor or epigenetic editing system modifies the chemical modification of DNA or histone residues associated with the DNA at a target gene harboring the target sequence, thereby activating or increasing expression of the target gene. In some embodiments, the epigenetic editor or epigenetic editing system comprises a DNA demethylase, a DNA dioxygenase, a DNA hydroxylase, or a histone demethylase domain.

Nucleic acids and polynucleotides

[0296] In some cases, the lipid delivery particle of the present disclosure comprises a payload comprising a nucleic acid. The nucleic acid as a payload can comprise or be composed of one or more nucleotides. Nucleotides are referred to by their commonly accepted single-letter codes: A represents adenine, C represents cytosine, G represents guanine, T represents thymine, U

represents uracil, I represents inosine. Unless otherwise indicated, nucleotide sequences are written from left to right in a 5' to 3' orientation. In some cases, the nucleic acid as a payload comprises a polynucleotide. The nucleic acid as a payload can comprise DNA or RNA. In some cases, the nucleic acid as a payload comprises or encodes a gene. The nucleic acid as a payload can comprise or encode any of the polynucleotides described elsewhere herein. The nucleic acid as a payload can be a vector encoding any of the polypeptide domains described elsewhere herein. In some cases, the nucleic acid as a payload is an engineered polynucleotide.

[0297] In some cases, the payload does not comprise a repair template. In some cases, the double stranded break is repaired through non-homologous end joining. In some cases, the payload comprises a repair template. The repair template can be double-stranded or single-stranded. The repair template can comprise a template sequence comprising a desired edit to be introduced in a target nucleic acid molecule. In some cases, the repair template is a homology-directed repair template. A homology-directed repair template can comprise a homology arm that is homologous to a sequence in the target nucleic acid. In some cases, the payload comprises a DNA-synthesis template comprising a DNA-synthesis template sequence. The DNA-synthesis template can comprise a desired edit to be introduced in a target nucleic acid molecule. The DNA-synthesis template can be a template for a DNA polymerase or a reverse transcriptase to carry out DNA synthesis. For example, in prime editing, a prime editor can use the DNA synthesis template sequence to synthesize a DNA strand that is complementary to the DNA synthesis template sequence. In some cases, the DNA strand is inserted into the target nucleic acid. In some cases, the nucleic acid comprising the DNA synthesis template sequence also comprises a primer-binding sequence. A primer-binding sequence can be complementary to a sequence in a primer strand to which a DNA polymerase or reverse transcriptase can add nucleotides. In some cases, the primer strand is part of a target nucleic acid molecule. A primer-binding sequence can be complementary to a sequence in the target nucleic acid.

[0298] In some cases, the payload comprises a double-stranded DNA containing a desired gene sequence to be inserted in the target nucleic acid molecule. In some cases, the double-stranded DNA is configured to couple to a transposase domain. In some cases, the payload is delivered in the same particle as the transposase domain. In some cases, the payload is delivered in a separate particle as the transposase domain.

[0299] In some cases, the payload comprises a polynucleotide that is configured to bind to a guidable polypeptide domain. In some cases, the polynucleotide directs a guidable polypeptide domain to a sequence in a target nucleic acid molecule. In some cases, the polynucleotide comprises a scaffold segment configured to bind to a guidable polypeptide domain (e.g., Cas9 or Cas12). In some cases, polynucleotide comprises a spacer sequence that is complementary to a

target sequence in the target nucleic acid molecule and is capable of hybridizing to the target sequence. The polynucleotide can be a natural molecule or an engineered or synthetic molecule. The polynucleotide can be derived or share sequence or structural similarities to CRISPR RNA (crRNA), a tracrRNA, or a scoutRNA encoded in a CRISPR system. In some cases, the polynucleotide is engineered to be a single RNA guide (sgRNA) comprising elements of the crRNA and the tracrRNA. In some cases, the polynucleotide comprises a scaffold segment and a spacer sequence. The scaffold segment can be configured to bind to a guidable polypeptide domain. The scaffold segment can be specific to a specific type of guidable polypeptide (e.g., Cas9 or Cas12). In some cases, the spacer sequence is programmed to be any sequence. In some cases, the spacer sequence is programmed to a sequence complementary to a target nucleic acid sequence.

[0300] In some cases, the payload comprises a polynucleotide that is a guide nucleic acid molecule for a prime editing system, *e.g.*, a prime editing guide RNA (PEgRNA). In some cases, the PEgRNA is capable of (i) identifying a target nucleotide sequence to be edited, and (ii) encoding new genetic information that replaces the targeted sequence. In some cases, a guide nucleic acid molecule for a prime editing system comprises two or more guide RNAs. In some cases, a guide nucleic acid molecule for a prime editing system comprises a nicking guide RNA. In some cases, a guide RNA comprises (A) a primer binding site, (B) a clamp segment, (C) a sequence encoding new genetic information that replaces the targeted sequence, (D) an aptamer, (E) spacer, or (F) scaffold, or any combinations thereof. In some cases, a guide RNA comprises a sequence encoding new genetic information that replaces the targeted sequence, a spacer, and scaffold. In some cases, a guide RNA comprises a spacer and scaffold. In some cases, the guide nucleic acid molecule is heterologous to the cell or host receiving the lipid delivery particle.

[0301] In some cases, the PEgRNA comprises an extended single guide RNA (sgRNA) containing a primer binding site (PBS) and a template sequence for nucleic acid polymerase (*e.g.*, reverse transcriptase or DNA polymerase). For example, a PEgRNA can comprise an architecture corresponding to 5'-[spacer]-[guide RNA core]-[extension arm]-3'. The spacer sequence can comprise about 20 nucleotides in length. The spacer sequence can bind to a protospacer in a target nucleic acid molecule. The spacer sequence can guide the nucleic acid-guided polypeptide (*e.g.*, Cas9) to the target nucleic acid molecule. The guide RNA core can be responsible for binding of the nucleic acid-guided polypeptide (*e.g.*, Cas9). The extension arm can comprise a primer binding site, an edit template, and a homology arm, in a 3' to 5' direction. The PEgRNA can further comprise, optionally, a 3' end modifier region, 5' end modifier region, a transcriptional signal at the 3' end. The PEgRNA can optionally comprise a secondary structure, such as, hairpins, stem/loops, toe loops, RNA-binding protein recruitment domains (*e.g.*, the MS2 aptamer which

recruits and binds to the MS2cp protein). In some cases, the PEgRNA comprises an aptamer and the prime editor further comprises an aptamer binding protein (*e.g.*, fused to Cas protein or reverse transcriptase). Guide RNAs including an aptamer include those described in International Publication No. WO2023205708, which is hereby incorporated herein by reference in its entirety. Homology arm can encode a portion of a resulting reverse transcriptase-encoded single strand DNA flap to be integrated into the target DNA site by replacing the endogenous strand. The portion of the single strand DNA flap encoded by the homology arm is complementary to the non-edited strand of the target DNA sequence, which facilitates the displacement of the endogenous strand and annealing of the single strand DNA flap in its place, thereby installing the edit. The edit template can comprise a sequence corresponding to new genetic information that replaces the targeted sequence, *i.e.*, a single strand RNA of the PEgRNA that codes for a complementary single strand DNA that is either the sense or the antisense strand of the new genetic information that replaces the targeted sequence and which is incorporated into the genomic DNA target locus through the prime editing process.

[0302] In some cases, during genome editing, the primer binding site allows the 3' end of the nicked DNA strand to hybridize to the PEgRNA, while the reverse transcriptase template serves as a template for the synthesis of edited genetic information. A prime editing system can allow DNA synthesis based on the reverse transcriptase template at a nick site a single 3' flap, which becomes integrated into a target nucleic acid on the same strand. In other embodiments, a prime editing system can be a multi-flap prime editing system that generate pairs or multiple pairs of 3' flaps on different strands, which form duplexes comprising desired edits and which become incorporated into target nucleic acid molecules, *e.g.*, at specific loci or edit sites in a genome. In some embodiments, the pairs or multiple pairs of 3' flaps form duplexes because they comprise reverse complementary sequences which anneal to one another once generated by the prime editors described herein. The duplexes can be incorporated into the target site by cell-driven mechanisms that naturally replace the endogenous duplex sequences located between adjacent nick sites. In certain embodiments, the new duplex sequences can be introduced at one or more locations (*e.g.*, at adjacent genomic loci or on two different chromosomal locations), and can comprise one or more sequences of interest, *e.g.*, protein-encoding sequence, peptide-encoding sequence, or RNA-encoding sequence.

[0303] In some cases, the payload comprises a polynucleotide comprising a scaffold segment, a spacer sequence, a DNA synthesis template, and a primer-binding sequence. In some cases, the scaffold segment and a spacer sequence are on a first nucleic acid molecule and the DNA synthesis template and the primer-binding sequence are on a second nucleic acid molecule.

[0304] In some cases, the guide RNA further comprises a clamp segment. In some cases, the guide RNA comprising, from 3' to 5', a primer binding site, a sequence encoding at least a portion of the first recombinase recognition sequence, a clamp segment, scaffold, and spacer. The clamp segment comprises a sequence that, after being reverse transcribed is at least partially complementary to a genomic site close to the primer binding site and where the spacer binds. Without wishing to be bound by a certain theory, the clamp segment can enhance integration efficiency of the new genetic material that replaces the target sequence at the double-stranded target DNA sequence relative to a guide RNA without the clamp segment. The clamp segment can allow for a reduced number of nucleotides in the primer binding site need to bind its genomic site and facilitate reverse transcription, which in turn enables design of a guide RNA that is shorter than conventional guide RNAs used for other gene editing methods. The clamp segment is described in International Publication No. WO2023215831, which is hereby incorporated herein by reference in its entirety.

[0305] In some cases, a guide RNA can complete the insertion of new genetic material that replaces the target sequence without another guide RNA when delivered to a cell together with a prime editor described herein. The guide RNA can complete the insertion of the new genetic material that replaces the target sequence with a second guide RNA that is a nicking guide RNA when delivered together with a prime editor described herein.

[0306] In some cases, a guide RNA comprises two or more guide RNAs. In some cases, the two or more guide RNAs comprise a first guide RNA encoding at least a first portion of new genetic material that replaces the target sequence. In some cases, the two or more guide RNAs comprise a second guide RNA encoding at least a second portion of the new genetic material that replaces the target sequence. In some cases, the first guide RNA and the second guide RNA work in a pair and collectively encode the new genetic material that replaces the target sequence, thereby inserting the new genetic material that replaces the target sequence into the genome of a cell receiving the lipid delivery particles in a site-specific manner. In some cases, the first and the second portion of the new genetic material that replaces the target sequence have at least 6bp overlap. In some cases, the first portion of the new genetic material that replaces the target sequence is 46 bp. In some cases, the first portion of the new genetic material that replaces the target sequence is 42 bp. In some cases, the first portion or the second portion of the new genetic material that replaces the target sequence is 36 bp, 38 bp, 40 bp, 42 bp, 44 bp, or 46 bp. The first guide RNA comprises a first spacer. The second guide RNA comprises a second spacer. The first spacer and the second spacer bind to two genomic target sites that are within 5-100 bp from each other. When the two or more guide RNAs are delivered to a cell together with a prime editor, the double strand DNA between the two genomic target sites are deleted and the full sequence of the

new genetic material that replaces the target sequence is inserted instead. The deletion can be mediated by the following steps: (a) reverse transcription of the sequence encoding the first portion of the new genetic material that replaces the target sequence in the first guide RNA and the sequence encoding the second portion of the new genetic material that replaces the target sequence in the second guide RNA, wherein the first and the second portion of the new genetic material that replaces the target sequence having at least 6bp overlap, (b) annealing of the two overlapped portion of the new genetic material that replaces the target sequence, (c) synthesis of the second strand comprising the full sequence of the new genetic material that replaces the target sequence, (d) excision of the original DNA sequence, and (e) ligation of the pair nicks. The mechanism, process, and components of this process include those described in International Publication Nos. WO2023122764, WO2023205710, and WO2023225670, each of which is hereby incorporated herein by reference in its entirety.

[0307] In some cases, the payload comprises a polypeptide domain described herein coupled to a polynucleotide domain described herein. In some cases, the payload comprises a polypeptide domain described herein complexed to a polynucleotide domain described herein. In some cases, the payload comprises a ribonucleoprotein. For example, the payload may comprise a guidable polypeptide domain complexed to a polynucleotide configured to bind to the guidable polypeptide domain (e.g., Cas9 complexed with an RNA guide).

[0308] Any of the payloads described herein can further comprise a plasma membrane recruitment element, a transmembrane domain, a signaling domain, a receptor domain, a packaging domain, or a targeting domain. Any of payloads described herein can comprise or be engineered to comprise a protein tag, a peptide tag, or small molecule tag. For example, a payload can comprise a nuclear localization signal (NLS), a nuclear export signal (NES), a cell penetrating peptide (CPP), a mitochondria penetrating peptide (MPP), a solubility tag, a fluorescent tag, or any combinations thereof.

Other Payloads

[0309] In some cases, the payload in the lipid delivery particle of the present disclosure comprises a recombinant protein. The payload can be a diagnostic imaging agent, such as a contrast agent. In some cases, the payload comprises a therapeutic agent, including, but not limited to, a nuclease, a recombinase, a growth factor, an antibody, a chimeric antigen receptor, a T cell receptor, a cytokine, a cytokine inhibitor or agonist, a transcription factor, an organelle, a nucleic acid molecule, a therapeutic DNA, a therapeutic RNA, a retrotransposon, a reverse transcriptase, an oligonucleotide, an aptazyme, an aptamer, or a ribozyme, a generic or specific kinase inhibitor, a small molecule drug, an immunomodulator, a tumor suppressor, a developmental regulator, a cancer vaccine, an anesthetic, an enzyme, a hormone, a ligand, a receptor, a T cell receptor, a

transposon, a retrotransposon, a DNA polymerase, a RNA dependent DNA polymerase, a homing endonuclease, interferons, chemokines, insulin, growth factors, an antisense oligonucleotide, an RNAi, a shRNA, and any combination thereof. The payload can be a prophylactic agent. In some cases, the payload comprises a biomarker. The payload can also comprise an exogenous antigen or an enzyme. In some cases, the payload comprises a metabolite molecule. In some cases, the payload comprises a lipid molecule. In some cases, the payload comprises a structural protein. In some cases, the payload comprises a hormone or a hormonal protein.

CHIMERIC PROTEIN

[0310] In aspects, the present disclosure provides a chimeric protein comprising a plasma membrane recruitment element and a payload that is a protein or a fragment thereof. In some aspects, the lipid delivery particle comprises a chimeric protein comprising a plasma membrane recruitment element and a payload that is a protein or a fragment thereof. In some cases, the plasma membrane recruitment element and the payload are fused directly in the chimeric protein. In other cases, the plasma membrane recruitment element and the payload are fused indirectly via a linker. In some cases, the linker between the plasma membrane recruitment element and the payload is a cleavable linker that is recognized by a protease.

[0311] The chimeric protein (*e.g.*, comprising a gag protein) can form at least part of a protein core of the lipid delivery particle. A lipid delivery particle can comprise two or more chimeric proteins. The chimeric protein can include a structural protein. The structural protein can comprise a plasma membrane recruitment element (*e.g.*, retroviral gag protein). The plasma membrane recruitment element can be fused to a payload. In some cases, the two or more chimeric proteins comprise the same structural protein. In some cases, the two or more chimeric proteins comprise different structural proteins. In some cases, the two or more chimeric proteins comprise different payloads. In some cases, the chimeric protein comprises a payload that comprises a nucleic acid-binding moiety. In some cases, the payload further comprises a guide nucleic acid molecule that forms a ribonucleoprotein complex with the nucleic acid-binding moiety. In some cases, the chimeric protein is suitable for delivery by a lipid delivery particle disclosed herein.

[0312] In some cases, the lipid delivery particle of the present disclosure further comprises a protease that recognizes the cleavable linker in the chimeric protein and cuts the chimeric protein at the cleavable linker. In some cases, the payload is present as a "free" entity separate from the plasma membrane recruitment element. As a result of the cleavage at the cleavable linker by the protease, the payload can be separated from the plasma membrane recruitment element. In some cases, the payload is present as a "free" entity separate from the plasma membrane recruitment element. For instance, the payload can be free and present within an inside of the protein core of the lipid

delivery particle. In some cases, the protease is part of a second chimeric protein comprising a second plasma membrane recruitment element and the protease, where the second plasma membrane recruitment element can be either different from or same as the plasma membrane recruitment element that is fused with the payload.

[0313] In some cases, the chimeric protein disclosed herein also comprises one or more non-cleavable linkers that operably link components together. The non-cleavable linker can be any suitable linker sequence that is used for chimeric protein construction, such as peptide linkers that consist of glycine (Gly) and serine (Ser) residues. In some embodiments, the non-cleavable linker comprises an amino acid sequence selected from the group consisting of: (GS)_x (SEQ ID NO: 426), (GGS)_x (SEQ ID NO: 427), (GGGS)_x (SEQ ID NO: 428), (GGSG)_x (SEQ ID NO: 429), and (SGGG)_x (SEQ ID NO: 430), and wherein x is an integer from 1 to 50. In some cases, the chimeric protein of the present disclosure comprises a nuclear export signal sequence that can direct transport of the chimeric protein out of the nucleus of a cell, *e.g.*, a producer cell.

[0314] In some cases, the chimeric protein disclosed herein has one of the following configurations of components positioned in an order from N-terminus to C-terminus:

[plasma membrane recruitment element]-[cleavable linker]-[payload];
 [plasma membrane recruitment element]-[n * NES]-[cleavable linker]-[payload];
 [plasma membrane recruitment element]-[cleavable linker]-[payload]-[n * NES];
 [plasma membrane recruitment element]-[cleavable linker]-[n * NES]-[payload];
 [plasma membrane recruitment element]-[cleavable linker 1]-[payload]-[cleavable linker 2]-[n * NES];
 [payload]-[cleavable linker]-[n * NES]-[plasma membrane recruitment element];
 [payload]-[n * NES]-[cleavable linker]-[plasma membrane recruitment element];
 [n * NES]-[payload]-[cleavable linker]-[plasma membrane recruitment element];
 [n * NES]-[cleavable linker 1]-[payload]-[cleavable linker 2]-[plasma membrane recruitment element]; and
 [payload]-[cleavable linker]-[plasma membrane recruitment element];

wherein n is an integer in the range of from 1 to 10, and denotes the number of repeats of the NES sequence. Non-cleavable linker sequence can be present or absent in any of the foregoing configurations between any two neighboring components. As provided herein, the payload sequence in the chimeric protein can have one or more NLS sequences, at its N-terminus, C-terminus, or both.

[0315] In some cases, a chimeric protein described herein can comprise: (i) a budding motif, (ii) a multimerization domain, (iii) a post-translation modification (*e.g.*, a post-translational motif), or (iv) any combination thereof. In some cases, a chimeric protein can comprise a budding motif and

a multimerization domain. In some cases, a chimeric protein can comprise a budding motif and a post-translation modification. In some cases, a chimeric protein can comprise a multimerization domain and a post-translation modification. In some cases, a chimeric protein can comprise a budding motif, a multimerization domain, and a post-translation modification.

Motifs and Modifications in Lipid Delivery Particles

[0316] In some aspects, provided herein are short polypeptide motifs and/or modifications that can be present within the chimeric protein as described herein. Without wishing to be bound by theory, the motifs in the chimeric proteins and/or modifications to the chimeric protein disclosed herein can enhance one or multiple characteristics of particle assembly and/or payload delivery, including, but not limited to, payload expression, payload stability, payload loading and offloading, particle formation and/or budding, and payload translocation to and from the nucleus.

[0317] In some cases, in the construct encoding the chimeric proteins disclosed herein, there are zero, or one or more codons, optionally neutral structural codons such as alanine, in between the sequences encoding the different components of the chimeric proteins, including, *e.g.*, the motifs disclosed herein (particle budding motif, membrane penetrating peptide, post-translational modification motif, multimerization motif, nuclear export signal, and/or nuclear localization signal), plasma membrane recruitment element, and payload.

Particle Budding Motif

[0318] A particle budding motif can be a short polypeptide fused to the payload or plasma membrane recruitment element. Viral particles and/or viral-like particles may bud through limiting host cell membranes and acquire a lipid envelope to exit a cell. In some cases, viral budding may utilize short peptide motifs (*e.g.*, late domains) located within the viral protein that function by recruiting cellular factors. Without wishing to be bound by theory, the particle budding motif as described herein can enhance the interaction with host factors that are responsible for the creation of lipid delivery particles. In some cases, the budding motif can be a PPXY, LPXY, PTAP, YPX_nL, or any combination thereof. In some embodiments, the budding motif can be YPX_nL, wherein *n* can be an integer from 1 to 5. In some cases, *n* = 1, *n* = 2, or *n* = 3. The particle budding motif can bind NEDD4, TSG101, and/or Alix. In some cases, the particle budding motif promotes the efficiency of particle assembly or budding (*e.g.*, increase particle titer). In some cases, the particle budding motif increases the particle titer by at least 1-fold, at least 2-fold, at least 3-fold, at least 4-fold, or at least 5-fold. In some cases, the particle budding motif increase the particle titer by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90%. The particle budding motif can comprise a late assembly (L) domain. L domains are conserved sequences that can engage the endosomal sorting complex required for transport

(ESCRT) components to promote budding. The particle budding motif can be found in retroviral gag proteins which recruit members of the ESCRT complex. The ESCRT complex is located on the cytoplasmic surface of the multivesicular body and can recognize ubiquitinated proteins into vesicles. The vesicles can then bud into the multivesicular body and can be released into the extracellular space. The ESCRT also plays a role in endolysosomal membrane and plasma membrane repair, neuronal pruning, nuclear envelope maintenance, and autophagy. The ESCRT complex comprises three complexes (ESCRT-I, ESCRT-II, ESCRT-III) which are active in the cytosol and the plasma membrane to assist in assembly and fission from the membrane. The particle budding motif can comprise PTAP, PPXY, LPYX, PPEE, LYPL, LYPSL, PPPY, PPEY, PSAP, LYPAL, PPAP, PPPE, YMYL, YRKL, YQCL, YCYL, LYRTL, YPXnL, or any combination thereof.

[0319] The lipid delivery particle can comprise the particle budding motif as described herein. In some cases, the particle budding motif is linked to the C-terminus of the chimeric protein. In some cases, the particle budding motif is linked to the N-terminus of the chimeric protein. In some cases, the particle budding motif is linked to the payload. In some cases, the particle budding motif is linked to a plasma membrane recruitment element (*e.g.*, a PH domain). In some cases, the particle budding motif is linked between the payload and the plasma membrane recruitment element. In some cases, the particle budding motif is linked to a nuclear export signal domain. In some cases, the particle budding motif is linked to a non-endogenous cleavage site. In some cases, the particle budding motif is linked to a post-translational modification motif. In some cases, the particle budding motif is present as tandem repeats in the chimeric protein. In some cases, the chimeric protein can comprise one, two, three, four, five, or more particle budding motifs. In some cases, addition of two, three, four, or five budding motifs to the chimeric protein may enhance a lipid delivery particle compared to a lipid delivery particle with a chimeric protein comprising only one budding motif. In some cases, flexible linkers can be present in-between the particle budding motifs. In some cases, a particle budding motif can have a linker to the N-terminus and/or the C-terminus.

[0320] Addition of budding motif to a chimeric protein described herein may increase an efficacy of a lipid delivery particle, for instance, increasing amount of payload loaded in the lipid delivery particle (*e.g.*, average amount per each particle among a population of particles), increasing amount of payload released by the lipid delivery particle upon fusion to a target cell (*e.g.*, average amount per each particle among a population of particles), increasing titer of the lipid delivery particle (*e.g.*, amount of lipid delivery particles per each preparation, *e.g.*, amount of particles per unit volume (*e.g.*, mL) of solution containing the particles that is prepared from a certain amount of production cells), or any combination thereof. In some cases, fusion of a budding motif in a

construct described herein may improve a base editing percentage (e.g., improve a capacity for a lipid delivery particle to deliver base editing payload to a transduced cell). In some cases, addition of budding motif can increase a mean base editing in a cell compared to a level of base editing in a cell transduced with a lipid delivery particle with no budding motif. Addition of one or more budding motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 99% greater than a mean base editing percentage from a lipid delivery particle with no budding motif. Addition of one or more budding motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at most about 99%, at most about 95%, at most about 90%, at most about 85%, at most about 80%, at most about 75%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less than about 5% greater than a mean base editing percentage from a lipid delivery particle with no budding motif. Addition of one or more budding motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 10% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a budding motif. Addition of one or more budding motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 30% to about 40%, about 30% to about 50%, about 30% to about 60%, about 30% to about 65%, about 30% to about 70%, about 30% to about 75%, about 30% to about 80%, about 30% to about 85%, about 30% to about 90%, about 30% to about 95%, about 30% to about 99%, about 40% to about 50%, about 40% to about 60%, about 40% to about 65%, about 40% to about 70%, about 40% to about 75%, about 40% to about 80%, about 40% to about 85%, about 40% to about 90%, about 40% to about 95%, about 40% to about 99%, about 50% to about 60%, about 50% to about 65%, about 50% to about 70%, about 50% to about 75%, about 50% to about 80%, about 50% to about 85%, about 50% to about 90%, about 50% to about 95%, about 50% to about 99%, about 60% to about 65%, about 60% to about 70%, about 60% to about 75%, about 60% to about 80%, about 60% to about 85%, about 60% to about 90%, about 60% to about 95%, about 60% to about 99%, about 65% to about 70%, about 65% to about 75%, about 65% to about 80%, about 65% to about 85%, about 65% to about 90%, about 65% to about 95%, about 65% to about 99%, about 70% to about 75%, about 70% to about 80%, about 70% to about 85%, about 70% to about 90%, about 70% to about 95%, about 70% to about 99%, about 75% to about 80%, about 75% to about 85%, about 75% to about 90%, about 75% to about 95%, about 75% to about 99%, about 80% to about 85%, about 80% to about 90%,

about 80% to about 95%, about 80% to about 99%, about 85% to about 90%, about 85% to about 95%, about 85% to about 99%, about 90% to about 95%, about 90% to about 99%, or about 95% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a budding motif.

[0321] In some cases, the particle budding motif is sourced from a virus. The virus can include, but is not limited to, HIV, EIAV, RSV, ALV, FUJSV, HTLV-1, MLV, EBOV, HTLV-2, SRV-2, CHMP4A, AVISY, fMLV, influenza, zika virus, dengue virus, WNV, and Tupaia. In some cases, the particle budding motif comprises a sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-32. In some cases, the particle budding motif comprises a sequence with at least 90% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-32. In some cases, the particle budding motif comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 1-32. In some cases, the particle budding motif comprises a sequence as set forth in any one of SEQ ID NOs: 1-32. In some cases, the particle budding motif comprises any of the sequences or a minimal sequence drawn from any of the sequences as set forth in **Table 6**.

Table 6. Exemplary particle budding motif sequences

SEQ ID NO.	Amino Acid Sequence	Virus Source	Budding Motif
1	LQSRPEPTAPPEESFRFGEETTTTPSQKQEPIDKE LYPLASLRSLFGSDPSSQ	HIV Gag	PTAP/PPEE/LYPL
2	PIQQKSQHNSVVQETPQTQNLYPDLSIKKEYN VKEKDQVEDLNLDLSLWE	EIAV Gag	LYPDLS
3	GETTVQRDAKMAPEETATPKTVGGTSCYHCGTAI GCNCATASAPPPPYVGSGLYPSLAGV	RSV / FUJSV / ALV Gag	PPPY/LYPSL
4	PAPPPSSPTHDPDSDPQIPPPYVEPTAPQVL	HTLV-1 Gag	PPPY/PTAP
5	LPPLPTAPVL	MLV Gag	PTAP
6	LPAPPEYMEAIYPA	EBOV VP40	PTAP/PPEY
7	LVKNQVSPSAPAAPVPTPICPTTTPPPPPPSPE AHVPPPYVEPTTTQCFPII	HTLV-2 Gag	PSAP/PPPY
8	PDWPPFLSRPPPNRTAATA	SRV-2 Gag	PPPY
9	LTRPPPNKATPSAPTVMVV	SRV-1 / MPMV Gag	PPPY/PSAP
10	MSGLGRLFGKGKKEKGPTPEEA IQKLKETEKI LI KKQEFLEQKIQQELQTAKKYGTKNKRAALQALRR KKRFEQQLAQTDGTLSTLEFQREAIENATTNAEV LRTMELAAQSMKKAYQDMDIDKVDLMTDI TEQQ EVAQQI SDAI SRPMGFDDVDEDELEEELEEQ EELAQE LLNVGDKEEPPSVKLP SVPS THLPAGPA PKVDEDEEALKQLAEWVSGDGS GSGVSAREP SA PSIPTPAYQSSPAGGHAPTPPPYAPRTMPPTK	CHMP4A W/ other motifs	PSAP/PPPY
11	MSGLGRLFGKGKKEKGPTPEEA IQKLKETEKI LI KKQEFLEQKIQQELQTAKKYGTKNKRAALQALRR KKRFEQQLAQTDGTLSTLEFQREAIENATTNAEV LRTMELAAQSMKKAYQDMDIDKVDLMTDI TEQQ EVAQQI SDAI SRPMGFDDVDEDELEEELEEQ EELAQE LLNVGDKEEPPSVKLP SVPS THLPAGPA PKVDEDEEALKQLAEWVS	CHMP4A	PSVP

12	GETTVQRDAKMAPEETATPKTVGGTSCYHCGTAS GCNCATATASAPPPYVGSGLCPSLAGVG	AVISY Gag	PPPY
13	GPVPTTSTLLLPPEYSSWGYPYEAPPSEYEQ	HUMAN ARRD1	PPEY/PPSY
14	LYPALTSPLNTKPRPQVLPSDGGPLIDLLEDPP PY	fMLV Gag	LYPAL/PPPY
15	KQVKENKTQPPVAYQYWPPAELQYRPPPEQYGY PGMPAPQGRAPYPQPPTRLNPTAPPSRQGSSEL HEIIDKSRK	kcon (HERV-K) Gag	PPPE/PPAP/PTAP
16	LQSRPEPTAPPEESFRFGEETTTTPSQKQEPIDKE LYPLASLRSLFGSDPSSQKMQIFVKTLTGKTI TL EVEP SDTI ENVKAKIQDKEGIPDQQR LIFAGKQ LEDGRTLSDYNIQKESTLHLVLRRLGG	HIV Gag with Ubiquilated Lysine	PTAP/PPEE/LYPL
17	KVLLSR TALLPQPSAPSAPPVLPSEEEESFHP FPPPYNPPAPPESSLV	Tcon Gag	PPPY/PAPP
18	TPGANERKYNMYLICYGFVEDVERTPETGKRK KIR TIA	Nipah M protein	YMYL
19	GLGDRKDECSMYMFLLGVEDSDPLGPPIG	Measles M Protein	YMFL
20	NNMDKAVKLYRKLKREITFHGAKEISLSYS	Influenza M protein	YRKL
21	YDQSGAASANKYRYMSLKQDHTVKEGTMD	Sindbis Structural Polypro.	YMSL
22	KKDDPVGFPGGYQCLLKVIIQLEKDPKTS	Tupaia M Protein	YQCL
23	ANLADVRSYCYLASVSDLSTGRFSITPSAP	WNV Genome Polyprotein	YCYL, PSAP
24	KRFSLGMLQGRGPLKLYMALVAFLRFLTIP	Dengue Genome Glycoprotein	YMAL
25	MAAEITRRGSAYMYLDRSDAGKAISFATT	Zika Genome Polyprotein	YMYL
26	PNCFNSSINNIHMEIQLKDALEKNQQLVYDQQ REVYVKGLLAKI FELEKKTETAHSLPQQTKKPE SEGYLQEEKQKC	Human Cep55	
27	PPQPPLPETKEKEQAPPAPSSLYRTLHLGS	Human Gag V	PPAP/LYRTL
28	KPPPYKPSAP	Basic PPPY + PSAP motif	PPPY/PSAP
29	LPTAPPEYMEAIYPAKMQIFVKTLTGKTI LVE PSDTIENVKAKIQDKEGIPDQQR LIFAGKQLED GRTLSDYNIQKESTLHLVLRRLGG	EBOV VP40 W/ Ubiquitylated Lysine	PTAP/PPEY
30	LPTAPPEYMEAIYPALPTAPPEYMEAIYPALPTA PPEYMEAIYPA	EBOV VP40 repeated	PTAP/PPEY
31	TVDISPDTVGTLSLIMLAQAQEVF FLKATRDMMK DAIIAKLANQAADYFGDAFKQCQYKDTLPKEVFP VLAAKHCIMQANAEYHQSILAKQQKFGEEIARL QHAAELIKTVASRYDEYVNVKDFSDKINRALAAA KKDNDFIYHDRVPLKDLDP IGKATLVKSTPVNV PISQKFTDLFEKMPVSVQQSLAAYNQRKADLVN RSIAQMREATTLAGDGGSGGVSAREPSAPSIPT PAYQSSPAGGHAPT PPPYAPRTMPPTK	Chimeric Alix Protein	
32	KPPPYK	Basic PPPY motif	PPPY

[0322] In some cases, the construct for the lipid delivery particle or the chimeric protein with a particle budding motif is designed with any of the following arrangements in **Table 12**.

Membrane Penetrating Peptide

[0323] A membrane penetrating peptide (e.g., a cell penetrating peptide) can be a short polypeptide fused to the payload or plasma membrane recruitment element that can facilitate protein translocation across a phospholipid bilayer. As referred herein, a “membrane penetrating peptide” can comprise a short peptide (e.g., 5-30 amino acids) comprising a charge, which can facilitate penetration into a cell across a cell membrane. Without wishing to be bound by theory, a membrane penetrating peptide may facilitate delivery of cargo through endocytosis. A

membrane penetrating peptide may have a positive charge. A membrane penetrating peptide may have a negative charge. A membrane penetrating peptide may have a neutral charge. In some cases, the membrane penetrating peptide can enhance delivery of cargo through energy dependent mechanisms (e.g., endocytosis), energy independent mechanisms (e.g., direct penetration), or any combination thereof.

[0324] In some cases, the membrane penetrating peptide is fewer than 40 amino acids. In some cases, the membrane penetrating peptide is fewer than 30 amino acids. In some cases, the membrane penetrating peptide is fewer than 20 amino acids. In some cases, the membrane penetrating peptide is about 40 amino acids. In some cases, the membrane penetrating peptide is about 35 amino acids. In some cases, the membrane penetrating peptide is about 30 amino acids. In some cases, the membrane penetrating peptide is about 25 amino acids. In some cases, the membrane penetrating peptide is about 20 amino acids.

[0325] In some cases, the membrane penetrating peptide promotes recruitment of the payload to the plasma membrane. In some cases, the membrane penetrating peptide increases the efficiency of particle assembly and/or budding. In some cases, the membrane penetrating peptide functions in a pH-inducible manner. In some cases, the pH that can modulate membrane penetrating activity of the membrane penetrating peptide is between 3 and 10, such as 3, 4, 5, 6, 7, 8, 9 or 10. Without wishing to be bound by theory, membrane penetrating peptides facilitate payload translocation via penetration of the membrane, endocytosis-mediated entry, or translocation through a transitory structure.

[0326] Membrane penetrating peptides can have amino acid compositions that either contain a high relative abundance of positively charged amino acids such as lysine or arginine or have sequences that contain an alternating pattern of polar, charged amino acids and non-polar, hydrophobic amino acids, as described in Derakhshankhah and Jafari (2018) *Biomedicine and Pharmacotherapy* 108:1090-1096. Membrane penetrating peptides can also contain only apolar residues with low net charge or hydrophobic amino acid groups that are helpful for cellular uptake. As described herein, the membrane penetrating peptides can be cationic, amphipathic, or hydrophobic.

[0327] The chimeric protein can comprise the membrane penetrating peptide as described herein. In some cases, the membrane penetrating peptide is linked to the C-terminus of the chimeric protein. In some cases, the membrane penetrating peptide is linked to the N-terminus of the chimeric protein. In some cases, the membrane penetrating peptide is linked to the payload. In some cases, the membrane penetrating peptide is linked to a plasma membrane recruitment element (e.g., a PH domain). In some cases, the membrane penetrating peptide is linked between

the payload and the plasma membrane recruitment element. In some cases, the membrane penetrating peptide is present as tandem repeats in the chimeric protein.

[0328] The construct for the lipid delivery particle or the chimeric protein with a membrane penetrating peptide will be designed with any of the following arrangements in **Table 13**. PH domain is abbreviated PH, membrane penetrating peptide is abbreviated MPP, and a protease cleavage site is abbreviated CS.

[0329] Addition of a membrane penetrating peptide to a chimeric protein described herein may increase an efficacy of a lipid delivery particle, for instance, increasing amount of payload loaded in the lipid delivery particle (*e.g.*, average amount per each particle among a population of particles), increasing amount of payload released by the lipid delivery particle upon fusion to a target cell (*e.g.*, average amount per each particle among a population of particles), increasing titer of the lipid delivery particle (*e.g.*, amount of lipid delivery particles per each preparation, *e.g.*, amount of particles per unit volume (*e.g.*, mL) of solution containing the particles that is prepared from a certain amount of production cells), or any combination thereof. In some cases, fusion of a membrane penetrating peptide in a construct described herein may improve a base editing percentage (*e.g.*, improve a capacity for a lipid delivery particle to deliver base editing payload to a transduced cell). In some cases, addition of a membrane penetrating peptide can increase a mean base editing in a cell compared to a level of base editing in a cell transduced with a lipid delivery particle with no budding motif. Addition of one or more membrane penetrating peptides to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 99% greater than a mean base editing percentage from a lipid delivery particle with no membrane penetrating peptide. Addition of one or more membrane penetrating peptides to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at most about 99%, at most about 95%, at most about 90%, at most about 85%, at most about 80%, at most about 75%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less than about 5% greater than a mean base editing percentage from a lipid delivery particle with no membrane penetrating peptide. Addition of one or more membrane penetrating peptides to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 10% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a membrane penetrating peptide. Addition of one or more membrane penetrating peptides to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing

percentage between about 30% to about 40%, about 30% to about 50%, about 30% to about 60%, about 30% to about 65%, about 30% to about 70%, about 30% to about 75%, about 30% to about 80%, about 30% to about 85%, about 30% to about 90%, about 30% to about 95%, about 30% to about 99%, about 40% to about 50%, about 40% to about 60%, about 40% to about 65%, about 40% to about 70%, about 40% to about 75%, about 40% to about 80%, about 40% to about 85%, about 40% to about 90%, about 40% to about 95%, about 40% to about 99%, about 50% to about 60%, about 50% to about 65%, about 50% to about 70%, about 50% to about 75%, about 50% to about 80%, about 50% to about 85%, about 50% to about 90%, about 50% to about 95%, about 50% to about 99%, about 60% to about 65%, about 60% to about 70%, about 60% to about 75%, about 60% to about 80%, about 60% to about 85%, about 60% to about 90%, about 60% to about 95%, about 60% to about 99%, about 65% to about 70%, about 65% to about 75%, about 65% to about 80%, about 65% to about 85%, about 65% to about 90%, about 65% to about 95%, about 65% to about 99%, about 70% to about 75%, about 70% to about 80%, about 70% to about 85%, about 70% to about 90%, about 70% to about 95%, about 70% to about 99%, about 75% to about 80%, about 75% to about 85%, about 75% to about 90%, about 75% to about 95%, about 75% to about 99%, about 80% to about 85%, about 80% to about 90%, about 80% to about 95%, about 80% to about 99%, about 85% to about 90%, about 85% to about 95%, about 85% to about 99%, about 90% to about 95%, about 90% to about 99%, or about 95% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a membrane penetrating peptide.

[0330] In some cases, the membrane penetrating peptide comprises a sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 33-56. In some cases, the membrane penetrating peptide comprises a sequence with at least 90% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 33-56. In some cases, the membrane penetrating peptide comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 33-56. In some cases, the membrane penetrating peptide comprises a sequence as set forth in any one of SEQ ID NOs: 33-56. In some cases, the membrane penetrating peptide comprises a sequence as set forth in **Table 7**.

Table 7. Exemplary membrane penetrating peptide sequences

SEQ ID NO.	Name	Amino Acid Sequence
33	YTA2	YTAIAWVKAFIRKLRLK
34	HIV-1 TAT	GRKKRRQRRRPPQ

35	Pep-c19	VVNKLIRNNKMNC
36	Vectofusin	KKALLHAALAHLLALAHHLALLKKA
37	L17E	IWLTALKFLGKHAAKHEAKQQLSKL
38	PCV2 NLSa	MTYPRRRFRRRHRPRS
39	9R	RRRRRRRR
40	11R	RRRRRRRRRR
41	Penetratin	RQIKIWFQNRRMKWKK
42	Pep-1a	KETWWETWWTEWSQPKKKRKV
43	Pep-1b	KETWETWWTEWSQPKKKRKV
44	MTD_1	AAAVPVLVAA
45	MTD_2	PLLLLLPAL
46	MTD_3	AVALVAVVAVA
47	MTD_4	LLAAAAALLA
48	MTD_5	AVVVPALVLAAP
49	MTD_6	AVLPLLAVP
50	MTD_7	AVKKKPAAA
51	pVEC	LLIILRRRIRKQAHASK
52	TP10	AGYLLGKINLKALAALAKKIL
53	MITO1	KLAKLAKKLAKLAK
54	nCTRL	DLSLARLATARLAI
55	KALA	WEAKLAKALAKALAKHLAKALAKALKACEA
56	HIV-1 TATplus	YGRKKRRQRRR

Post-Translational Modifications Motif

[0331] A post-translational modification motif disclosed herein can be a short peptide present within the chimeric protein that leads the payload to be modified following translation (*e.g.*, a post-translational modification). A post-translational modification can be any modification following translation. For example, the post-translational modification may be any chemical modification of the amino acids in a protein product after the protein product is translated. In some cases, the post-translational modification motif is between about 10 and about 40 amino acids in length. Without wishing to be bound by theory, the post-translational modification alters the payload in a way that can promote recruitment to the plasma membrane. In some cases, the post-translational modification motif, upon the corresponding post-translational modification to the chimeric protein, can alter localization of the payload. The post-translational modification can be a covalent modification of one or more amino acids within the chimeric protein following biosynthesis. The post-translational modification motif can promote plasma membrane localization of cytosolic proteins. The post-translational modification motif can be enzymatic. As

described herein, a post-translational modification motif can be to one or more amino acids within the payload, the plasma membrane recruitment element, or both. In some cases, a post-translational modification comprises phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, lipidation, acetylation, or proteolysis. In some cases, a post-translational modification comprises geranylgeranylation, myristoylation, glypiation, isoprenylation, palmitoylation, farnesylation, or any combination thereof. Acetylation can comprise transfer of an acetyl group to nitrogen. Myristoylation can comprise an attachment of a myristoyl group to a N-terminal glycine residue of a protein. Isoprenylation can comprise an attachment of a prenyl group (*e.g.*, a hydrophobic molecule) to a protein. Farnesylation can comprise an attachment of a farnesyl group to a C-terminal cysteine residue of a protein. Phosphorylation can comprise addition of a phosphate group and can occur on serine, threonine, or tyrosine residues. Glycosylation can comprise addition of a sugar moiety (*e.g.*, monosaccharide, polysaccharide, oligosaccharide, or carbohydrate) to a residue. Ubiquitination can comprise the addition of a ubiquitin polypeptide to a lysine residue of a protein. Nitrosylation can comprise incorporation of a nitrosyl moiety of nitric oxide to a protein. Nitrosylation can occur on free cysteine residues to produce S-nitrothiols. Methylation can comprise the transfer of one-carbon methyl groups to a protein and can increase hydrophobicity of the protein. Lipidation can comprise incorporation of a lipid moiety to a protein. Proteolysis can comprise cleaving of peptide bonds of a protein and can assist antigen processing, apoptosis, surface protein shedding, and cell signaling. Geranylgeranylation can comprise attachment of at least one lipophilic geranylgeranyl isoprene unit from geranylgeranyl diphosphate to at least one cysteine residue. Glypiation can comprise covalent bonding of a glycosylphosphatidylinositol (GPI) anchor to a protein. Palmitoylation can comprise the attachment of at least one fatty acid (*e.g.*, palmitic acid) to a cysteine residue of a protein.

[0332] In some cases, the construct for the lipid delivery particle or the chimeric protein with a particle budding motif and a post-translational modification is designed with any of the following arrangements in **Table 15**. PH domain is abbreviated PH, particle budding motif is abbreviated BD, post-translational modification motif is abbreviated PTM, and a protease cleavage site is abbreviated CS.

[0333] The chimeric protein can comprise the post-translational modification motif as described herein. In some cases, the post-translational modification motif is linked to the C-terminus of the chimeric protein. In some cases, the post-translational modification motif is linked to the N-terminus of the chimeric protein. In some cases, the post-translational modification motif is linked to the payload. In some cases, the post-translational modification motif is linked to a plasma membrane recruitment element (*e.g.*, a PH domain). In some cases, the post-translational modification motif is linked to both a payload and a plasma membrane recruitment element (*e.g.*,

a PH domain). In some cases, the post-translational modification motif is linked to some other part(s) of the chimeric protein than a payload and a plasma membrane recruitment element (*e.g.*, a PH domain). In some cases, the post-translational modification motif is linked to some other part(s) of the chimeric protein besides a payload and a plasma membrane recruitment element (*e.g.*, a PH domain). In some cases, the post-translational modification motif is linked between the payload and the plasma membrane recruitment element. In some cases, the post-translational modification motif is linked to a nuclear export signal domain. In some cases, the post-translational modification motif is linked to a non-endogenous cleavage site. In some cases, the post-translational modification motif is linked to a particle budding motif sequence. In some cases, the post-translational modification motif is present as tandem repeats in the chimeric protein.

[0334] In some cases, the post-translational modification motif comprises a sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 57-86. In some cases, the post-translational modification motif comprises a sequence with at least 90% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 57-86. In some cases, the post-translational modification motif comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 57-86. In some cases, the post-translational modification motif comprises an amino acid sequence as set forth in any one of SEQ ID NOs: 57-86, with at least one but not more than 8 nucleic acid differences. In some cases, the post-translational modification motif comprises a sequence as set forth in any one of SEQ ID NOs: 57-86. In some cases, the post-translational modification motif comprises a sequence as set forth in **Table 8**.

[0335] The construct for the lipid delivery particle or the chimeric protein with a post-translational modification motif will be designed with any of the following arrangements in **Table 14**. PH domain is abbreviated PH, post-translational modification motif is abbreviated PTM, and a protease cleavage site is abbreviated CS.

[0336] Addition of a post-translational modification motif to a chimeric protein described herein may increase an efficacy of a lipid delivery particle, for instance, increasing amount of payload loaded in the lipid delivery particle (*e.g.*, average amount per each particle among a population of particles), increasing amount of payload released by the lipid delivery particle upon fusion to a target cell (*e.g.*, average amount per each particle among a population of particles), increasing titer of the lipid delivery particle (*e.g.*, amount of lipid delivery particles per each preparation, *e.g.*, amount of particles per unit volume (*e.g.*, mL) of solution containing the particles that is prepared from a certain amount of production cells), or any combination thereof. In some cases, fusion of a post-translational modification motif in a construct described herein may improve a base editing

percentage (e.g., improve a capacity for a lipid delivery particle to deliver base editing payload to a transduced cell). In some cases, addition of a post-translational modification motif can increase a mean base editing in a cell compared to a level of base editing in a cell transduced with a lipid delivery particle with no budding motif. Addition of one or more post-translational modification motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 99% greater than a mean base editing percentage from a lipid delivery particle with no post-translational modification motif. Addition of one or more post-translational modification motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at most about 99%, at most about 95%, at most about 90%, at most about 85%, at most about 80%, at most about 75%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less than about 5% greater than a mean base editing percentage from a lipid delivery particle with no post-translational modification motif. Addition of one or more post-translational modification motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 10% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a post-translational modification motif. Addition of one or more post-translational modification motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 30% to about 40%, about 30% to about 50%, about 30% to about 60%, about 30% to about 65%, about 30% to about 70%, about 30% to about 75%, about 30% to about 80%, about 30% to about 85%, about 30% to about 90%, about 30% to about 95%, about 30% to about 99%, about 40% to about 50%, about 40% to about 60%, about 40% to about 65%, about 40% to about 70%, about 40% to about 75%, about 40% to about 80%, about 40% to about 85%, about 40% to about 90%, about 40% to about 95%, about 40% to about 99%, about 50% to about 60%, about 50% to about 65%, about 50% to about 70%, about 50% to about 75%, about 50% to about 80%, about 50% to about 85%, about 50% to about 90%, about 50% to about 95%, about 50% to about 99%, about 60% to about 65%, about 60% to about 70%, about 60% to about 75%, about 60% to about 80%, about 60% to about 85%, about 60% to about 90%, about 60% to about 95%, about 60% to about 99%, about 65% to about 70%, about 65% to about 75%, about 65% to about 80%, about 65% to about 85%, about 65% to about 90%, about 65% to about 95%, about 65% to about 99%, about 70% to about 75%, about 70% to about 80%, about 70% to about 85%, about 70% to about 90%, about 70% to about 95%, about 70% to about 99%, about 75% to about 80%, about 75% to about 85%, about 75% to about 90%,

about 75% to about 95%, about 75% to about 99%, about 80% to about 85%, about 80% to about 90%, about 80% to about 95%, about 80% to about 99%, about 85% to about 90%, about 85% to about 95%, about 85% to about 99%, about 90% to about 95%, about 90% to about 99%, or about 95% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a post-translational modification motif.

Table 8. Exemplary post-translational modification motif sequences

SEQ ID NO.	SEQ Origin	Amino Acid Sequence	Post-Translational Modification	Location on PH-Freight
57	HIV1	MGARASVLSGGELDKWE	Myristoylation	N-terminal
58	HIV2	MGARGSVLSGKKTDELE	Myristoylation	N-terminal
59	MMTV	MGQKLFVSVLQRLLSER	Myristoylation	N-terminal
60	MPMV	MGQELSQHERYVEQLKQ	Myristoylation	N-terminal
61	HERVK	MGQTKSKIISKYASYLS	Myristoylation	N-terminal
62	MoMuLV	MGQTVTTPLSLTGLHWK	Myristoylation	N-terminal
63	HTLV1	MGQIFSRASAPIRPPR	Myristoylation	N-terminal
64	BLV	MGNPSYNPPAGISPSD	Myristoylation	N-terminal
65	RSV	MEAVIKVISSACKTYCG	Acetylation	N-terminal
66	EIAV	MGDPLTWSKALKKLEKV	Acetylation	N-terminal
67	FV	MASGSNVEEYELDVEAL	Acetylation	N-terminal
68	Myr1	MGSSKSKPKDEAVIKVISSACKTYCG	Acetylation	N-terminal
69	HERVV1	MGNQSSVPQDSPLGCIL	Myristoylation	N-terminal
70	hYes	MGCIKSKENKSPAIKYR	Myristoylation	N-terminal
71	hSrc	MGSNKS KPKDASQRRR	Myristoylation	N-terminal
72	hLck	MGCGCSSHPEDDWENI	Myristoylation	N-terminal
73	hFyn	MGCVQCKDKEATKLTEE	Myristoylation	N-terminal
74	hLyn	MGCIKSKGKDSLSDGV	Myristoylation	N-terminal
75	hYrk	MGCVQCKDKEATKLTEE	Myristoylation	N-terminal
76	hBlk	MGLVSSKKPKDKEPKIKE	Myristoylation	N-terminal
77	hHck	MGGRSSCEDPGCPRDEE	Myristoylation	N-terminal
78	hFgr	MGCVFCKKLEPVATAKE	Myristoylation	N-terminal
79	hArf	MGNIFANLFGKLFKGKE	Myristoylation	N-terminal
80	hFSP1	MGSQVSVESGALHVIV	Myristoylation	N-terminal
81	hRhoB	CINCKVL	isoprenylation and palmitoylation	C-terminal
82	hH-Ras	QHKLRKLNPPDESGPGCMSCKCVLS	isoprenylation	C-terminal
83	hN-Ras	QYRMKKLNSSDDGTQGCMGLPCVVM	isoprenylation	C-terminal

84	hK-Ras4a	QYRLKKISKEEKTPGCVKIKKCI IM	isoprenylation	C-terminal
85	hK-Ras4b	KHKEKMSKDGGKKKKKSKTKCVI M	isoprenylation	C-terminal
86	hH-Ras	LRSRAQASNSGCMSCCKCVLS	farnesylation motif of HRAS	C-terminal

Multimerization Motif

[0337] The chimeric protein described herein can comprise a multimerization motif. In some cases, the multimerization motif can perform protein-protein interactions. As referred to herein, a “multimerization motif” can comprise any amino acid sequence wherein the sequence recognizes another amino acid sequence of a monomer and promotes multimerization (e.g., formation of a multimer). For example, multimerization may occur at a plasma membrane and enhance assembly and budding of particles described herein. In some cases, the multimerization motif forms a parallel homodimer, such as a coiled-coil. In some cases, the multimerization motif forms dimers, trimers, tetramers, or other oligomers. In some cases, the multimerization motif is a leucine zipper oligomerization motif. The leucine zipper oligomerization motif can comprise a series of leucines spaced at intervals of every seventh amino acid along an α -helix. In some cases, a leucine zipper oligomerization motif can mediate dimerization or oligomerization. In some cases, the coiled-coil is a leucine zipper oligomerization motif (e.g., a GCN4 leucine zipper or a mutant thereof). Upon dimerization, the leucine zipper oligomerization motif can form a parallel-coiled coil. In some cases, the leucine zipper oligomerization motif (e.g., the GCN4 leucine zipper motif). Without wishing to be bound by theory, the GCN4 leucine zipper can be used to induce multimerization that can assist in assembly of the lipid delivery particles described herein. As such, the GCN4 leucine zipper can complement removal of nucleic acid binding domains that are necessary for multimerization in native gag proteins. Multimerization motifs can form dimers, trimers, tetramers, or any combination thereof.

[0338] In some cases, the leucine zipper oligomerization motif is linked to the C-terminus of the chimeric protein. In some cases, the leucine zipper oligomerization motif is linked to the N-terminus of the chimeric protein. In some cases, the leucine zipper oligomerization motif is linked to the payload. In some cases, the leucine zipper oligomerization motif is linked to a plasma membrane recruitment element (e.g., a PH domain). In some cases, the leucine zipper oligomerization motif is linked to both a payload and a plasma membrane recruitment element (e.g., a PH domain). In some cases, the leucine zipper oligomerization motif is linked to some other part(s) of the chimeric protein than a payload and a plasma membrane recruitment element (e.g., a PH domain). In some cases, the leucine zipper oligomerization motif is linked to some

other part(s) of the chimeric protein besides a payload and a plasma membrane recruitment element (*e.g.*, a PH domain). In some cases, the leucine zipper oligomerization motif is linked between the payload and the plasma membrane recruitment element. In some cases, the leucine zipper oligomerization motif is linked to a nuclear export signal domain. In some cases, leucine zipper oligomerization motif is linked to a non-endogenous cleavage site. In some cases, the leucine zipper oligomerization motif is linked to a particle budding motif sequence. In some cases, leucine zipper oligomerization motif is present as tandem repeats in the chimeric protein.

[0339] The construct for the lipid delivery particle or the chimeric protein with a multimerization motif will be designed with any of the following arrangements in **Table 16**. In this Example, a leucine zipper domain will be the multimerization motif. PH domain is abbreviated PH, particle budding motif is abbreviated BD, post-translational modification motif is abbreviated PTM, leucine zipper oligomerization motif is abbreviated LZ, and a protease cleavage site is abbreviated CS. The notation (BD-Payload) will be any one of the combinations of budding and payload configurations. The notation (PH/PTM) will be any one of the combinations of PTMs and payloads, either singly or in combination.

[0340] Addition of multimerization motif to a chimeric protein described herein may increase an efficacy of a lipid delivery particle, for instance, increasing amount of payload loaded in the lipid delivery particle (*e.g.*, average amount per each particle among a population of particles), increasing amount of payload released by the lipid delivery particle upon fusion to a target cell (*e.g.*, average amount per each particle among a population of particles), increasing titer of the lipid delivery particle (*e.g.*, amount of lipid delivery particles per each preparation, *e.g.*, amount of particles per unit volume (*e.g.*, mL) of solution containing the particles that is prepared from a certain amount of production cells), or any combination thereof. In some cases, fusion of a multimerization motif in a construct described herein may improve a base editing percentage (*e.g.*, improve a efficacy for a lipid delivery particle to deliver base editing payload to a transduced cell). In some cases, addition of multimerization motif can increase a mean base editing in a cell compared to a level of base editing in a cell transduced with a lipid delivery particle with no multimerization motif. Addition of one or more multimerization motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 99% greater than a mean base editing percentage from a lipid delivery particle with no multimerization motif. Addition of one or more multimerization motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage from at most about 99%, at most about 95%, at most

about 90%, at most about 85%, at most about 80%, at most about 75%, at most about 70%, at most about 60%, at most about 50%, at most about 40%, at most about 30%, at most about 20%, at most about 10%, at most about 5%, or less than about 5% greater than a mean base editing percentage from a lipid delivery particle with no multimerization motif. Addition of one or more multimerization motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 10% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a multimerization motif. Addition of one or more multimerization motifs to a chimeric protein of a lipid delivery particle described herein may increase a mean base editing percentage between about 30% to about 40%, about 30% to about 50%, about 30% to about 60%, about 30% to about 65%, about 30% to about 70%, about 30% to about 75%, about 30% to about 80%, about 30% to about 85%, about 30% to about 90%, about 30% to about 95%, about 30% to about 99%, about 40% to about 50%, about 40% to about 60%, about 40% to about 65%, about 40% to about 70%, about 40% to about 75%, about 40% to about 80%, about 40% to about 85%, about 40% to about 90%, about 40% to about 95%, about 40% to about 99%, about 50% to about 60%, about 50% to about 65%, about 50% to about 70%, about 50% to about 75%, about 50% to about 80%, about 50% to about 85%, about 50% to about 90%, about 50% to about 95%, about 50% to about 99%, about 60% to about 65%, about 60% to about 70%, about 60% to about 75%, about 60% to about 80%, about 60% to about 85%, about 60% to about 90%, about 60% to about 95%, about 60% to about 99%, about 65% to about 70%, about 65% to about 75%, about 65% to about 80%, about 65% to about 85%, about 65% to about 90%, about 65% to about 95%, about 65% to about 99%, about 70% to about 75%, about 70% to about 80%, about 70% to about 85%, about 70% to about 90%, about 70% to about 95%, about 70% to about 99%, about 75% to about 80%, about 75% to about 85%, about 75% to about 90%, about 75% to about 95%, about 75% to about 99%, about 80% to about 85%, about 80% to about 90%, about 80% to about 95%, about 80% to about 99%, about 85% to about 90%, about 85% to about 95%, about 85% to about 99%, about 90% to about 95%, about 90% to about 99%, or about 95% to about 99% compared to a mean base editing percentage from a lipid delivery particle without a multimerization motif

[0341] In some cases, the multimerization motif comprises a sequence with at least 80% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 281-285. In some cases, the multimerization motif comprises a sequence with at least 90% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 281-285. In some cases, the multimerization motif comprises a sequence with at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to a sequence as set forth in any one of SEQ ID NOs: 281-285. In some cases, the multimerization motif comprises an amino acid

sequence as set forth in any one of SEQ ID NOs: 281-285, with at least one but not more than 8 nucleic acid differences. In some cases, the multimerization motif comprises a sequence as set forth in any one of SEQ ID NOs: 281-285. In some cases, the multimerization motif comprises a sequence as set forth in **Table 9**.

Table 9. Exemplary multimerization motif sequences

SEQ ID NO.	Name	Description	Amino Acid Sequence
281	dDZF1	Antiparallel heterodimerization domain 1 (binds to dDZF2)	FKCEHCRI LFLDHVMFTI HMGCHGFRDPFKCNMC GEKCDGVPVGLFVHMARNAHGEKPFYCEHCEITFR DVVMYSLHKG YHGFRDPFECNICGYHSQDRYEFS SHIVRGEH
282	dDZF2	Antiparallel heterodimerization domain 2 (binds to dDZF1)	HHCQHCDMYFADNII LYTIHMGCHSCDDVFKCNMC GEKCDGVPVGLFVHMARNAHGEKPTKCVHCGIVFL DEVMYALHMSCHGFRDPFECNICGYHSQDRYEFS SHIVRGEH
283	GCN4 Leucine Zipper	Multimerization dimer	RMKQLEDKVEELLSKNYHLENEVARLKKLVGER
284	GCN4 Leucine Zipper	Multimerization trimer	RMKQIEDKIEEILSKIYHIENEIARIKKLIGER
285	GCN4 Leucine Zipper	Multimerization tetramer	RMKQIEDKLEELLSKLYHIENELARIKKLLGER

Table 12. Exemplary Construct Arrangements for Chimeric Protein with Budding Motif

	Arrangement with Added NES		Arrangement with Added CS
PH-BD-Payload	PH-BD-NES-Payload	PH-BD-Payload-NES	PH-BD-NES-CS-Payload
PH-Payload-BD	PH-NES-Payload-BD	PH-Payload-NES-BD	PH-NES-CS-Payload-BD
BD-Payload-PH	BD-NES-Payload-PH	BD-Payload-NES-PH	BD-NES-CS-Payload-PH
BD-PH-Payload	BD-PH-NES-Payload	BD-PH-Payload-NES	BD-PH-NES-CS-Payload
Payload-PH-BD	NES-Payload-PH-BD	Payload-NES-PH-BD	NES-CS-Payload-PH-BD
Payload-BD-PH	NES-Payload-BD-PH	NES-Payload-BD-PH	NES-CS-Payload-BD-PH

Table 13. Exemplary Construct Arrangements for Chimeric Protein with Membrane Penetrating Peptide

	Arrangement with Added NES		Arrangement with Added CS
PH-MPP-Payload	PH-MPP-NES-Payload	PH-MPP-Payload-NES	PH-MPP-NES-CS-Payload
PH-Payload-MPP	PH-NES-Payload-MPP	PH-Payload-NES-MPP	PH-NES-CS-Payload-MPP
MPP-Payload-PH	MPP-NES-Payload-PH	MPP-Payload-NES-PH	MPP-NES-CS-Payload-PH
MPP-PH-Payload	MPP-PH-NES-Payload	MPP-PH-Payload-NES	MPP-PH-NES-CS-Payload
Payload-PH-MPP	NES-Payload-PH-MPP	Payload-NES-PH-MPP	NES-CS-Payload-PH-MPP
Payload-MPP-PH	NES-Payload-MPP-PH	NES-Payload-MPP-PH	NES-CS-Payload-MPP-PH

Table 14. Exemplary Construct Arrangements for Chimeric Protein with Post-Translational Modification

	Arrangement with Added NES		Arrangement with Added CS
PH-PTM-Payload	PH-PTM-NES-Payload	PH-PTM-Payload-NES	PH-PTM-NES-CS-Payload
PH-Payload-PTM	PH-NES-Payload-PTM	PH-Payload-NES-PTM	PH-NES-CS-Payload-PTM
PTM-Payload-PH	PTM-NES-Payload-PH	PTM-Payload-NES-PH	PTM-NES-CS-Payload-PH
PTM-PH-Payload	PTM-PH-NES-Payload	PTM-PH-Payload-NES	PTM-PH-NES-CS-Payload
Payload-PH-PTM	NES-Payload-PH-PTM	Payload-NES-PH-PTM	NES-CS-Payload-PH-PTM
Payload-PTM-PH	NES-Payload-PTM-PH	NES-Payload-PTM-PH	NES-CS-Payload-PTM-PH

Table 15. Exemplary Construct Arrangements for Chimeric Protein with Post-Translational Modification Motif and Budding Motif

	Arrangement with Added NES		Arrangement with Added CS
PTM-PH-BD-Payload	PTM-PH-BD-NES-Payload	PTM-PH-BD-Payload-NES	PTM-PH-BD-NES-CS-Payload
PTM-BD-PH-Payload	PTM-BD-PH-NES-Payload	PTM-BD-PH-Payload-NES	PTM-BD-PH-NES-CS-Payload
PTM-BD-Payload-PH	PTM-BD-NES-Payload-PH	PTM-BD-Payload-NES-PH	PTM-BD-NES-CS-Payload-PH
PTM-Payload-BD-PH	PTM-NES-Payload-BD-PH	PTM-Payload-NES-BD-PH	PTM-NES-CS-Payload-BD-PH
Payload-BD-PH-PTM	NES-Payload-BD-PH-PTM	Payload-NES-BD-PH-PTM	NES-CS-Payload-BD-PH-PTM
Payload-PH-BD-PTM	NES-Payload-PH-BD-PTM	Payload-NES-PH-BD-PTM	NES-CS-Payload-PH-BD-PTM
PH-BD-Payload-PTM	PH-BD-NES-Payload-PTM	PH-BD-Payload-NES-PTM	PH-BD-NES-CS-Payload-PTM
BD-PH-Payload-PTM	BD-PH-NES-Payload-PTM	BD-PH-Payload-NES-PTM	BD-PH-NES-CS-Payload-PTM

Table 16. Exemplary Construct Arrangements for Chimeric Protein with Multimerization Motif

	Arrangement with Added NES		Arrangement with Added CS
(PH/PTM)-(BD-Payload)-LZ	PH/PTM)-(BD-NES-Payload)-LZ	(PH/PTM)-(BD-Payload-NES)-LZ	PH/PTM)-(BD-NES-CS-Payload)-LZ
(PH/PTM)-LZ-(BD-Payload)	(PH/PTM)-LZ-(BD-NES-Payload)	(PH/PTM)-LZ-(BD-Payload-NES)	(PH/PTM)-LZ-(BD-NES-CS-Payload)
(BD-Payload)-LZ-(PH/PTM)	(BD-NES-Payload)-LZ-(PH/PTM)	(BD-Payload-NES)-LZ-(PH/PTM)	(BD-NES-CS-Payload)-LZ-(PH/PTM)
LZ-(BD-Payload)-(PH/PTM)	LZ-(BD-NES-Payload)-(PH/PTM)	LZ-(BD-Payload-NES)-(PH/PTM)	LZ-(BD-NES-CS-Payload)-(PH/PTM)

Nuclear Export Signal

[0342] Direction of nuclear transport within the cell can be governed by nuclear targeting signals within freight proteins. As used herein, the term “nuclear export signal” refers to a sequence of amino acids that targets a freight protein for export from the nucleus. In some cases, a nuclear export signal (NES) is a short target peptide sequence containing four hydrophobic residues. These residues target the protein for export from the nucleus to the cytoplasm through the nuclear pore complex. A chimeric protein provided herein can comprise 1 NES, 2 NESs, 3 NESs, 4 NESs, 5 NESs, 6 NESs, 7 NESs, 8 NESs, 9 NESs, or 10 NESs. In some cases, the NES is located at the N-terminus, C-terminus, or in an internal region of the chimeric protein. In some cases, a NES is coupled between the plasma membrane recruitment element and the payload in the chimeric protein. In some cases, there is a cleavable linker between the plasma membrane recruitment element and the payload in the chimeric protein, and one or more NESs present on the same of the cleavable linker as the plasma membrane recruitment element.

[0343] In some cases, the NES sequence that is used in the chimeric protein comprises LQLPPLERLTL derived from HIV-1 Rev protein, or any of the sequences having at least 80% identity thereto. In some cases, the NES sequence comprises LALKLAGLDI derived from PKI α , or any of the sequences having at least 80% identity thereto. In some cases, the NES sequence that is used in the chimeric protein comprises an amino acid sequence as set forth in **Tables 10A-10E**. In some cases, the NES sequence comprises an amino acid sequence having 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any sequence listed in **Tables 10A-10E**. In some cases, the NES sequence described herein comprises a sequence with greater than 80% sequence identity to any sequence listed in **Tables 10A-10E**. In some cases, the NES sequence that is used in the chimeric protein comprises an amino acid sequence as set forth in any one of SEQ ID NOs: 87-267. In some cases, the NES sequence comprises an amino acid sequence having 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to the sequence listed set forth in any one of SEQ ID NOs: 87-267. In some cases, the NES sequence described herein comprises a sequence with greater than 80% sequence identity to the sequence listed set forth in any one of SEQ ID NOs: 87-267. The transport of payload proteins within a cell is enabled through both NES and nuclear export receptors. In some cases, the NES described herein is associated with a nuclear export receptor (*e.g.*, CRM-1). In some cases, the NES may be conditionally active or inactive. In some cases, the NES sequence disclosed herein comprises a sequence such as those described in T la Cour, et al., *Nucleic Acids Res.* 2003;31(1):393-396; and Xu D, et al. *Mol Biol Cell.* 2012 Sep;23(18):3673-6, each of which is incorporated herein by reference in its entirety. Any of the NES sequences described in the NES

sequence database (NESdb[®]; prodata.swmed.edu/LRNes) or (NESbase; services.healthtech.dtu.dk/datasets/NESbase-1.0) can be used in a chimeric protein disclosed herein, *e.g.*, for the purpose of packaging a payload into the molecular assembly, *e.g.*, the lipid delivery particle.

Table 10A. Exemplary nuclear export signal sequences

SEQ ID NO.	Name	Amino Acid Sequence
87	RSV gag (2x sequence)	KKRKKKKRK
88	RSV gag (3x sequence)	KKRKKKKRKKRK
89	RSV gag (4x sequence)	KKRKKKKRKKRKKRK
90	Human FGF1 (2x sequence)	GNYYKPKLGNYYKPKL
91	Human FGF1 (3x sequence)	GNYYKPKLGNYYKPKLGNYYKPKL
92	HIV REV (NES)	LQLPPLERLTL
93	Influenza A NP	VKGVGTMVMELIRMI
94	RSV gag	LTDWARVREEL
95	Protein Kinase Inhibitor α	LALKLAGLDL
96	BRCA1	QLVEELLKIIICAFQL
97	HTLV 1 Rex Protein	LSAQLYSSLSLD
98	C Myc + GGSG + HIV REV NES	PAAKRVKLDGGGSLQLPPLERLTL
99	C Myc + GGSG + RSV gag NES	PAAKRVKLDGGGSLTDWARVREEL
100	C Myc + GGSG + Protein Kinase Inhibitor α NES	PAAKRVKLDGGGSLALKLAGLDL
101	C Myc + Influenza A NP NES	PAAKRVKLDVKGVGTMVMELIRMI
102	C Myc + BRCA1 NES	PAAKRVKLDQLVEELLKIIICAFQL
103	C Myc + HTLV 1 Rex Protein NES	PAAKRVKLDLSAQLYSSLSLD
104	Human FGF1 + GGSG + HIV REV NES	GNYYKPKLGGGSLQLPPLERLTL
105	Human FGF1 + GGSG + RSV gag NES	GNYYKPKLGGGSLTDWARVREEL
106	Human FGF1 + GGSG + Protein Kinase Inhibitor α NES	GNYYKPKLGGGSLALKLAGLDL
107	Human FGF1 + Influenza A NP NES	GNYYKPKLGGSGVKGVGTMVMELIRMI
108	Human FGF1 + BRCA1 NES	GNYYKPKLGGGSLVEELLKIIICAFQL
109	Human FGF1 + HTLV 1 Rex Protein NES	GNYYKPKLGGGSLSAQLYSSLSLD

Table 10B. Exemplary nuclear export signal sequences

SEQ ID NO.	Name	Amino Acid Sequence
110	Snurportin	MEELSQALASSFSV
111	HIV-REV	PLQLPPLERLTL
112	PKI α	NELALKLAGLDI
113	p53	ERFEMFRELNEALEL
114	RanBP1	DHAEKVAEKLALSV

115	BRCA1	QLVEELLKIICAFQL
116	MAPKK1/MEK1	TNLEALQKKLEEELEL
117	MK2	DVKEEMTSALATMRV
118	Stat1	STNGSLAAEFRHLQL
119	mPER2	PSVQELTEQIHRLLM
120	PLC- δ 1	MNFKELKDFLKELNI
121	p73	ENFEILMKLKESLEL
122	Smad1	FETVYELTKMCTIRM
123	λ PKC	SGKASSSLGLQDFDL
124	Yap1p	PKYSDIDVDGLCSEL
125	Cyclin D1	VDLACTPTDVRDVDI
126	p38	YGEKTTQRDLTELEI
127	E2F4	RRIYDITNVLEGI GL
128	RSV M	AKIIPYSGLLLVITV
129	FAK1	LRSEEVHWLHVDMGV
130		LQSEEVHWLHLDMGV
131	FAK1	LQVRKYSLDLASLIL
132	Vpr	AGVEAIIIRILQQLLF
133		TGVEALIRILQQLLF
134	nsP2	IVLNQLCVRFFGLDL
135	NPM	SLGGFEITPPVVLRL
136	NPM mutA	EAIQDLCLAVEEVSL
137	CHP1	DELLQVLRMMVGVNI
138	E1A	SVMLAVQEGIDLLTF
139	ADAR1	LSSHFAQELS I1
140	E7	QSTHVDIRTLEDLLM
141		ESSAEDLRTLQQLFL
142	N protein	EFSLPTHHTVRLIRV
143	Hxk2	MSSGYYLGEILRLAL
144	Topo I α	DTVLDILRDFEFELRL
145		NSVNEILSEFYVRL
146	Survivin	CAFLSVKKQFEELTL
147	NC2 β	ISPEHVIQALESLGF
148	Tbx5	AHWMRQLVSFQKLKL
149	Paxillin	ATRELDELMASLDF
150	BPV E1	YQNIELITFINALKL
151	CaMKI α	FNATAVVRHMRKLQL
152	STRAD α	SGIFGLVTNLEEEV
153	LEI	EESYTLNSDLARLGV
154		EESYDLTSHLARLGV
155	MK5	GIQQAHAQQLANMRI
156		DVKEEMTSALATMRV
157	Dab1	AAEPVILDRLDFQL
158	p28GANK	MEGCVSNLMV
159		EGCVSNLMV
160	SIRT2	DMDFLRNLFSTLSL
161	Hst2	EQLLEIVHDLNLSL
162	PP2A α	NVMKYFTDLFDYLPPL
163	ORF-9b	KVYPIILRLGSNLSL

Table 10C. Exemplary nuclear export signal sequences

SEQ ID NO.	Amino Acid Sequence	Name
------------	---------------------	------

164	YAGFSLPHAILRIDL	Actin
165	EIVRDIKEKLCYVAL	Actin
166	EAINKLESNLRELQI	c-Abl
167	EAINKLENNLRELQI	
168	SDQKQEQLLLKKMYL	Stat1
169	KQVLWDRTFSLFQQL	Stat1
170	AQLQNLTKRIDSLPL	APC protein
171	NDENEHQLSLRTVSL	NPM
172	ISFTEFVKVLEKVDV	CHP1g
173	MESAITLWQFLLQL	Net
174	VPKELMQQIENFEKI	Hxk2
175	QARFILLEKIDGKII	Topo I α
176	QVKFIKMII EKELTV	
177	NHRMKNLREISQLGI	Gal3
178	NHRVKKLNEISKLGI	
179	TEKHLQKYL RQDLRL	DcpS
180	RQERKRPLLDLHIEL	FLIP-L
181	ANMRIQDLKVSLKPL	MK5
182	ATMRVDYEQIKIKKI	
183	LQGEEFVCLKSII LL	Era
184	THYGQKAILFLPLPV	FGF-1
185	PSAHEITGLADSLQL	6-Oct
186	VRLHDVLHSDKKLTL	Cdk5-p27
187	LINRNDELKLANFGL	Cdk5-p27
188	LEPLKKLECLKSLDL	APRIL

Table 10D Exemplary nuclear export signal sequences

SEQ ID NO.	Amino Acid Sequence	Protein from which NES can be derived
189	MEELS QALASSFSV	Snurportin
190	PLQLPPLERLTL	HIV-REV
191	NELALKLAGLDI	PKI α
192	ERFEMFRELNEALEL	p53
193	DHAEKVAEKLEALSV	RanBP1
194	QLVEELLKI ICAFQL	BRCA1
195	TNLEALQKKLEEELEL	MAPKK1/MEK1
196	DVKEEMTSALATMRV	MK2
197	STNGSLAAEFRHLQL	Stat1
198	PSVQELTEQIHRLLM	mPER2
199	MNFKELKDFLKELNI	PLC- δ 1
200	ENFEILMKLKESLEL	p73
201	FETVYELTKMCTIRM	Smad1
202	SGKASSSLGLQDFDL	λ PKC
203	PKYSDIDVDGLCSEL	Yap1p
204	VDLACTPTDVRDVDI	Cyclin D1
205	YGEKTTQRDLTELEI	p38
206	RRIYDITNVLEGIGL	E2F4
207	AKIIPYSGLLLVITV	RSV M
208	LRSEEVHWHVDMGV	FAK1
209	LQSEEVHWHVDMGV	
210	LQVRKYSLDLASLIL	FAK1
211	AGVEAIIIRILQQLLF	Vpr
212	TGVEALIRILQQLLF	

213	IVLNQLCVRFFGLDL	nsP2
214	SLGGFEITPPVVLRL	NPM
215	EAIQDLCLAVEEVSL	NPM mutA
216	DELLQVLRMMVGVNI	CHP1
217	SVMLAVQEGIDLLTF	E1A
218	LSSHFAQELSI1	ADAR1
219	QSTHVDIRTLLEDLLM	E7
220	ESSAEDLRTLQQFL	
221	EFSLPHTHTVRLIRV	N protein
222	MSSGYLGEILRLAL	Hxk2
223	DTVLDILRDFELRL	Topo Iia
224	NSVNEILSEFYVRL	
225	CAFLSVKKQFEELTL	Survivin
226	ISPEHVIQALESLGF	NC2 β
227	AHWMRQLVSFQKLKL	Tbx5
228	ATRELDELMASLSD	Paxillin
229	YQNIELITFINALKL	BPV E1
230	FNATAVVRHMRKLQL	CaMKI α
231	SGIFGLVTNLEELEV	STRAD α
232	EESYTLNSDLARLGV	LEI
233	EESYDLTSHLARLGV	
234	GIQQAHAEQLANMRI	MK5
235	DVKEEMTSALATMRV	
236	AAEPVILDLRDLFQL	Dab1
237	MEGCVSNLMV	p28GANK
238	EGCVSNLMV	
239	DMDFLRNLFSQLSL	SIRT2
240	EQLEIVHDLENLSL	Hst2
241	NVMKYFTDLFDYLP	PP2A α
242	KVYPILRLGSNLSL	ORF-9b

Table 10E Exemplary nuclear export signal sequences

SEQ ID NO.	Amino Acid Sequence	Protein from which NES can be derived
243	YAGFSLPHAILRIDL	Actin
244	EIVRDIKEKLCYVAL	Actin
245	EAINKLESNLRELQI	c-Abl
246	EAINKLENNLRELQI	
247	SDQKQEQLLLKKMYL	Stat1
248	KQVLWDRTFSLFQQL	Stat1
249	AQLQNLTKRIDSLPL	APC protein
250	NDENEHQSLRTVSL	NPM
251	ISFTEFVKVLEKVDV	CHP1g
252	MESAITLWQFLLQL	Net
253	VPKELMQQIENFEKI	Hxk2
254	QARFILLEKIDGKII	Topo Iia
255	QVKFKMIIIEKELTV	
256	NHRMKNLREISQLGI	Gal3
257	NHRVKKLNEISKLGI	
258	TEKHLQKYLRQDLRL	DcpS
259	RQERKRPLLDLHIEL	FLIP-L
260	ANMRIQDLKVSLKPL	MK5
261	ATMRVDYEQIKIKKI	

262	LQGEEFVCLKSII LL	Era
263	THYGQKAILFLPLPV	FGF-1
264	PSAHEITGLADSLQL	6-Oct
265	VRLHDLVLS DKKLTL	Cdk5-p27
266	LINRNGELKLANFGL	Cdk5-p27
267	LEPLKKLECLKSLDL	APRIL

[0344] In some cases, a chimeric protein disclosed herein includes at least one NES sequences, such as, 2 or more, 3 or more, 4 or more, or 5 or more NES sequences. In some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the N-terminus and/or the C- terminus of the chimeric protein. In some cases, the chimeric protein disclosed herein comprises only one NES sequence. In some cases, the chimeric protein disclosed herein comprises two NES sequences. In some cases, the chimeric protein disclosed herein comprises three NES sequences. In some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the N-terminus of the chimeric protein. In some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the C-terminus of the chimeric protein. In some cases, one or more NES sequences (3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) both the N-terminus and the C-terminus of the chimeric protein. In some cases, an NES sequence is positioned at the N-terminus and an NES sequence is positioned at the C-terminus of the chimeric protein.

[0345] In some cases, a payload is a protein that is delivered as part of the chimeric protein disclosed herein, *e.g.*, operably linked to a structural protein (*e.g.*, human endogenous retroviral structural protein or a Plasma membrane recruitment element). In some embodiments, the one or more NES sequences are positioned at or near the one or both ends of the payload protein sequence inside the chimeric protein. For example, in some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the N-terminus and/or the C- terminus of the payload protein sequence. In some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the N-terminus of the payload protein sequence. In some cases, one or more NES sequences (2 or more, 3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) the C-terminus of the payload protein sequence. In some cases, one or more NES sequences (3 or more, 4 or more, or 5 or more NES sequences) are positioned at or near (*e.g.*, within 50 amino acids of) both the N-terminus and the C-terminus of the payload protein sequence. In some cases, an NES sequence is positioned at the N-terminus and an NES sequence is positioned at the C-terminus of the payload

protein sequence. In some cases, the chimeric protein disclosed herein comprises only one NES sequence. In some cases, the chimeric protein comprises only one NES sequence, and the NES sequence is positioned at or near (*e.g.*, within 50 amino acids of) the N-terminus of the payload protein.

[0346] In eukaryotic cells, transport of proteins between the nucleus and the cytoplasm can be mediated by transport factors in the karyopherin- β family, which are also known as importins and exportins. The direction of nuclear–cytoplasmic transport can be dictated by nuclear targeting signals within the payload proteins. Nuclear export sequences (NESs) can direct export of proteins from the nucleus to the cytoplasm. NESs can bind directly to the export karyopherin CRM1 (also known as exportin 1), which can escort payload proteins through the nuclear pore complex.

Nuclear Localization Signal

[0347] In some instances, a payload described herein comprises one or more nuclear localization sequences (NLS). As used herein, the term “nuclear localization signal” refers to a sequence of amino acids that targets a payload (*e.g.*, a protein or a short polypeptide) to localize to the nucleus. In some cases, an NLS facilitates the import of a polypeptide comprising an NLS into the cell nucleus. A polypeptide may comprise 1 NLS, 2 NLSs, 3 NLSs, 4 NLSs, 5 NLSs, 6 NLSs, 7 NLSs, 8 NLSs, 9 NLSs, or 10 NLSs. In some cases, the NLS is located at the N-terminus, C-terminus, or in an internal region of the polypeptide. In some cases, a NLS is coupled to a nucleic acid binding domain described elsewhere herein. In some cases, a NLS is coupled to a nucleic acid modifying domain described elsewhere herein. In some cases, a NLS is coupled to a guidable polypeptide domain, a deaminase domain, or a reverse transcriptase domain. In some cases, a NLS is covalently linked to a nucleic acid binding domain described elsewhere herein. In some cases, a NLS is covalently linked to a nucleic acid modifying domain described elsewhere herein. In some cases, a NLS is covalently linked to a guidable polypeptide domain, a deaminase domain, or a reverse transcriptase domain. In some cases, a nucleic acid binding domain does not comprise an NLS. In some cases, a nucleic acid binding domain does not comprise an NLS. In some cases, a guidable polypeptide domain, a deaminase domain, or a reverse transcriptase domain does not comprise an NLS. Examples of NLS are provided in **Table 11** below. In some cases, NLS sequence can comprise an amino acid sequence having 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any sequence listed in **Table 11**. In some cases, the NLS sequence described herein can comprise a sequence with greater than 80% sequence identity to any sequence listed in **Table 11**. In some cases, the NLS sequence described herein can comprise any of the sequences listed in **Table 11**. In some cases, NLS sequence can

comprise an amino acid sequence having 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any sequence as set forth in SEQ ID NOs: 268-280.

Table 11. Exemplary Nuclear Localization Signal Sequences

SEQ ID NO.	Name	Amino Acid Sequence
268	Influenza A NP	TKGTRKRSYEQM
269	Influenza A NP	KGINDRNFWRGENGRRTR
270	Nucleoplasmin NLS	KRPAATKKAGQAKKKK
271	Simian Virus	PKKKRKV
272	RSV gag	KKRK
273	Human FGF1	GNYKKPKL
274	Plasmid Derived NLS Variant	KRTADGSEFEPKKKRKV
275	SV40 Large T Antigen	PKKKRKVEDPYC
276	HTLV Rex Protein	PKTRRQTRRRARNRPPT
277	C Myc Protein NLS	PAAKRVKLD
278	Influenza A NP (2x sequence)	TKGTRKRSYEQMTKGTRKRSYEQM
279	Simian Virus (2x sequence)	PKKKRKVPKKKRKV
280	Simian Virus (3x sequence)	PKKKRKVPKKKRKVPPKKRKVPKKKRKV

[0348] In some cases, a chimeric protein disclosed herein includes a nuclear localization sequence (NLS). In some cases, the NLS facilitates delivery of the chimeric protein, or a payload released from the chimeric protein (for instance, released from the chimeric protein following cleavage of a cleavable linker), into the nucleus of a target cell.

[0349] In some cases, a payload is a protein and is delivered as part of the chimeric protein disclosed herein, *e.g.*, operably linked to a structural protein (*e.g.*, plasma membrane recruitment element). In some embodiments, the one or more NLS sequences are positioned at or near the one or both ends of the payload protein sequence of the chimeric protein. In some cases, a chimeric protein includes (*e.g.*, is fused to) between 2 and 5 NLS sequences (*e.g.*, 2-4, or 2-3 NLSs). Examples of NLS sequences include an NLS sequence derived from: the NLS of the SV40 virus large T-antigen, having the amino acid sequence PKKKRKV (SEQ ID NO: 271); the NLS from nucleoplasmin (*e.g.*, the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK (SEQ ID NO: 270); the c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: 277) or RQRRNELKRSP (SEQ ID NO: 435); the hRNPA1M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 436); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 437) of the IBB domain from importin- α ; the sequences VSRKRPRP (SEQ ID NO: 438) and PPKKARED (SEQ ID NO: 439) of the myoma T protein; the sequence PQPKKKPL

(SEQ ID NO: 440) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO: 441) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO: 442) and PKQKKRK (SEQ ID NO: 443) of the influenza virus NS1; the sequence RKLKKKIKKL (SEQ ID NO: 444) of the Hepatitis virus delta antigen; the sequence REKKKFLKRR (SEQ ID NO: 445) of the mouse Mx1 protein; the sequence KRKGDE VDGVDDEV AKKKS KK (SEQ ID NO: 446) of the human poly(ADP-ribose) polymerase; and the sequence RKCLQAGMNLEARKTKK (SEQ ID NO: 447) of the steroid hormone receptors (human) glucocorticoid, and sequences having at least 80% identity to the foregoing. In some cases, an NLS comprises the amino acid sequence MDSLLMNRKFLY QFKNVRWAKGRRETYLC (SEQ ID NO: 448).

[0350] Other examples of an NLS sequence include KRTADGSEFESPKKKRKV (SEQ ID NO: 449), KKTELQTTNAENKTKKL (SEQ ID NO: 450), KRGINDRNFWRGENGRKTR (SEQ ID NO: 451), RSGKIAAIVVKRPRK (SEQ ID NO: 452), and MDSLLMNRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 453), SPKKKRKVEAS (SEQ ID NO: 454), encoded by AGCCCCAAGAAgAAGAGaAAGGTGGAGGCCAGC (SEQ ID NO: 455), GPKKKRKVAAA (SEQ ID NO: 456), as well as any of those described in Cokol *et al.*, *EMBO Rep.*, 2000, 1(5): 411-415 and Freitas *et al.*, *Current Genomics*, 2009, 10(8): 550-7; Lu, J., *et al.*, *Cell Commun Signal* 19, 60 (2021); international publication no. WO/2001/038547, each of which is incorporated herein by reference in its entirety, and sequences having at least 80% identity to the foregoing.

[0351] In some embodiments, the chimeric protein comprises one NES sequence and two NLS sequences. In some cases of these embodiments, the NES sequence, NLS sequences, and the payload protein sequence are positioned in an order from N-terminus to C-terminus as follows: NES-NLS-payload protein-NLS. In some embodiments, the chimeric protein comprises two or more NES sequences and two NLS sequences. In some cases of these embodiments, the NES sequences, NLS sequences, and the payload protein sequence are positioned in an order from N-terminus to C-terminus as follows: n X NES (n >=2)-NLS-payload protein-NLS.

Cleavable Linker

[0352] In some cases, the chimeric protein comprises a cleavable linker in between two or more components. For instance, the chimeric protein can comprise a cleavable linker between a payload protein sequence and a plasma membrane recruitment element sequence (*e.g.*, retroviral gag protein sequence). In some cases, the cleavable linker separates the plasma membrane recruitment element sequence from a NLS sequence, and/or a NES sequence at its N-terminus or C-terminus. The cleavable linker can separate the payload protein sequence from the plasma membrane recruitment element sequence, NLS sequence, and/or NES sequence at its N-terminus or C-

terminus. Examples of cleavable linker sequences that can be used in the chimeric protein include TSTLLMENSS (SEQ ID NO: 431), PRSSLYPALTP (SEQ ID NO: 432), VQALVLTQ (SEQ ID NO: 433), and PLQVLTLNIERR (SEQ ID NO: 434), and sequences having at least 80% identity to any one of the foregoing.

[0353] The cleavable linker sequence provided herein can be a cleavable sequence that is recognized and cleaved by any applicable protease, such as a viral protease, a bacterial protease, or a eukaryotic protease (*e.g.*, a protease derived from a plant, an animal, or a fungus). In some cases, the cleavable sequence is recognized by a retroviral protease (*pro*), such as *pro* protein derived from Moloney murine leukemia virus (MMLV) or Friend murine leukemia virus (FMLV). In some cases, the cleavable sequence is recognized by a mammalian endogenous protease (*e.g.*, furin, cathepsin, metalloprotease, ADAM10, or presenilin). In some cases, the lipid delivery particle further comprises a protease that recognizes the cleavable linker sequence, such as *pro* protein derived from Moloney murine leukemia virus (MMLV) or Friend murine leukemia virus (FMLV), or protease that is of other viral origin, bacterial origin, or eukaryotic origin. Examples of cleavable linker sequences that can be present in the chimeric protein include TSTLLMENSS, PRSSLYPALTP, VQALVLTQ, and PLQVLTLNIERR, as well as variant sequences having at least 80% identity to the foregoing sequences.

NUCLEIC ACIDS

[0354] Nucleic acids encoding the aforementioned lipid delivery particles are also disclosed.

[0355] In some cases, the invention features a nucleic acid comprising nucleotide sequences that encode the lipid delivery particle, plasma membrane recruitment element, or other component of the chimeric protein disclosed herein. The nucleic acid molecule can encode one, two, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif; and/or (iv) a multimerization motif.

[0356] In some cases, the nucleic acid is an isolated nucleic acid or a non-naturally occurring nucleic acid. Non-naturally occurring nucleic acids are well known to those of skill in the art. In some cases, the nucleic acid is an *in vitro* transcribed nucleic acid. The term “isolated” means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell. In some cases, provided herein is an isolated nucleic acid molecule encoding a lipid delivery particle comprising a lipid membrane, a plasma membrane recruitment element, and a chimeric protein comprising a particle budding motif, a membrane penetrating

peptide, and/or a post-translational modification motif. In some cases, provided herein is an isolated nucleic acid molecule encoding a chimeric protein comprising a particle budding motif, a membrane penetrating peptide, and/or a post-translational modification motif.

[0357] In some cases, the nucleic acid molecule is a recombinant nucleic acid molecule. In some cases, provided herein is a recombinant nucleic acid molecule encoding a lipid delivery particle comprising a lipid membrane, a plasma membrane recruitment element, and a chimeric protein comprising a particle budding motif, a membrane penetrating peptide, and/or a post-translational modification motif. In some cases, provided herein is a recombinant nucleic acid molecule encoding a chimeric protein comprising a particle budding motif, a membrane penetrating peptide, and/or a post-translational modification motif. In some instances, the recombinant nucleic acid further comprises a leader sequence. In some instances, the recombinant nucleic acid further comprises a promoter sequence. In some instances, the recombinant nucleic acid further comprises a sequence encoding a poly(A) tail. In some instances, the recombinant nucleic acid further comprises a 3'UTR sequence.

[0358] In some cases, the lipid delivery particle is encoded by a messenger RNA (mRNA). In one aspect the mRNA encoding the lipid delivery particle is introduced into a cell for production of a lipid delivery particles. In one embodiment, the in vitro transcribed RNA lipid delivery particle can be introduced to a cell as a form of transient transfection. The RNA is produced by in vitro transcription using a polymerase chain reaction (PCR)-generated template. DNA of interest from any source can be directly converted by PCR into a template for in vitro mRNA synthesis using appropriate primers and RNA polymerase. The source of the DNA can be, for example, genomic DNA, plasmid DNA, phage DNA, cDNA, synthetic DNA sequence or any other appropriate source of DNA. The desired template for in vitro transcription is a lipid delivery particle of the present disclosure. In one embodiment, the DNA to be used for PCR contains an open reading frame. The DNA can be from a naturally occurring DNA sequence from the genome of an organism. In one embodiment, the nucleic acid can include some or all of the 5' and/or 3' untranslated regions (UTRs). The nucleic acid can include exons and introns.

[0359] In some cases, the chimeric protein described herein is encoded by a messenger RNA (mRNA). In one aspect the mRNA encoding the chimeric protein is introduced into a cell for production of a lipid delivery particles comprising the chimeric protein. In one embodiment, the in vitro transcribed RNA chimeric protein can be introduced to a cell as a form of transient transfection. The RNA is produced by in vitro transcription using a polymerase chain reaction (PCR)-generated template. DNA of interest from any source can be directly converted by PCR into a template for in vitro mRNA synthesis using appropriate primers and RNA polymerase. The source of the DNA can be, for example, genomic DNA, plasmid DNA, phage DNA, cDNA,

synthetic DNA sequence or any other appropriate source of DNA. The desired template for in vitro transcription is a chimeric protein of the present disclosure. In one embodiment, the DNA to be used for PCR contains an open reading frame. The DNA can be from a naturally occurring DNA sequence from the genome of an organism. In one embodiment, the nucleic acid can include some or all of the 5' and/or 3' untranslated regions (UTRs). The nucleic acid can include exons and introns.

[0360] In another aspect, the application features producer cells and vectors containing the nucleic acids described herein. The nucleic acids may be present in a single vector or separate vectors present in the same host cell or separate host cell.

VECTORS

[0361] In various aspects, the lipid delivery vesicles can be embodied as plasmid vectors useful for transient or permanent protein expression in mammalian cells.

[0362] Further provided herein are vectors comprising the compositions or nucleotide sequences encoding a lipid delivery particle or a chimeric protein described herein. In some cases, the vectors comprise nucleotides encoding a lipid delivery particle described herein. In some cases, the vectors comprise nucleotides encoding a chimeric protein described herein. The vectors include, but are not limited to, a virus, plasmid, cosmid, lambda phage or a yeast artificial chromosome (YAC).

[0363] Numerous vector systems can be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as, for example, bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (Rous Sarcoma Virus, MMTV or MOMLV) or SV40 virus. Another class of vectors utilizes RNA elements derived from RNA viruses such as Semliki Forest virus, Eastern Equine Encephalitis virus and Flaviviruses. In some cases, the vector is selected from the group consisting of a DNA, a RNA, a plasmid, a lentivirus vector, adenoviral vector, or a retrovirus vector. In some cases, the vector further comprises a promoter. In some cases, the vector is an in vitro transcribed vector. In some cases, a nucleic acid sequence in the vector further comprises a poly(A) tail. In some cases, a nucleic acid sequence in the vector further comprises a 3'UTR.

[0364] Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression vectors may be transfected or introduced into an appropriate host cell. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation, retroviral transduction, viral transfection, gene gun, lipid based transfection or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity.

[0365] Methods and conditions for culturing the resulting transfected cells and for recovering the antibody molecule produced are known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed, based upon the present description.

CELLS

[0366] In another aspect, described herein are producer cells comprising the lipid delivery particles described herein. In some cases, the producer cells comprise a nucleic acid sequence encoding a lipid delivery particle described herein. The nucleic acids may be present in a single vector or separate vectors present in the same host cell or separate host cell. The host cell can be a eukaryotic cell, *e.g.*, a mammalian cell, an insect cell, a yeast cell, or a prokaryotic cell, *e.g.*, *E. coli*. For example, the mammalian cell can be a cultured cell or a cell line. Exemplary mammalian cells include lymphocytic cell lines (*e.g.*, NSO), Chinese hamster ovary cells (CHO), COS cells, oocyte cells, and cells from a transgenic animal, *e.g.*, mammary epithelial cell.

[0367] The disclosure also provides producer cells comprising a nucleic acid encoding a chimeric protein as described herein. In some cases, the producer cells are genetically engineered to comprise nucleic acids encoding the antibody molecule.

[0368] In some cases, the producer cells are genetically engineered by using an expression cassette. The phrase “expression cassette,” refers to nucleotide sequences, which are capable of affecting expression of a gene in hosts compatible with such sequences. Such cassettes may include a promoter, an open reading frame with or without introns, and a termination signal. Additional factors necessary or helpful in effecting expression may also be used, such as, for example, an inducible promoter.

[0369] The present also provides producer cells comprising the vectors described herein.

[0370] The cell can be, but is not limited to, a eukaryotic cell, a bacterial cell, an insect cell, or a human cell. Suitable eukaryotic cells include, but are not limited to, Vero cells, HeLa cells, COS cells, CHO cells, HEK293 cells, BHK cells and MDCKII cells. Suitable insect cells include, but are not limited to, Sf9 cells. In some cases, the producer cell is a 293T Gesicle cell or a 293FT cell.

[0371] Also provided herein are cells (*e.g.*, isolated cells, preferably mammalian, *e.g.*, human, cells) that express, *e.g.*, that have been induced to overexpress, components of the lipid delivery particles and/or chimeric proteins described herein (*e.g.*, over expressed from exogenous sources, such as plasmids or stably integrated transgenes), and a payload fused to a human endogenous GAG, a PH domain, or alternative plasma membrane recruitment element. In some cases, the cell overexpresses a chimeric protein with one, two, or all of: a particle budding motif, a membrane

penetrating peptide, and/or a post-translational modification motif. In some cases, the cell does not express a gag protein except for gag proteins that are encoded in the human genome or gag proteins that are encoded by a consensus sequence that is derived from gag proteins found in the human genome (overexpressed from exogenous sources, such as plasmids or stably integrated transgenes). In some cases, the cells are primary or stable human cell lines, e.g., Human Embryonic Kidney (HEK) 293 cells, HEK293 T cells, or BeWo cells.

PRODUCTION OF LIPID DELIVERY PARTICLES

[0372] In some aspects, provided herein are composition, methods of production, methods of purification related to the lipid delivery particles provided herein. In some cases, the lipid delivery particles can be produced from producer cell lines that are either transiently transfected with at least one plasmid or stably expressing constructs that have been integrated into the producer cell line genomic DNA.

[0373] Producer cell lines can be generated by stably integrating genetic material with a gene of interest into a host cell line. In some cases, the genetic material is transiently expressed in a producer cell line. In some cases, the genetic material is expressed via viral methods. In some cases, the genetic material is expressed via non-viral methods. In some cases, a producer cell line grows in a serum-free medium or in suspension. A producer cell line can be grown in serum-free medium and suspension simultaneously. In some cases, producer cell lines can be generated with adherent cells (*e.g.*, cells cultured in media and attached to a substrate).

[0374] Producer cells can be used to produce the lipid delivery particles described herein. In some cases, generating a producer cell line comprises transfecting cells (*e.g.*, cells of a mammalian cell type) with genetic material of the present disclosure, culturing the cells to produce the lipid delivery particles, obtaining a media from the mammalian cell producing the lipid delivery particles, collecting and filtering the harvested media, and, optionally, purifying the lipid delivery particles to retain structural integrity. In some cases, the method of producing the lipid delivery particle further comprises providing new media to promote transient production of the lipid delivery particles. In some cases, the mammalian cell type includes a HT1080 cell, a COS cell, a HeLa cell, a Chinese Hamster Ovary (CHO) cell, or a HEK 293 cell. HEK293 cells are cells derived from human embryonic kidney cells grown in tissue culture. In some cases, the HEK293 cell is a HEK293, 293E, 293T, 293F, 293FT, or 293T Gesicle cell. The producer cell line can be transformed with a viral vector or non-viral method in any number of means including calcium phosphate and the like.

[0375] Following transfection, the cells can be cultured under conditions for production of lipid delivery particles. Exemplary culturing conditions can include refeeding cells in appropriate

media, addition of CO₂, and humidity. In some cases, culturing conditions includes addition of antibiotics, anti-fungals, and/or growth factors. The medium can be harvested after 24, 48, 72, or 96 hours, or at any appropriate time point to allow sufficient production of the lipid delivery particles.

[0376] Optionally, the lipid delivery particles in the media can be isolated and collected using any number of techniques known in the art. In some cases, the lipid delivery particles are purified, wherein the lipid delivery particles are washed or resuspended in an appropriate buffer or media or at particular concentration.

[0377] In an aspect, disclosed herein are methods of manufacturing producer cell lines that comprise the lipid delivery particles of the present disclosure. Adherent cells can be first transfected to produce lipid delivery particles. In some cases, transfection occurs by the addition or expression of exogenous nucleic acid sequences via non-viral methods (*e.g.*, by electroporation, microinjection, or a chemical system such as DEAE-dextran or cationic polymers). In some cases, transfection occurs by the addition or expression of exogenous nucleic acid sequences via viral methods (*e.g.*, by infecting the cells with a viral vector, such as an adenoviral vector, adeno-associate viral vector, a lentiviral vector, a herpes viral vector, or a HSV vector). In some cases, the cells are from a HEK293 cell line (*e.g.*, HEK293, 293E, or 293T). In some cases, to transfect DNA into the host cells, the cells are cultured in a medium. In some cases, cells can be cultured in the medium for 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25 hours. In some cases, cells can be cultured in the medium for between 10-20 hours. In some cases, cells can be cultured in the medium for 18 hours.

[0378] Following incorporation into the transfection medium, cells are transferred to a new solution. In some cases, the new solution is new media. In some cases, the new media promotes the production of the lipid delivery particles. In some cases, the cells incorporate into the new media for between 10-50 hours. In some cases, the cells incorporate into the new media for 10, 20, 30, 35, 40, 45, or 50 hours. In some cases, the cells incorporate into the new media for 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 or 40 hours. The media can then be harvested. The harvested media can be filtered, and the lipid delivery particles can be collected. Filtration can comprise microfiltration and/or depth filtration. In some cases, the lipid delivery particles can undergo further purification and/or concentration methods that maintain the structural integrity of the particles.

[0379] In some aspects, provided herein is a method of loading a lipid delivery particle with components such as a payload. RNA and protein from a producer cell can get packaged and/or incorporated into lipid delivery vehicles of the present disclosure. In some cases, the components of the lipid delivery particles, such as a payload, is loaded via the packaging and assembly process

of the lipid delivery particle. For instance, the payload can be a polypeptide or protein that is packaged into the lipid delivery particle as a part of a chimeric protein as disclosed herein.

[0380] In some cases, the payload is assembled into the lipid delivery particle as an independent entity, *e.g.*, not as a part of a chimeric protein. In other embodiments, the lipid delivery particle provided herein is loaded with a payload by utilizing any suitable method for delivering a biological or chemical payload through a lipid membrane, such as nucleofection, electroporation, lipid-based, polymer-based, or CaCl_2 transfection, sonication, freeze thaw, incubation at various temperatures, or heat shock of lipid delivery particles mixed with payload. In some cases, the nucleic acid molecules, such as a template RNA described herein, are loaded into the lipid delivery particle by direct loading, such as electroporation of the lipid delivery particle *in vitro*. In some cases, the nucleic acid molecules are loaded into the lipid delivery particle by binding to a nucleic acid binding protein (*e.g.*, Cas protein) that is part of the lipid delivery particle or is already loaded into the lipid delivery particle.

[0381] There can be more than one type of loading techniques utilized for loading payloads (*e.g.*, for loading more than one type of payloads) into the lipid delivery particle. For instance, in some cases, a first payload is a polypeptide that is assembled into the lipid delivery particle as a part of a chimeric protein, and a second payload is a separate protein or nucleic acid (RNA or DNA) that interacts with (*e.g.*, binds) the first payload, and thus is loaded into the lipid delivery particle via the interaction between the first payload and the second payload. Alternatively, the second payload can be loaded into the lipid delivery particle via a transfection-like technique or any other suitable method.

[0382] In aspects, also provided herein are methods of using a lipid delivery particle or pharmaceutical composition according to some embodiments of the present disclosure, comprising contacting a cell with the lipid delivery particle described herein. In some cases, the cell is a mammalian cell, such as a human cell. In some cases, the cell is within a subject in need of treatment for a disease or a condition. In some cases, contact comprising administering the lipid delivery particle described herein to the subject, such as via injections.

[0383] In aspects, also provided herein are methods of administering a lipid delivery particle, systems, or pharmaceutical compositions according to some embodiments of the present disclosure. In some cases, the method comprises administering the lipid delivery particle, system, or pharmaceutical composition described herein to a subject in need thereof, such as via injections.

[0384] In aspects, also provided herein are methods of producing a lipid delivery particle or pharmaceutical composition according to some embodiments of the present disclosure. In some cases, the method comprises contacting a producer cell with compositions described herein.

Methods of Purification

[0385] In an aspect, described herein are methods of purifying lipid delivery particles. In some cases, the lipid delivery particles are produced from producer cell lines that are either transiently transfected with at least one plasmid or stably expressing constructs that have been integrated into the producer cell line genomic DNA. In some cases, the producer cell culture medium is harvested 24-, 48-, 72-, or 96-hours post-transfection. In some cases, the producer cell culture medium is harvested between 40- and 48-hours post-transfection. The harvested medium can undergo centrifugation steps to remove producer cell debris while maintaining the structural integrity of the lipid delivery particle. In some cases, during harvesting, the producer cell medium is centrifuged, *e.g.*, at 500g for 5 minutes. The clarified lipid delivery particle containing supernatant can then be collected and filtered. In some cases, the lipid delivery particles are further concentrated. In some cases, the lipid delivery particles are further concentrated by ultracentrifugation. In some cases, the lipid delivery particles are concentrated 50-fold, 100-fold, 200-fold, 500-fold, 1000-fold, 2000-fold, 3000-fold, or 5000-fold. In some cases, the concentrated lipid delivery particles are resuspended, *e.g.*, in cold PBS. In some cases, the concentrated lipid delivery particles are frozen, *e.g.*, frozen at a rate of $-1^{\circ}\text{C}/\text{min}$ and stored at -80°C .

[0386] In some cases, the purification methods can comprise chromatographic methods (*e.g.*, anion exchange chromatography), ultrafiltration methods (*e.g.*, tangential flow filtration), clarifying normal flow filtration, and/or sterilizing membrane filtration. Anion exchange chromatography can separate substances based on net-surface charge, using an ion-exchange resin. Tangential flow filtration can separate molecules using ultrafiltration membranes. In some cases, the membrane pore size used for tangential flow filtration can retain a biological product of a size less than 1000 kDa, less than 750 kDa, less than 500 kDa, less than 250 kDa, less than 200 kDa, less than 150 kDa, less than 100 kDa, or less than 50 kDa. Normal flow filtration assists in the clarification of biofluid by convecting the substance directly toward a membrane under an applied pressure. In some cases, normal flow filtration can comprise a membrane pore size of greater than 0.1 μm , greater than 0.2 μm , greater than 0.3 μm , greater than 0.4 μm , greater than 0.5 μm , greater than 0.6 μm , greater than 0.7 μm , greater than 0.8 μm , greater than 0.9 μm , greater than 1.0 μm , greater than 1.5 μm , or greater than 2.0 μm . In some cases, normal flow filtration can comprise a membrane pore size of 0.2 μm , 0.45 μm , 0.8 μm , 1.2 μm , or 2.0 μm . Sterilizing membrane filtration can be used to sterilize heat-sensitive liquid without exposure to denaturing heat. In some cases, sterilizing membrane filtration can comprise a membrane pore size of about 0.1 μm , about 0.2 μm , about 0.3 μm , about 0.4 μm , or about 0.5 μm . In some cases, sterilizing membrane filtration can comprise a membrane pore size of about 0.2 μm or 0.22 μm .

COMPOSITIONS AND SYSTEMS

[0387] In aspects, also provided herein are nucleic acid molecules that encode one or more of the components of the lipid delivery particles of the present disclosure. For instance, a nucleic acid molecule encoding the chimeric protein is provided. A nucleic acid molecule encoding the envelope protein is also provided.

[0388] In aspects, provided herein are compositions or systems that include nucleic acid molecules that encode one or more of the components of the lipid delivery particles of the present disclosure.

[0389] The compositions or systems can be used for producing a lipid delivery particle of the present disclosure, for instance, by transfecting or otherwise delivering the nucleic acid molecules in the compositions or systems into a producer cell. The nucleic acid molecules can be expressed in the producer cell, the result of which assemble, package, and subsequently cause the producer cell to release the lipid delivery particle.

[0390] In some cases, a lipid delivery particle of the present disclosure facilitates gene editing efficiency greater than 40%, greater than 50%, greater than 60%, or more. In some cases, a lipid delivery particle of the present disclosure facilitates gene editing efficiency greater than 70%. In some cases, a lipid delivery particle of the present disclosure facilitates gene editing efficiency comprising 8-fold increase of base editing efficiency when compared to a conventional VLP (*e.g.*, the VLPs described in Mangeot, P. E. *et al.* Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat. Commun.* **10**, 45 (2019).). In some cases, a lipid delivery particle of the present disclosure facilitates gene editing efficiency comprising 8-fold increase of prime editing efficiency when compared to a conventional VLP. In some cases, a lipid delivery particle of the present disclosure exhibits reduced immunogenicity in transduced target cells. In some cases, a lipid delivery particle of the present disclosure produces reduced off-target genome editing in target cells when delivering genome editing system into the target cells when compared to a conventional VLP. In some cases, a lipid delivery particle of the present disclosure leads to more than 100-fold reduction in Cas-independent off-target editing when compared to a conventional VLP. In some cases, a lipid delivery particle of the present disclosure leads to at least 10-fold, such as 12- to 900-fold, lower Cas-dependent off-target editing when compared to a conventional VLP.

PHARMACEUTICAL FORMULATIONS

[0391] In an aspect, disclosed herein is a pharmaceutical formulation comprising the lipid delivery particle disclosed herein and optionally further comprising a pharmaceutically acceptable carrier,

excipient, or additive. The term “pharmaceutical formulation”, as used herein, refers to a composition formulated for pharmaceutical use. The terms such as “excipient,” “carrier,” “pharmaceutically acceptable carrier” or the like are used interchangeably herein. Pharmaceutical formulations comprise an immunologically effective amount of one or more cells, vectors, lipid delivery particles, or compositions disclosed herein, and optionally one or more other components which are pharmaceutically acceptable. In some cases, the pharmaceutical formulation comprises additional agents, *e.g.*, for specific delivery, increasing half-life, or other therapeutic benefit. In some cases, the pharmaceutical formulation may comprise one or more of dimethylsulfoxide (DMSO), dextrose, water, succinate, poly I: poly C, poly-L-lysine, carboxymethylcellulose, and/or chloride.

[0392] As used herein, a “pharmaceutically acceptable carrier” is an agent that is compatible with the other ingredients of the formulation and not injurious to the tissue of the subject (*e.g.*, physiologically compatible, sterile, physiologic pH, etc.) In some cases, a pharmaceutically acceptable carrier comprises any vehicle, such as a liquid or solid filler, diluent, excipient, manufacturing aid (*e.g.*, lubricant, talc magnesium, calcium or zinc stearate, or steric acid), or solvent encapsulating material, involved in carrying or transporting the compound from one site (*e.g.*, the delivery site) of the body, to another site (*e.g.*, organ, tissue or portion of the body).

[0393] Some exemplary materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanthin; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids; (23) serum alcohols, such as ethanol; and (24) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation.

[0394] Pharmaceutical formulation disclosed herein can comprise one or more pH buffering compounds to maintain the pH of the formulation at a predetermined level that reflects

physiological pH, such as in the range of about 5.0 to about 8.0. The pH of the pharmaceutical formulation can be about 4, about 5, about 6, about 7, about 8 or about 9. The pH buffering compound used in the aqueous liquid formulation can be an amino acid or mixture of amino acids, such as histidine or a mixture of amino acids such as histidine and glycine. The pH buffering compound can be an agent which does not chelate calcium ions. Exemplary pH buffering compounds include imidazole and acetate ions. The pH buffering compound can be present in any amount suitable to maintain the pH of the formulation at a predetermined level.

[0395] The pharmaceutical formulations described herein can be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient(s) into association with an excipient and/or one or more other accessory ingredients, and then, optionally, shaping and/or packaging the product into a desired single- or multi-dose unit. Pharmaceutical formulations can additionally comprise a pharmaceutically acceptable excipient, which, as used herein, includes any and all solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, solid binders, lubricants, and the like, as suited to the particular dosage form desired.

METHOD OF TREATMENT, PREVENTION, OR DIAGNOSIS

[0396] A lipid delivery particle provided herein can find use in a variety of fields and methods. In some cases, the lipid delivery particle of the present disclosure can be used to deliver one or more payloads, such as a ribonucleoprotein complex to a cell. In some cases, the target cells to which the lipid delivery particles are delivered are *in vitro* cells, *ex vivo* cells, or *in vivo* cells. The lipid delivery particles of the present disclosure can be applicable for delivery of freights into a variety of cell types, such as, animal cells, plant cells, bacteria cells, algal cells, or fungal cells.

[0397] In aspects, also provided herein are methods of treating a subject by administering a lipid delivery particle described herein, a system described herein, a composition described herein, or pharmaceutical composition according to some embodiments of the present disclosure.

[0398] In some cases, the present disclosure provides methods of treating, preventing, or diagnosing a condition, disease, or disorder. In some cases, a composition, kit, or method described herein can be used to treat, prevent, or diagnose a condition, disease, or disorder. The condition, disease, or disorder can comprise a cancer, an immune disorder, an autoimmune disorder, a metabolic disorder, a hormonal disorder, an inflammatory disorder, a developmental disorder, a reproductive disorder, an imprinting disorder, a genetic disorder, a neurological disorder, or a neurodegenerative disorder. In some cases, the condition, disease, or disorder comprises a liver disorder, an eye disorder, a heart disorder, a kidney disorder, a skin disorder, a blood disorder, a fibrotic disorder, a skeletal disorder, or a muscle disorder. In some cases, the condition, disease, or disorder is caused by a genetic mutation (*e.g.*, an insertion, deletion,

or point mutation). In some cases, the condition, disease, or disorder is hereditary. In some cases, the condition, disease, or disorder is caused by a virus or bacteria or fungus. In some cases, the condition, disease, or disorder is caused by aberrant gene expression. In some cases, the condition, disease, or disorder is a result of age. In some embodiments, the condition, disease, or disorder is chronic.

[0399] The subject in the method of present disclosure can be an animal. In some embodiments, the subject is an animal cell. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human. In some embodiments, the subject is an aquaculture animal (fish, crabs, shrimp, oysters etc.), a mammal. In some embodiments, the animal's cell is from, for example, a pet or zoo animal (cats, dogs, lizards, birds (*e.g.*, parrots), lions, tigers and bears etc.), from a farm or working animal (horses, cows (*e.g.*, dairy and beef cattle) pigs, chickens, turkeys, hens or roosters, goats, sheep, etc.), or a human. In some embodiments, the target cell as disclosed herein is in a subject to whom the method of the present disclosure is applicable.

[0400] The methods described herein can be therapeutic or veterinary methods for treating a subject. In some embodiments, the methods described herein are used to treat a disease resulting from a non-functional, poorly functional, or poorly expressed protein or gene product, for instance, resulting from a genetic mutation in one or more cells of the subject. In some embodiments, the methods described herein are used to treat a genetic disease (*e.g.*, a mutation, a substitution, a deletion, an expansion, or a recombination), a monogenic disease, an inherited metabolic disease, a cancer, a neurodegenerative disease, a cardiovascular disease, a pulmonary disease, a renal disease, a liver disease, a genetic disease, a vascular disease, ophthalmic disease, musculoskeletal disease, lymphatic disease, auditory and inner ear disease, a metabolic disease, an inflammatory disease, an autoimmune disease, or an infectious disease. In some cases, provided herein are pharmaceutical compositions and methods for treating a retinal disease, *e.g.*, Leber congenital amaurosis, by administering a pharmaceutical composition formulated for subretinal injection.

KIT

[0401] In one aspect, disclosed herein are kits comprising the lipid delivery particles, the compositions, or the pharmaceutical formulation of the present disclosure. In some cases, the kit comprises the lipid delivery particles, compositions, or pharmaceutical formulations of the present disclosure; and an informational medium containing instructions for administering the lipid delivery particle, composition, or pharmaceutical formulation to a subject. The kit can include a label indicating the intended use of lipid delivery particle, composition, or pharmaceutical formulation in the kit. Label can include any writing, or recorded material supplied on or with the kit, or which otherwise accompanies the kit.

[0402] A kit of the present disclosure can include, alternatively or additionally, diagnostic agents and/or other therapeutic agents. In some cases, the kit includes cells or pharmaceutical formulations of the present disclosure and a diagnostic agent that can be used in a diagnostic method for diagnosing a condition, disease, or disorder in a subject.

[0403] In some cases, the kit comprises a vector, composition, cell, or formulation as described herein. In some cases, the kit is useful for producing a vector, composition, cell, or formulation as described herein. In some cases, the kit further comprises cell culture media.

[0404] In some cases, a kit described herein includes one or more additional active ingredients, pharmaceutically acceptable carrier, diluent, excipient or adjuvant. The additional active agent may include immune stimulating cytokines (e.g., IL-2, IFN α 2, GM-CSF), targeted small molecules and biological molecules (such as components of signal transduction pathways, e.g. modulators of tyrosine kinases and inhibitors of receptor tyrosine kinases), an anti-inflammatory agent, a cytotoxic agent, a radiotoxic agent, or an immunosuppressive agent.

[0405] The ability of an adjuvant to increase an immune response to an antigen is typically manifested by a significant or substantial increase in an immune-mediated reaction, or reduction in disease symptoms.

[0406] The pharmaceutical acceptable carrier, diluent, excipient or adjuvant may include; sterile diluents such as water, saline solution, preferably physiological saline, Ringer's solution, isotonic sodium chloride, fixed oils such as synthetic mono or diglycerides, polyethylene glycols, glycerin, or other solvents; antibacterial agents such as benzyl alcohol or methyl paraben; and agents for the adjustment of tonicity such as sodium chloride or dextrose.

[0407] The kits can be compartmentalized for ease of use and can include one or more containers with reagents. In certain cases, all of the kit components are packaged together. Alternatively, one or more individual components of the kit can be provided in a separate package from the other kits components. The kits can also include instructions for using the separate kit components.

METHODS OF ADMINISTERING

[0408] In some cases, the composition or pharmaceutical formulation described herein is prepared for administration to a subject. In some cases, the pharmaceutical formulation is prepared to induce a therapeutic or prophylactic effect in a subject. Suitable routes of administering the pharmaceutical formulation described herein include transdermal, intravesical, intravenous, intravascular, intraosseous, topical, subcutaneous, intradermal, intralesional, intraarticular, intraperitoneal, transmucosal, gingival, intradental, intracochlear, transtympanic, intraorgan, epidural, intrathecal, intramuscular, periocular, intratumoral, intracerebral, intravitreal, and intracerebroventricular administration. In some cases, the pharmaceutical formulation described

herein is administered locally to a diseased site (*e.g.*, site of infection or tumor site). In some cases, the pharmaceutical composition described herein is delivered in a controlled release system. In some cases, a pump is used. In some cases, polymeric materials is used for controlled release. In some cases, the pharmaceutical composition described herein is administered to a subject by injection, by means of a catheter, by means of a suppository, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including a membrane, such as a sialastic membrane, or a fiber. In some cases, the pharmaceutical formulation is formulated in accordance with routine procedures as a formulation adapted for intravenous or subcutaneous administration to a subject. In some cases, pharmaceutical formulations for administration by injection are solutions in sterile isotonic use as solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients can be supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachets indicating the quantity of active agent. In some cases, if the pharmaceutical is to be administered by infusion, it is dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. In some cases, if the pharmaceutical formulation is administered by injection, an ampoule of sterile water for injection or saline is provided so that the ingredients can be mixed prior to administration.

[0409] A pharmaceutical formulation as described herein can be administered or packaged as a unit dose, for example, in reference to a pharmaceutical formulation to physically discrete units suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent, carrier, or vehicle.

EXAMPLES

[0410] The following examples are provided to further illustrate some embodiments of the present disclosure, but are not intended to limit the scope of the disclosure; it will be understood by their exemplary nature that other procedures, methodologies, or techniques known to those skilled in the art may alternatively be used.

Example 1: Identification and Selection of Constructs for the Chimeric Protein

[0411] Initial candidates for the chimeric protein of the lipid delivery particles will be selected based on different arrangements of payload and plasma membrane recruitment elements. Exemplary arrangements of constructs for the chimeric protein can be found in **FIGs. 1A-1F**.

[0412] PH domain is abbreviated PH, particle budding domain is abbreviated BD, and a protease cleavage site is abbreviated CS.

Example 2: Manufacture of Lipid Delivery Particles

[0413] This is a series of steps for the production of lipid delivery particles. Lipid delivery particles will be generated using steps shown in **FIG. 2**.

[0414] Adherent cells will be first transfected to produce lipid delivery particles (*e.g.*, ectosomes). A nucleic acid molecule will be generated that encodes a chimeric protein that comprises a PH domain, a payload protein, and one, two, or all of: (i) a particle budding motif; (ii) a membrane penetrating peptide; (iii) a post-translational modification motif; and/or (iv) a multimerization motif. In this example, the nucleic acid molecule encodes a chimeric protein comprising a PH domain and a particle budding motif. Other nucleic acid molecules that will be generated encode an envelope and a target specific guide RNA. The nucleic acid molecules encoding the chimeric protein, the envelope, and the target specific guide RNA are introduced into HEK293T producer cells to generate the lipid delivery particles comprising the payload protein. DNA will be transiently expressed via non-viral methods. The non-viral methods will include chemical-based methods including DEAE-dextran, cationic lipids, or cationic polymers and physical methods including microinjection or electroporation. In this example, cells will be transfected using electroporation.

[0415] The cells will incorporate into the transfection media for approximately 18 hours. Following culturing in the media, the media will be changed out for a new solution or media. This second solution helps promote additional transient production of lipid delivery particles.

[0416] After approximately 30 hours have passed from incorporation in the second media, the lipid delivery particles will be harvested and the harvest will be filtered and collected. Harvesting methods include precipitation, depth filtration, microfiltration, or centrifugation. For this example, cells are harvested using depth filtration. Cell culture supernatant will be passed through a porous material with decreasing pore size to filter. The porous material will be composed of a layer of cellulose, a filter of diatomaceous earth, and a polymeric resin. Depth filtration systems will also be used including Millistak+ Pod from Millipore Sigma, Stax from Pall Corporation, 3M Zeta Plus from Cuno Inc., and Sartoclear P from Sartorius Stedim Biotech.

[0417] The harvested cells will be run through a purification protocol that retains the structural integrity of collected lipid delivery particles. Purification methods include FACS, BACS, MACS, size exclusion chromatography, ultrafiltration methods (*e.g.*, tangential flow filtration), clarifying normal flow filtration, sterilizing membrane filtration, or anion exchange chromatography.

Example 3: Administration of Lipid Delivery Particles to Patient with a Genetic Disease

[0418] Lipid delivery particles harvested from cell culture will be administered systemically to a subject with a genetic disease or disorder.

[0419] The lipid delivery particles with chimeric protein with a particle budding motif; a membrane penetrating peptide; and/or a post-translational modification motif will be provided to a subject with familial hypercholesterolemia (FH) caused by a pathogenic mutation in *PCSK9*. The lipid delivery particle generated and optimized will be provided to the subject intravenously. The gene editing payload protein will thus be delivered to the target cells by the lipid delivery particles, where the guide RNA delivers the gene editing machinery to *PCSK9*, and rectifies the pathogenic mutation. A sample from the subject will be taken after treatment to assess whether the genomic alteration is sufficient to restore normal function of the gene or to block pathogenic function of the pathogenic gene. A control sample from the subject will be taken from the subject prior to treatment.

[0420] The treated and control samples will be harvested for genomic DNA using Qiagen Blood & Cell Culture DNA kit. The DNA will be used as a template for PCR, wherein *HEXA* target specific primers flank the insertion locus. PCR will be performed using a high fidelity, proof-reading Taq polymerase (Prime Star, Takara) to isolate the target region of interest. The PCR inserts will be purified and prepared for a next generation sequencing run using a library preparation kit (Nextera XT, Illumina).

Example 4: Analysis of Nuclear Localization Signal (NLS) and Nuclear Export Signal (NES)

[0421] Experiments were conducted to examine the effects of NLS and NES elements on PH-payload. An example of a construct encoding a chimeric protein comprising a payload, pleckstrin homology domain (PH), and NLS element is shown in **FIG. 6A**. In this Example, the payload was adenine base editor, ADE8e fused to the mutant pleckstrin homology (PH) domain of Akt1 (also named as AKT) with E17K substitution (AKT PH (E17K)). There can be zero, or one or more codons, preferably neutral structural codons such as alanine, between the sequences provided for the constructs depicted in the figure.

[0422] Percentage of adenine to guanine base editing in 293T cells was measured in the treated cells. NES and NLS, as well as 1-3x NLS were analyzed. NLS derived from Influenza NP (e.g., containing SEQ ID NO: 268), Nucleoplasmin, human c-myc, and human FGF1 (x2) showed greater base editing percentage compared to original particles without an NLS or NES element (**FIG. 6B**).

Example 5: Analysis of Budding Domains on Lipid Delivery Particles and Editing Efficiency

[0423] Experiments were conducted to determine the effects of addition of budding domains with PH-payload chimeric proteins on lipid particle delivery. Without wishing to be bound by a particular theory, ectosome budding can be actively mediated through multiple mechanisms that include the ESCRT pathway. Ectosome budding from a producer cell might be increased by adding small peptide motifs that recruit ESCRT machinery and accessories to payload fusions. Budding domains that interact with human ESCRT pathway include PPXY, PXAP, and LYPXL. Budding domains were fused to the C-terminus of payload. Approximately 20 budding domain sequences (BD) were analyzed. Tandem repeats of the top performing budding domain was further compared against particles without budding domains. Experimental procedures were conducted in a similar manner across all treatment groups, *e.g.*, starting with similar amount of production cells, collecting and processing the production cells in a similar manner to produce the preparation containing the respective delivery particles.

[0424] Experiments were conducted with high dose and low dose of delivery particles. The high dose was 30 μ l of 100x polyethylene glycol (PEG)-concentrated particles and the low dose was 1 μ l of 50x polyethylene glycol (PEG)-concentrated particles. For the high dose budding domain experiment in U2OS cells, PPPY + PSAP budding domain showed the greatest editing percentage (93.63%), compared to that of control and domains derived from MLV gag, influenza, and EBOV VP40 (**FIG. 7A**). For the high dose budding domain experiment in U2OS cells, PPPY + PSAP budding domain also showed the greatest editing percentage (70.80%), compared to that from control and domains derived from HERV V gag, sindbis polypro., MLV gag, and EBOV VP40 (**FIG. 7B**).

[0425] A further experiment was conducted to examine different chimeric proteins, both with and without budding domains, on editing percentage. Experimental procedures were conducted in a similar manner across all treatment groups, *e.g.*, starting with similar amount of production cells, collecting and processing the production cells in a similar manner to produce the preparation containing the respective delivery particles. Example chimeric protein constructs are shown in **FIG. 8A**. Base constructs that were tested included payload ABE8e fused to PH domains, either wild-type or mutant E17K. Payload-PH constructs including budding domains with 1x, 3x, or 4x repeats of the budding domain. The six conditions that were tested include: payload alone (ABE8e), AKT WT (payload fused to wild-type Akt1 PH domain), AKT E17K (payload fused to mutant E17K Akt1 PH domain), AKT E17K 1x BD (payload fused to mutant E17K Akt1 PH domain with C-terminus budding domain), AKT E17K 3x BD (payload fused to mutant E17K Akt1 PH domain with 3 repeats of C-terminus budding domain), and AKT E17K 4x BD (payload

fused to mutant E17K Akt1 PH domain with 4 repeats of C-terminus budding domain). Constructs were tested at doses of 0.1 μ l, 1.0 μ l, and 10.0 μ l. At 1.0 μ l, the AKT E17K 3x BD showed the greatest editing percentage compared to that of all other conditions, with significantly greater editing percentage compared to constructs with no budding domain (**FIG. 8B**). In **FIG. 8B**, the bars from left-to-right for each dose designate: ABE8e (No PH), AKT WT, AKT E17K, AKT E17K 1x BD, AKT E17K 3x BD, and AKT E17K 4x BD. Adding a 30-amino acid 3x budding domain increased mean base editing from 51% to 67%. The average maximum editing at 10.0 μ l was also increased from 86% to 96%. This experiment demonstrated that addition of a budding domain element significantly enhanced the efficacy of lipid delivery particles.

Example 6: Analysis of Multimerization Domains on Lipid Delivery Particles and Editing Efficiency

[0426] Experiments were conducted to determine the effects of addition of multimerization domains with PH-payload constructs on lipid particle delivery. Viruses can multimerize their structural components via nucleocapsid and capsid lateral interactions. Without wishing to be bound by a particular theory, multimerization at the plasma membrane can promote assembly and budding of particles. Multimerization of editor payload can lead to higher local concentration of editor at a target site (e.g., nucleus) and higher on-target gene editing. Experimental procedures were conducted in a similar manner across all treatment groups, *e.g.*, starting with similar amount of production cells, collecting and processing the production cells in a similar manner to produce the preparation containing the respective delivery particles.

[0427] Multimerization domain (MD), GCN4 leucine zipper, may dimerize or tetramerize based on mutations installed in its heptad repeats. Two payload-MD fusions were assessed against three particle constructs without a multimerization domain or fusion. **FIG. 9** shows an example of dimerized construct with MD domains at the C-terminus. Similar to the experimental design of Example 5, an experiment was conducted to examine different constructs, both with and without multimerization domains, on editing percentage. Example constructs are shown in **FIG. 10A**. Base constructs included payload ABE8e fused to Akt1 PH domain, either wild-type or mutant E17K. The five conditions that were tested included: payload alone (ABE8e with no PH domain), AKT WT (payload fused to wild-type Akt1 PH domain), AKT E17K (payload fused to mutant E17K PH domain), MD dimer (payload fused to mutant E17K Akt1 PH domain with C-terminus dimerized multimerization domain), and MD tetramer (payload fused to mutant E17K Akt1 PH domain with C-terminus tetramer multimerization domain). Constructs were tested at doses of 0.1 μ l, 1.0 μ l, and 10.0 μ l. At 1.0 μ l, the MD dimer condition showed the greatest editing percentage compared to that of all other conditions, with significantly greater editing percentage compared to

constructs with no multimerization domain (**FIG. 10B**). In **FIG. 10B**, the bars from left-to-right for each dose designate: ABE8e (No PH), AKT WT, AKT E17K, MD Dimer, and MD Tetramer. The MD dimer addition to the particle construct led to significantly greater editing percentage (e.g., adenine to guanine conversion) compared to the editing percentage from AKT E17K construct ($p < 0.0024$). At 1 μl , adding a 28 amino acid dimerization domain increased mean base editing from 51% to 71%. The average maximum base editing at 10.0 μl also increased from 86% to 95%. This experiment demonstrated that addition of a multimerization domain element significantly enhanced the efficacy of lipid delivery particles.

Example 7: Analysis of Post-Translation Modifications on Lipid Delivery Particles and Editing Efficiency

[0428] Experiments were conducted to determine the effects of addition of post-translational modification motifs (PTMs) with PH-payload constructs on lipid particle delivery. All experimental procedures were conducted in a similar manner across all treatment groups, e.g., starting with similar amount of production cells, collecting and processing the production cells in a similar manner to produce the preparation containing the respective delivery particles. The PTM was fused to the N-terminus of the payload. An example construct included PTM-payload (ABE8e)-PH-NLS (**FIG. 11A**). An experiment was conducted to compare the effect of fusing PTM, budding domain (BD), or multimerization domain (MD) to the delivery construct, compared to a construct with payload-PH (original). Tested doses were 0.2 μl , 2.0 μl , and 20.0 μl . At 2.0 μl , all construct modifications (e.g., PTM, BD, MD) showed improved editing percentage as compared to that from original (e.g., control payload-PH particle) (**FIG. 11B**). In **FIG. 11B**, the set of three bars from left-to-right for each group of doses designate: PTM, BD, MD, and Original.

[0429] While preferred embodiments of the present disclosure have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. It should be understood that various alternatives to the embodiments of the present disclosure may be employed in practicing the present disclosure. It is intended that the following claims define the scope of the present disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

What is claimed is:

1. A lipid delivery particle that comprises:
 - (a) a lipid membrane; and
 - (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload;
wherein the chimeric protein further comprises:
 - (i) a particle budding motif;
 - (ii) a membrane penetrating peptide;
 - (iii) a post-translational modification motif that promotes assembly of the lipid delivery particle and/or promotes loading of the payload into the lipid delivery particle;
 - (iv) a multimerization motif, or
 - (v) any combination thereof.
2. A lipid delivery particle that comprises:
 - (a) a lipid membrane; and
 - (b) a chimeric protein, wherein the chimeric protein comprises a plasma membrane recruitment element and a payload;
wherein the chimeric protein further comprises:
 - (i) a particle budding motif;
 - (ii) a membrane penetrating peptide;
 - (iii) a post-translational modification motif selected from the group consisting of: myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, and farnesylation motif;
 - (iv) a multimerization motif, or
 - (v) any combination thereof.
3. The lipid delivery particle of claim 1 or 2, wherein the lipid delivery particle is an ectosome, an exosome, or a viral-like particle.
4. The lipid delivery particle of any one of claims 1-3, wherein the plasma membrane recruitment element comprises a pleckstrin homology (PH) domain.
5. The lipid delivery particle of any one of claims 1-4, wherein the plasma membrane recruitment element comprises a mutated PH domain or an engineered PH domain.
6. The lipid delivery particle on any one of claims 4 or 5, wherein the PH domain is from a protein selected from the group consisting of human phospholipase C δ 1, human Akt1, human Akt1 with E17K substitution, human 3-phosphoinositide-dependent protein

kinase 1, human Daap1, mouse Grp1, human Grp1, human OSBP, human Btk1, human FAPP1, human CERT, human PKD, human PHLPP1, human SWAP70, and human MAPKAP1.

7. The lipid delivery particle of any one of claims 4-6, wherein the plasma membrane recruitment element further comprises a retroviral gag protein.
8. The lipid delivery particle of claim 7, wherein the PH domain is linked to an N-terminus of the retroviral gag protein.
9. The lipid delivery particle of any one of claims 1-8, wherein the PH domain is linked to a gag-payload fusion peptide.
10. The lipid delivery particle of any one of claims 1-9, wherein the plasma membrane recruitment element is linked via its C-terminus to an N-terminus of the payload.
11. The lipid delivery particle of any one of claims 1-9, wherein the plasma membrane recruitment element is linked via its N-terminus to a C-terminus of the payload.
12. The lipid delivery particle of any one of claims 1-11, wherein the payload comprises a therapeutic agent selected from the group consisting of an antibody, a cytokine, a therapeutic protein, a small molecule therapeutic, a therapeutic DNA, and a therapeutic RNA.
13. The lipid delivery particle of any one of claims 1-12, wherein the payload comprises a ribonucleoprotein, a zinc-finger, a transcription activator-like effector (TALE), a nucleic acid modifying domain, or a guidable polypeptide domain.
14. The lipid delivery particle of any one of claims 1-13, wherein the chimeric protein comprises the particle budding motif.
15. The lipid delivery particle of any one of claims 1-14, wherein the particle budding motif is PTAP, PPXY, LPYX, PPEE, LYPL, LYPSL, PPPY, PPEY, PSAP, LYPAL, PPAP, PPPE, YMYL, YRKL, YQCL, YCYL, LYRTL, YPX_nL, or any combination thereof, and wherein n is an integer from 1 to 5.
16. The lipid delivery particle of claim 14 or 15, wherein the particle budding motif is linked to the payload or the plasma membrane recruitment element.
17. The lipid delivery particle of claim 14 or 15, wherein the particle budding motif is linked between the payload and the plasma membrane recruitment element.
18. The lipid delivery particle of claim 14 or 15, wherein the particle budding motif is linked to the chimeric protein based on a construct arrangement listed in Table 12 or Table 15.
19. The lipid delivery particle of any one of claims 1-18, wherein the chimeric protein comprises the membrane penetrating peptide.

20. The lipid delivery particle of claim 19, wherein the membrane penetrating peptide is linked to the payload or the plasma membrane recruitment element.
21. The lipid delivery particle of claim 19, wherein the membrane penetrating peptide is linked between the payload and the plasma membrane recruitment element.
22. The lipid delivery particle of claim 19, wherein the membrane penetrating peptide is linked to the chimeric protein based on a construct arrangement listed in Table 13.
23. The lipid delivery particle of any one of claims 1-22, wherein the chimeric protein comprises the post-translational modification motif.
24. The lipid delivery particle of claim 23, wherein the chimeric protein comprises the post-translational modification motif linked to one or more amino acids of the payload or one or more amino acids of the plasma membrane recruitment element.
25. The lipid delivery particle of claim 23, wherein the chimeric protein comprises the post-translational modification motif linked to one or more amino acids of a region in the chimeric protein outside of the payload and the plasma membrane recruitment element.
26. The lipid delivery particle of claim 23, wherein the chimeric protein comprises the post-translational modification motif linked to one or more amino acids of a region in the chimeric protein between the payload and the plasma membrane recruitment element.
27. The lipid delivery particle of claim 23 wherein the post translational modification motif is linked to the chimeric protein based on a construct arrangement listed in Table 14 or Table 15.
28. The lipid delivery particle of any one of claims 1-27, wherein the chimeric protein comprises the multimerization motif.
29. The lipid delivery particle of any one of claims 1-28, wherein the multimerization motif comprises a leucine zipper oligomerization motif.
30. The lipid delivery particle of claim 29, wherein the multimerization motif is linked to the payload or the plasma membrane recruitment element.
31. The lipid delivery particle of claim 29, wherein the multimerization motif is linked between the payload and the plasma membrane recruitment element.
32. The lipid delivery particle of claim 29, wherein the multimerization motif is linked to the chimeric protein based on a construct arrangement listed in Table 16.
33. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif and the membrane penetrating peptide.
34. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif and the post-translational modification motif.

35. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif and the multimerization motif.
36. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the membrane penetrating peptide and the post-translational modification motif.
37. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the membrane penetrating peptide and the multimerization motif.
38. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the post-translational modification motif and the multimerization motif.
39. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the post-translational modification motif.
40. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the multimerization motif.
41. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif, the post-translational modification motif, and the multimerization motif.
42. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the post-translational modification motif, the membrane penetrating peptide, and the multimerization motif.
43. The lipid delivery particle of any one of claims 1-32, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, and the multimerization motif.
44. The lipid delivery particle of any one of claims 1-43, wherein the chimeric protein comprises a tandem repeat of the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, or any combination thereof.
45. The lipid delivery particle of claim 44, wherein the tandem repeat of the particle budding motif further comprises a linker sequence.
46. The lipid delivery particle of any one of claims 1-44, wherein the chimeric protein further comprises a nuclear localization signal (NLS) or a nuclear export signal (NES).
47. The lipid delivery particle of claim 46, wherein the nuclear localization signal is linked to the payload or the plasma membrane recruitment element.

48. The lipid delivery particle of claim 46 or 47, wherein the nuclear localization signal comprises a sequence having at least about 80% sequence identity to the sequence set forth in any one of SEQ ID NOs: 268-280.
49. The lipid delivery particle of any one of claims 46-48, wherein the nuclear localization signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 268-280.
50. The lipid delivery particle of claim 48 or 49, wherein the nuclear localization signal comprises a sequence as set forth in any one of SEQ ID NOs: 268-280.
51. The lipid delivery particle of any one of claims 46-50, wherein the lipid delivery particle comprises 2, 3, 4, 5, or more NLS sequences.
52. The lipid delivery particle of claim 46, wherein the nuclear export signal is linked to the payload or the plasma membrane recruitment element.
53. The lipid delivery particle of claim 46, wherein the nuclear export signal is linked between the payload and the plasma membrane recruitment element.
54. The lipid delivery particle of claim 46, wherein the nuclear export signal is linked to the particle budding motif, the membrane penetrating particle, and/or the post-translational modification motif.
55. The lipid delivery particle of claim 46, wherein the nuclear export signal is linked to a cleavage site.
56. The lipid delivery particle of any one of claims 46 and 52-55, wherein the nuclear export signal comprises a sequence having at least about 80% sequence identity to the sequence set forth in any one of SEQ ID NOs: 87-267.
57. The lipid delivery particle of any one of claims 46 and 52-55, wherein the nuclear export signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 87-267.
58. The lipid delivery particle of claim 56 or 57, wherein the nuclear export signal comprises a sequence as set forth in any one of SEQ ID NOs: 87-267.
59. The lipid delivery particle of claim 46, 52-58, wherein the lipid delivery particle comprises 2, 3, 4, 5, or more NES sequences.
60. The lipid delivery particle of any one of claims 1-59, wherein the chimeric protein comprises a linker sequence.
61. The lipid delivery particle of claim 60, wherein the linker sequence comprises a cleavable linker.

62. The lipid delivery particle of any one of claims 1 and 3-61, wherein the post-translational modification motif comprises myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, farnesylation motif, or any combination thereof.
63. The lipid delivery particle of any one of claims 1-62, wherein the particle budding motif comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 1-32.
64. The lipid delivery particle of any one of claims 1-63, wherein the particle budding motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 1-32.
65. The lipid delivery particle of claim 63 or 64, wherein the particle budding motif comprises a sequence as set forth in any one of SEQ ID NOs: 1-32.
66. The lipid delivery particle of any one of claims 1-65, wherein the membrane penetrating peptide comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 33-56.
67. The lipid delivery particle of any one of claims 1-66, wherein the membrane penetrating peptide comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 33-56.
68. The lipid delivery particle of claim 66 or 67, wherein the membrane penetrating peptide comprises a sequence as set forth in any one of SEQ ID NOs: 33-56.
69. The lipid delivery particle of any one of claims 1-68, wherein the post-translational modification motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 57-86.
70. The lipid delivery particle of any one of claims 1-69, wherein the post-translational modification motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 57-86.
71. The lipid delivery particle of claim 69 or 70, wherein the post-translational modification motif comprises a sequence as set forth in SEQ ID NOs: 57-86.
72. The lipid delivery particle of any one of claims 1-71, wherein the multimerization motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 281-285.
73. The lipid delivery particle of any one of claims 1-72, wherein the multimerization motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 281-285.
74. The lipid delivery particle of claim 71 or 72, wherein the multimerization motif comprises a sequence as set forth in SEQ ID NOs: 281-285.

75. The lipid delivery particle of any one of claims 1-74, wherein the particle budding motif interacts with an endosomal sorting complex required for transport (ESCRT).
76. The lipid delivery particle of any one of claims 1-75, wherein the lipid delivery particle further comprises an envelope protein.
77. The lipid delivery particle of claim 76, wherein the envelope protein comprises a VSV glycoprotein (VSV-G), a human immunodeficiency virus (HIV) GP160 glycoprotein, a Baboon Endogenous Retrovirus (BaEVTR) glycoprotein, a fusion protein of Vesicular stomatitis Indiana virus and Rabies virus glycoprotein (FuG-E), an ecotropic Murine Leukemia Virus envelope protein (MLV ENV ecotropic), a human T-cell lymphotropic virus type 1 (HTLV-1) glycoprotein, an amphotropic Murine Leukemia Virus envelope protein (MLV ENV amphotropic), a moloney murine leukemia virus 10A1 strain glycoprotein (MLV 10A1), a Baculovirus envelope glycoprotein (GP64), a pantropic MLV envelope protein, a xenotropic MLV envelope protein, a xenotropic murine leukemia virus (XMLV) envelope protein, a Moloney murine leukemia virus (MMLV) envelope protein, a moloney murine sarcoma virus (MoMSVg) envelope protein, a simian endogenous type D retrovirus protein (RD-114), a gibbon ape leukemia virus (GALV) envelope protein, a feline leukemia virus (FLV) envelope protein, a mouse mammary tumor virus (MMTV) envelope protein, an avian leukosis virus envelope protein, a Rous Sarcoma virus envelope protein, or an endogenous feline virus envelope protein (RD114 ENV), or a mutant thereof.
78. The lipid delivery particle of claim 76, wherein the envelope protein comprises a HIV-1 envelope protein, a respiratory syncytial virus glycoprotein, a SARS-CoV glycoprotein, a rabies glycoprotein, a Mokola virus glycoprotein, a Semliki Forest virus glycoprotein, a Venezuelan equine encephalitis virus glycoprotein, an Ebola virus glycoprotein, a Marburg virus glycoprotein, or a mutant thereof.
79. The lipid delivery particle of claim 76, wherein the envelope protein comprises a glycoprotein of a mammalian endogenous retrovirus or mutant thereof.
80. The lipid delivery particle of claim 79, wherein the glycoprotein of the mammalian endogenous retrovirus is a glycoprotein of a human endogenous retrovirus (hERV).
81. The lipid delivery particle of claim 80, wherein the glycoprotein of a hERV comprises a hENVH1, a hENVH2, a hENVH3, a hENVK1, a hENVK2, a hENVK3, a hENVK4, a hENVK5, a hENVK6, a hENVT, a hENVW, a hENVFRD, a hENVR, a hENVR(b), a hENVR(c)1, a hENVR(c)2 or a hENVK_{con}, or a biologically active mutant thereof.
82. The lipid delivery particle of any one of claims 76-81, wherein the envelope protein is fused to a targeting moiety, optionally an scFv.

83. The lipid delivery particle of claim 76, wherein the envelope protein comprises a non-viral envelope protein.
84. A chimeric protein comprising:
- (a) a plasma membrane recruitment element;
 - (b) a payload; and
 - (c) one, two, three, or all of:
 - (i) a particle budding motif;
 - (ii) a membrane penetrating peptide;
 - (iii) a post-translational modification motif that promotes assembly of a lipid delivery particle and/or promotes loading of the payload into the lipid delivery particle; or
 - (iv) a multimerization motif.
85. A chimeric protein comprising:
- (a) a plasma membrane recruitment element;
 - (b) a payload; and
 - (c) one, two, three, or all of:
 - (i) a particle budding motif;
 - (ii) a membrane penetrating peptide;
 - (iii) a post-translational modification motif selected from the group consisting of: myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, and farnesylation motif; or
 - (iv) a multimerization motif.
86. The chimeric protein of claim 84 or 85, wherein the plasma membrane recruitment element comprises a pleckstrin homology (PH) domain.
87. The chimeric protein of any one of claims 84-86, wherein the plasma membrane recruitment element comprises a mutated PH domain or an engineered PH domain.
88. The chimeric protein on any one of claims 84-87, wherein the PH domain is from a protein selected from a group consisting of human phospholipase C δ 1, human Akt1, human Akt1 with E17K substitution, human 3-phosphoinositide-dependent protein kinase 1, human Daap1, mouse Grp1, human Grp1, human OSBP, human Btk1, human FAPP1, human CERT, human PKD, human PHLPP1, human SWAP70, and human MAPKAP1.
89. The chimeric protein of any one of claims 86-88, wherein the plasma membrane recruitment element further comprises a retroviral gag protein.
90. The chimeric protein of claim 89, wherein the PH domain is linked to an N-terminus of the retroviral gag protein.

91. The chimeric protein of any one of claims 84-90, wherein the PH domain is linked to a gag-payload fusion construct.
92. The chimeric protein of any one of claims 84-91, wherein the plasma membrane recruitment element is linked via its C-terminus to an N-terminus of the payload.
93. The chimeric protein of any one of claims 84-92, wherein the plasma membrane recruitment element is linked via its N-terminus to a C-terminus of the payload.
94. The chimeric protein of any one of claims 84-93, wherein the payload comprises a therapeutic agent selected from the group consisting of an antibody, a cytokine, a therapeutic protein, a small molecule therapeutic, a therapeutic DNA, and a therapeutic RNA.
95. The chimeric protein of any one of claims 84-94, wherein the payload comprises a ribonucleoprotein, a zinc-finger, a transcription activator-like effector (TALE), or a CRISPR-based genome editing or modulating protein.
96. The chimeric protein of any one of claims 84-95, wherein the chimeric protein comprises the particle budding motif.
97. The chimeric protein of any one of claims 84-96, wherein the particle budding motif is PTAP, PPXY, LPYX, PPEE, LYPL, LYPSL, PPPY, PPEY, PSAP, LYPAL, PPAP, PPPE, YMYL, YRKL, YQCL, YCYL, LYRTL, YPX_nL, or any combination thereof, and wherein n is an integer from 1 to 5.
98. The chimeric protein of claim 96 or 97, wherein the particle budding motif is linked to the payload or the plasma membrane recruitment element.
99. The chimeric protein of claim 96 or 97, wherein the particle budding motif is linked between the payload and the plasma membrane recruitment element.
100. The chimeric protein of claim 96 or 97, wherein the particle budding motif is linked to the chimeric protein based on a construct arrangement listed in Table 12 or Table 15.
101. The chimeric protein of any one of claims 84-100, wherein the chimeric protein comprises the membrane penetrating peptide.
102. The chimeric protein of claim 101, wherein the membrane penetrating peptide is linked to the payload or the plasma membrane recruitment element.
103. The chimeric protein of claim 101, wherein the membrane penetrating peptide is linked between the payload and the plasma membrane recruitment element.
104. The chimeric protein of claim 101, wherein the membrane penetrating peptide is linked to the chimeric protein based on a construct arrangement listed in Table 13.
105. The chimeric protein of any one of claims 84-104, wherein the chimeric protein comprises the post-translational modification motif.

106. The chimeric protein of claim 105, wherein the chimeric protein comprises the post-translational modification motif to one or more amino acids of the payload or one or more amino acids of the plasma membrane recruitment element.
107. The chimeric protein of claim 105, wherein the chimeric protein comprises the post-translational modification motif to one or more amino acids of a region in the chimeric protein outside of the payload and the plasma membrane recruitment element.
108. The chimeric protein of claim 105, wherein the chimeric protein comprises the post-translational modification motif to one or more amino acids of a region in the chimeric protein between the payload and the plasma membrane recruitment element.
109. The chimeric protein of claim 105 wherein the post-translational modification motif is linked to the chimeric protein based on a construct arrangement listed in Table 14 or Table 15.
110. The chimeric protein of any one of claims 84-109, wherein the chimeric protein comprises the multimerization motif.
111. The chimeric protein of any one of claims 84-110, wherein the multimerization motif comprises a leucine zipper oligomerization motif.
112. The chimeric protein of claim 110 or 111, wherein the multimerization motif is linked to the payload or the plasma membrane recruitment element.
113. The chimeric protein of claim 110 or 111, wherein the multimerization motif is linked between the payload and the plasma membrane recruitment element.
114. The chimeric protein of claim 110 or 111, wherein the multimerization motif is linked to the chimeric protein based on a construct arrangement listed in Table 16.
115. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif and the membrane penetrating peptide.
116. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif and the post-translational modification motif.
117. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif and the multimerization motif.
118. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the membrane penetrating peptide and the post-translational modification motif.
119. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the membrane penetrating peptide and the multimerization motif.
120. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the post-translational modification motif and the multimerization motif.

121. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the post-translational modification motif.
122. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, and the multimerization motif.
123. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif, the post-translational modification motif, and the multimerization motif.
124. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the post-translational modification motif, the membrane penetrating peptide, and the multimerization motif.
125. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, and the multimerization motif.
126. The chimeric protein of any one of claims 84-114, wherein the chimeric protein comprises a tandem repeat of the particle budding motif, the membrane penetrating peptide, the post-translational modification motif, or any combination thereof.
127. The chimeric protein of claim 126, wherein the tandem repeat of the particle budding motif further comprises a linker sequence.
128. The chimeric protein of any one of claims 84-127, wherein the chimeric protein further comprises a nuclear localization signal (NLS) or a nuclear export signal (NES).
129. The chimeric protein of claim 128, wherein the nuclear localization signal is linked to the payload or the plasma membrane recruitment element.
130. The chimeric protein of claim 128 or 129, wherein the nuclear localization signal comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 268-280.
131. The chimeric protein of any one of claims 128-130, wherein the nuclear localization signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in SEQ ID NOs: 268-280.
132. The chimeric protein of claim 130 or 131, wherein the nuclear localization signal comprises a sequence as set forth in SEQ ID NOs: 268-280.
133. The lipid delivery particle of any one of claims 128-132, wherein the chimeric protein comprises 2, 3, 4, 5, or more NLS sequences.

134. The chimeric protein of claim 128, wherein the nuclear export signal is linked to the payload or the plasma membrane recruitment element.
135. The chimeric protein of claim 128, wherein the nuclear export signal is linked between the payload and the plasma membrane recruitment element.
136. The chimeric protein of claim 128, wherein the nuclear export signal is linked to the particle budding motif, the membrane penetrating particle, the post-translational modification motif, and/or the multimerization motif.
137. The chimeric protein of claim 128, wherein the nuclear export signal is linked to a cleavage site.
138. The chimeric protein of any one of claims 128 and 134-137, wherein the nuclear export signal comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 87-267.
139. The chimeric protein of claim 128 and 134-138, wherein the nuclear export signal comprises a sequence having at least one but not more than 8 modifications of an amino acid sequence as set forth in SEQ ID NOs: 87-267.
140. The chimeric protein of claim 138 or 139, wherein the nuclear export signal comprises a sequence as set forth in SEQ ID NOs: 87-267.
141. The chimeric protein of any one of claims 128, and 134-140, wherein the chimeric protein comprises 2, 3, 4, 5, or more NES sequences.
142. The chimeric protein of any one of claims 84-141, wherein the chimeric protein further comprises a linker sequence.
143. The chimeric protein of claim 142, wherein the linker sequence comprises a cleavable linker.
144. The chimeric protein of any one of claims 84-143, wherein the post-translational modification comprises myristoylation motif, acetylation motif, isoprenylation motif, palmitoylation motif, farnesylation motif, or any combination thereof.
145. The chimeric protein of any one of claims 84-144, wherein the particle budding motif comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 1-32.
146. The chimeric protein of any one of claims 84-145, wherein the particle budding motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 1-32.
147. The chimeric protein of claim 145 or 146, wherein the particle budding motif comprises a sequence as set forth in any one of SEQ ID NOs: 1-32.

148. The chimeric protein of any one of claims 84-147, wherein the membrane penetrating peptide comprises a sequence having at least about 80% sequence identity to the sequence as set forth in any one of SEQ ID NOs: 33-56.
149. The chimeric protein of any one of claims 84-148, wherein the membrane penetrating peptide comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 33-56.
150. The chimeric protein of claim 148 or 149, wherein the membrane penetrating peptide comprises a sequence as set forth in any one of SEQ ID NOs: 33-56.
151. The chimeric protein of any one of claims 84-150, wherein the post-translational modification motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 57-86.
152. The chimeric protein of any one of claims 84-151, wherein the post-translational modification motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 57-86.
153. The chimeric protein of claim 151 or 152, wherein the post-translational modification motif comprises a sequence as set forth in SEQ ID NOs: 57-86.
154. The chimeric protein of any one of claims 84-153, wherein the multimerization motif comprises a sequence having at least about 80% sequence identity to SEQ ID NOs: 281-285.
155. The chimeric protein of any one of claims 84-154, wherein the multimerization motif comprises a sequence having at least one but not more than 10 modifications of an amino acid sequence as set forth in any one of SEQ ID NOs: 281-285.
156. The chimeric protein of claim 154 or 155, wherein the multimerization motif comprises a sequence as set forth in SEQ ID NOs: 281-285.
157. The chimeric protein of any one of claims 84-156, wherein the particle budding motif interacts with an endosomal sorting complex required for transport (ESCRT).
158. A method for preparing a lipid delivery particle comprising a payload, the method comprising:
- (a) providing a producer cell that produces the lipid delivery particle of any one of claims 1-83; and
 - (b) collecting the lipid delivery particle from the producer cell.
159. The method of claim 158, wherein the producer cell comprises a nucleic acid sequence encoding the chimeric protein.
160. The method of claim 159, wherein the method further comprises expressing in the producer cell the nucleic acid molecule encoding the chimeric protein.

161. The method of any one of claims 158-160, wherein the producer cell comprises a nucleic acid sequence encoding the envelope protein.
162. The method of claim 161, wherein the method further comprises expressing in the producer cell the nucleic acid molecule encoding the envelope protein.
163. The method of any one of claims 158-162, wherein the method further comprises culturing the producer cell in a medium and maintaining the producer under conditions to produce the lipid delivery particle.
164. The method of any one of claims 158-163, wherein the method further comprises harvesting the medium and purifying the extracellular vesicle.
165. The method of claim 164, wherein the purifying retains the structural integrity of the extracellular vesicle.
166. A lipid delivery particle produced by the method of any one of claims 158-165.
167. A nucleic acid sequence encoding the chimeric protein of any one of claims 1-157.
168. A composition comprising:
- (a) a first nucleic acid sequence encoding an envelope protein; and
 - (b) a second nucleic acid sequence encoding a chimeric protein,
- wherein the chimeric protein comprises:
- (i) a plasma membrane recruitment element;
 - (ii) a payload; and
 - (iii) one, two, three, or all of:
 - 1. a particle budding motif;
 - 2. a membrane penetrating peptide;
 - 3. a post-translational modification motif; or
 - 4. a multimerization motif.
169. A vector comprising the nucleic acid sequence of claim 167 or 168.
170. The vector of claim 169, wherein the vector is a vector selected from a plasmid, a cosmid, a bacterial vector, a viral vector, or an artificial chromosome.
171. A cell comprising the lipid delivery particle of any one of claims 1-83.
172. A cell produced by the method of any one of claims 158-165.
173. A pharmaceutical formulation comprising the lipid delivery particle of any one of claims 1-83, and a pharmaceutically acceptable carrier, excipient, adjuvant or diluent.
174. A kit comprising:
- (a) the lipid delivery particle of any one of claims 1-83, or the pharmaceutical formulation of claim 173; and
 - (b) instructions for use of the lipid delivery particle or the pharmaceutical formulation.

175. The lipid delivery particle of any one of claims 1-83, wherein the chimeric protein enhances an efficacy of the lipid delivery particle compared to an efficacy of an otherwise identical lipid delivery particle without the chimeric protein.
176. The lipid delivery particle of claim 175, wherein the efficacy of the lipid delivery particle is a base editing percentage.



FIG. 1A



FIG. 1B



FIG. 1C



FIG. 1D



FIG. 1E

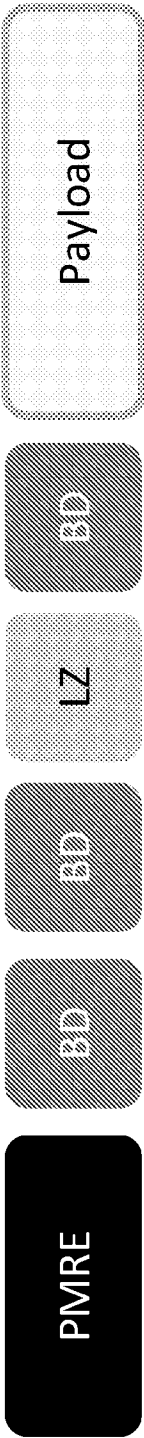


FIG. 1F

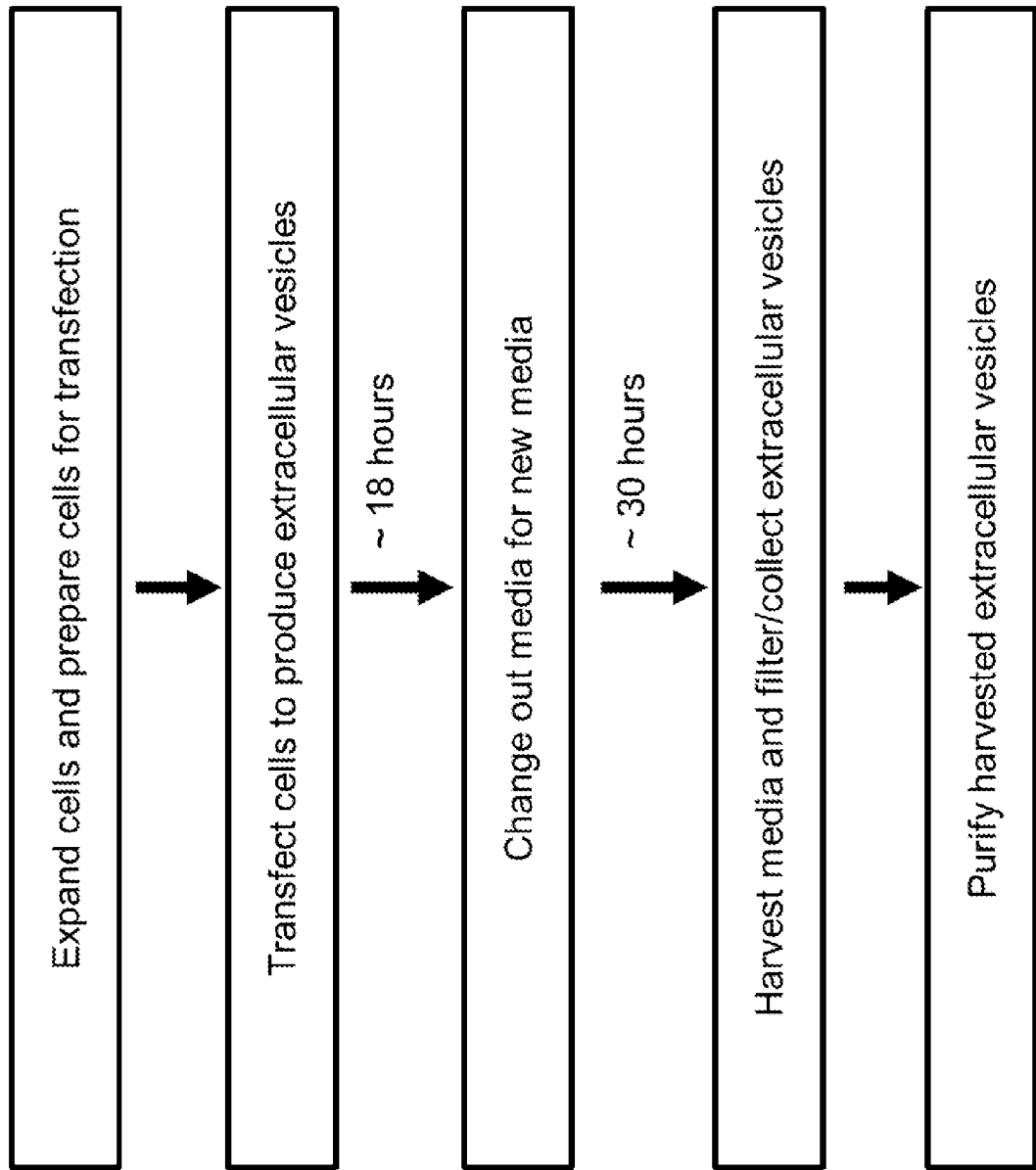


FIG. 2

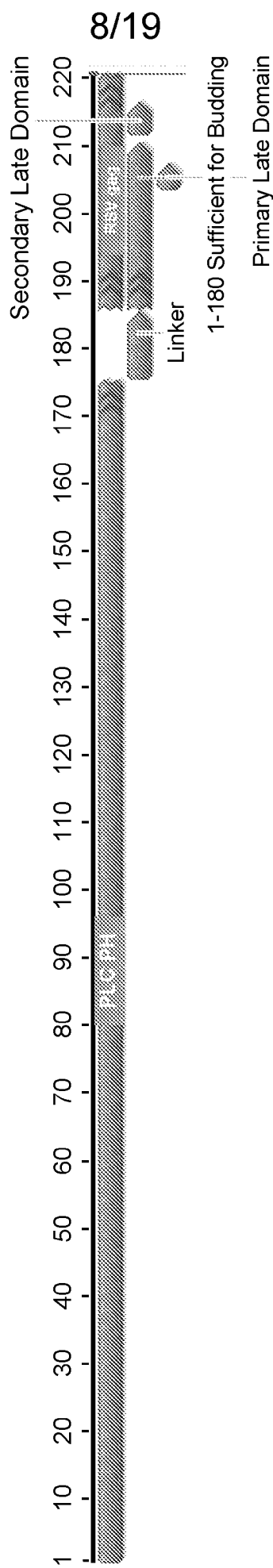


FIG. 3

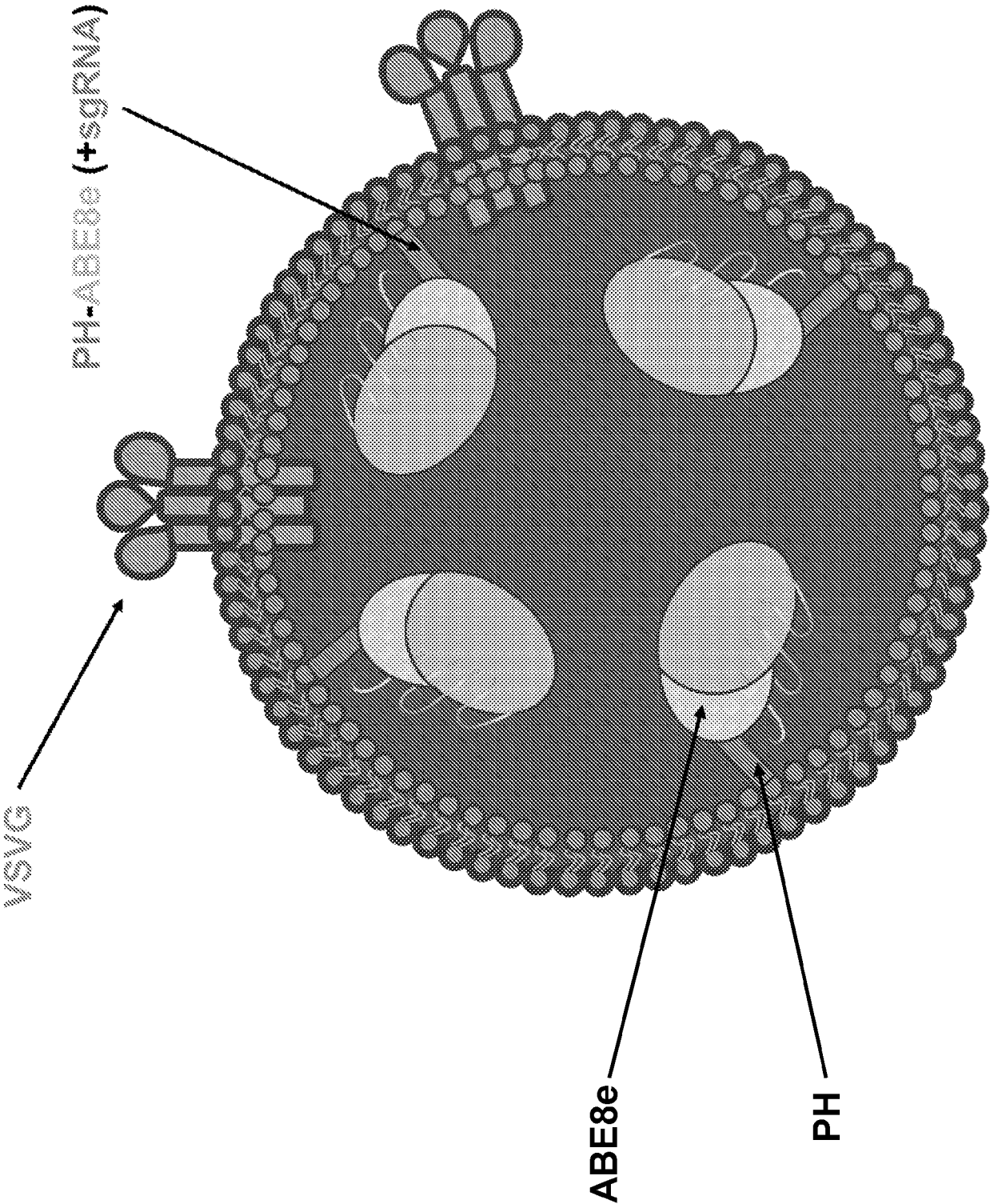


FIG. 4

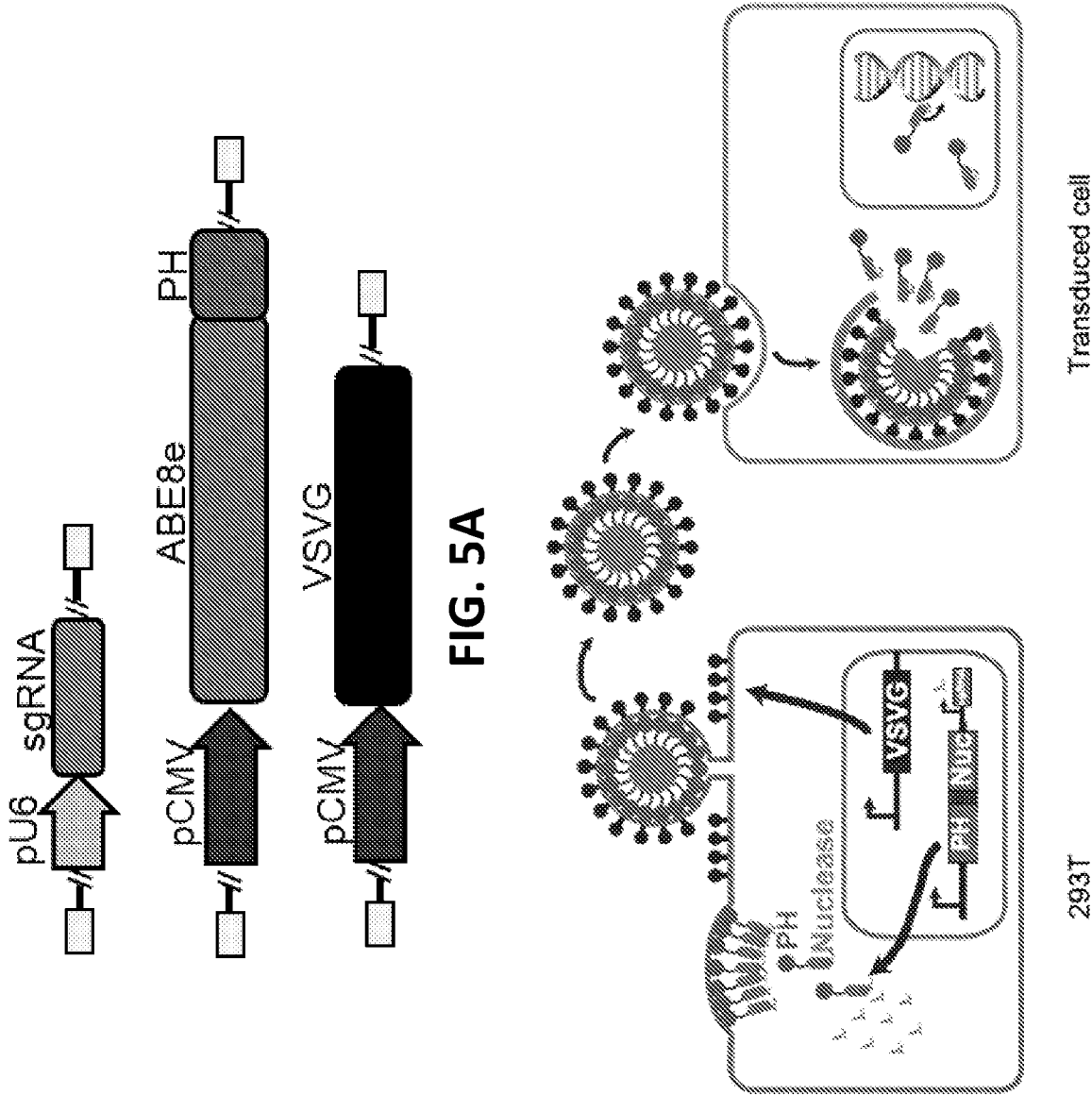


FIG. 5B

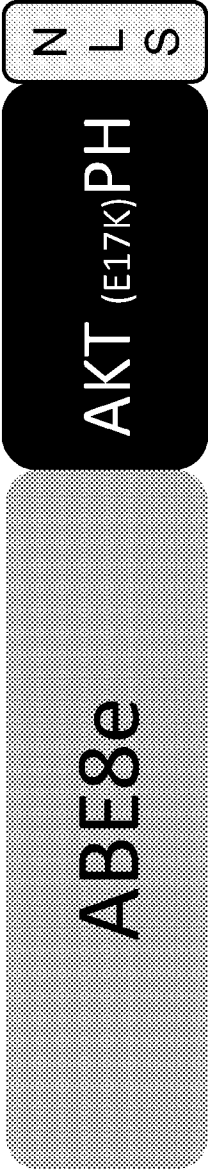


FIG. 6A

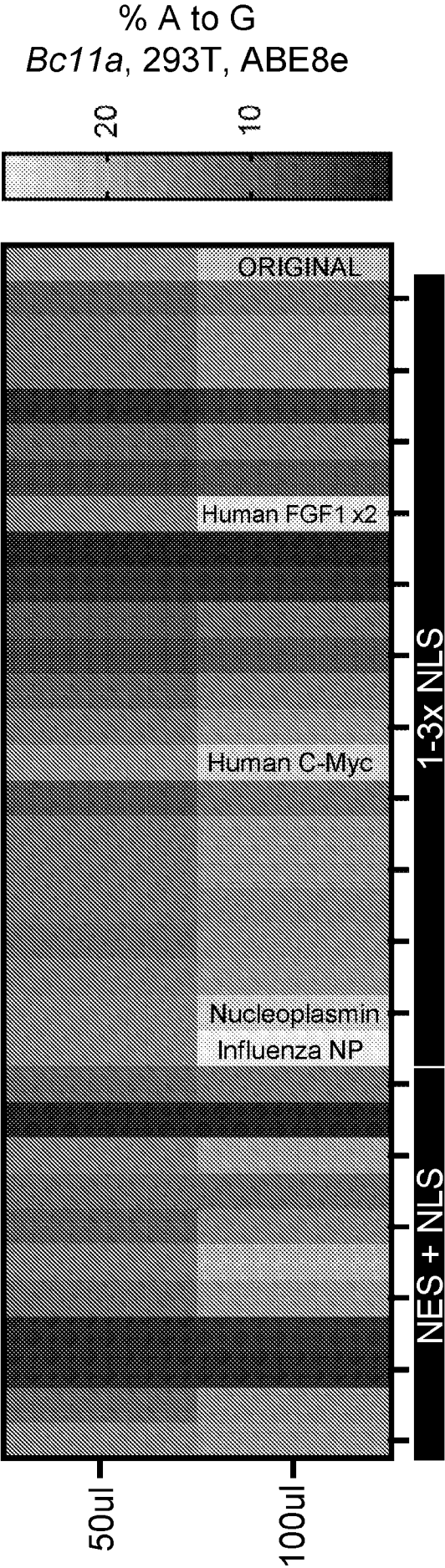


FIG. 6B

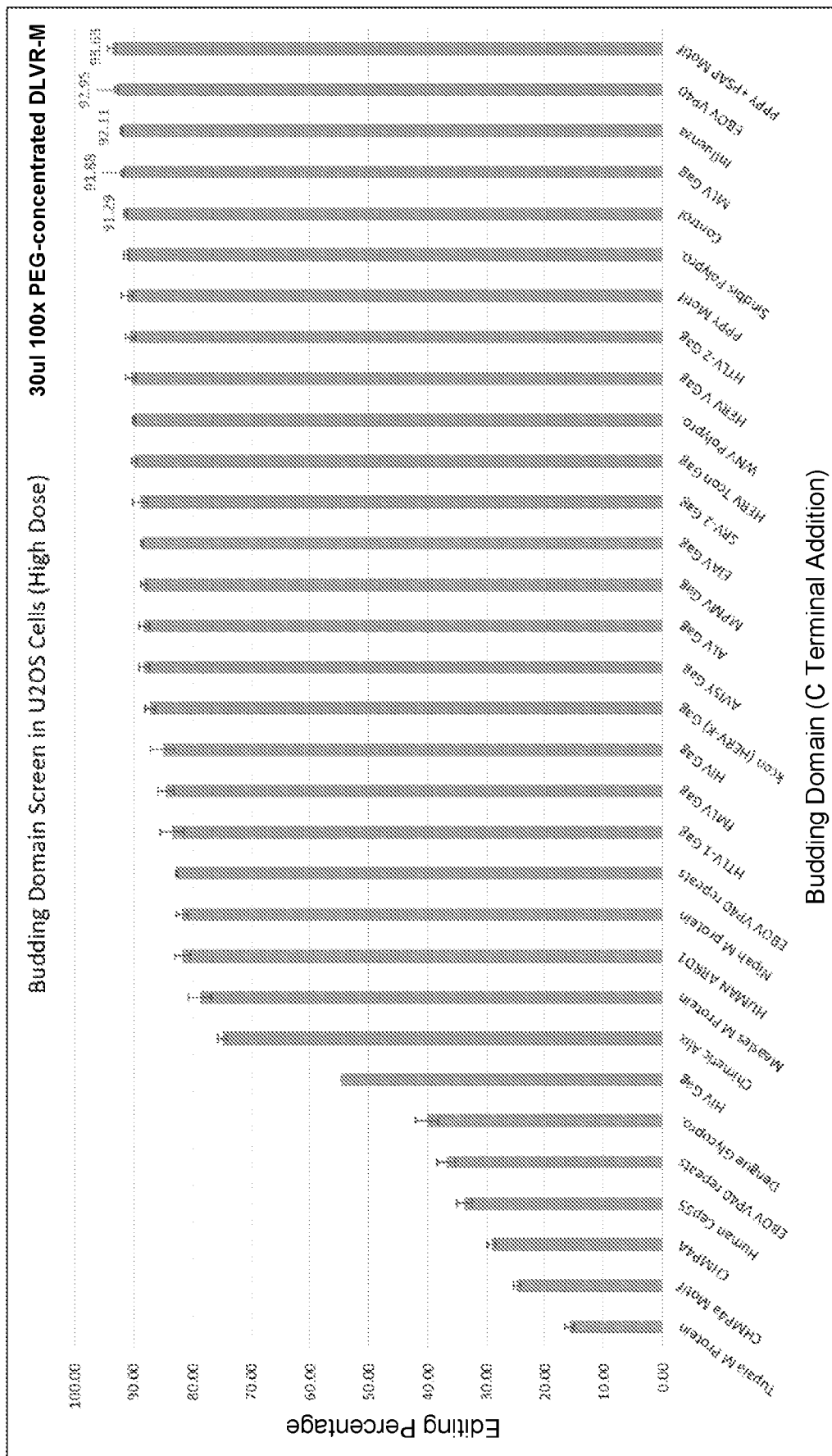


FIG. 7A

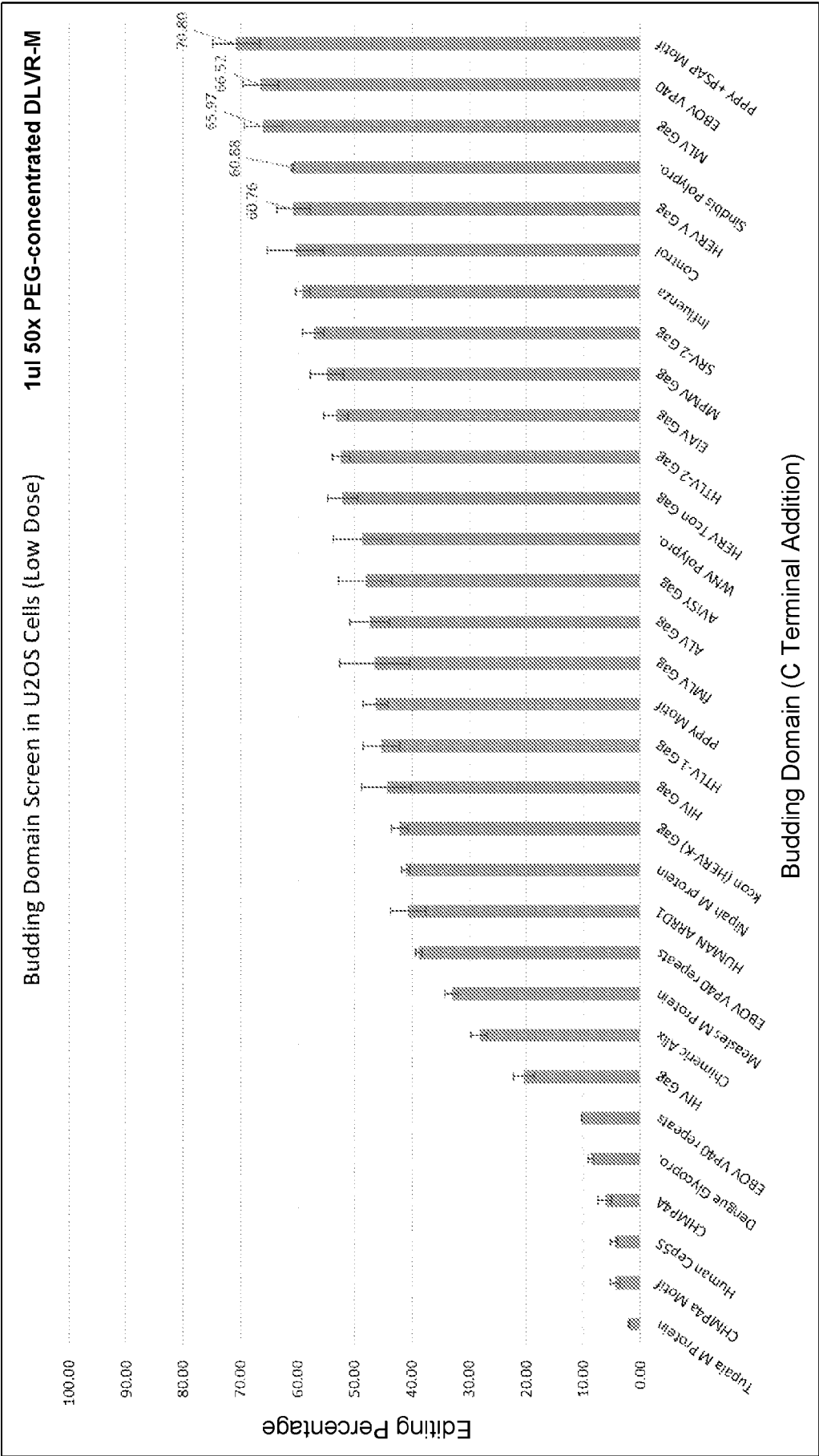


FIG. 7B

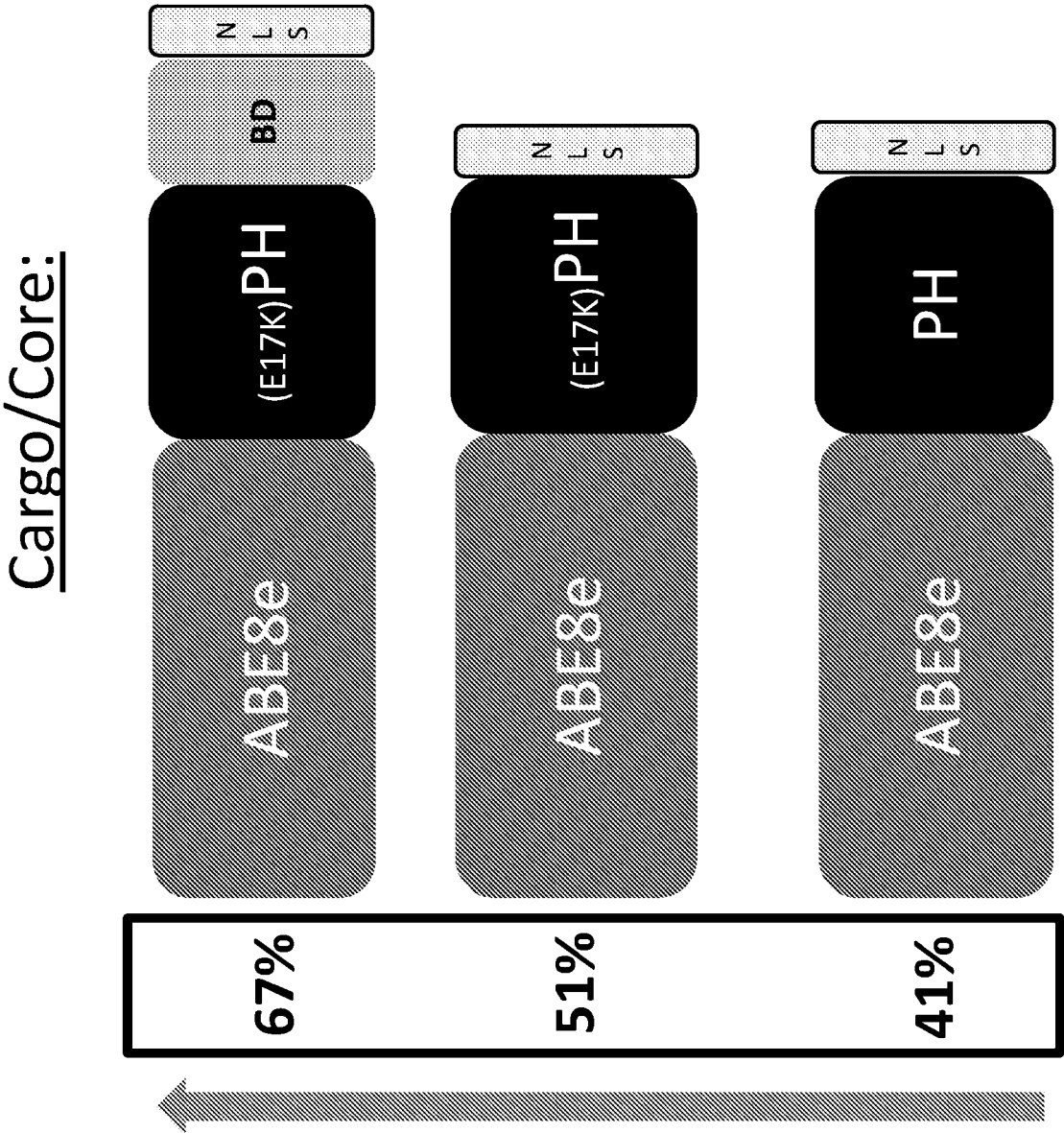


FIG. 8A

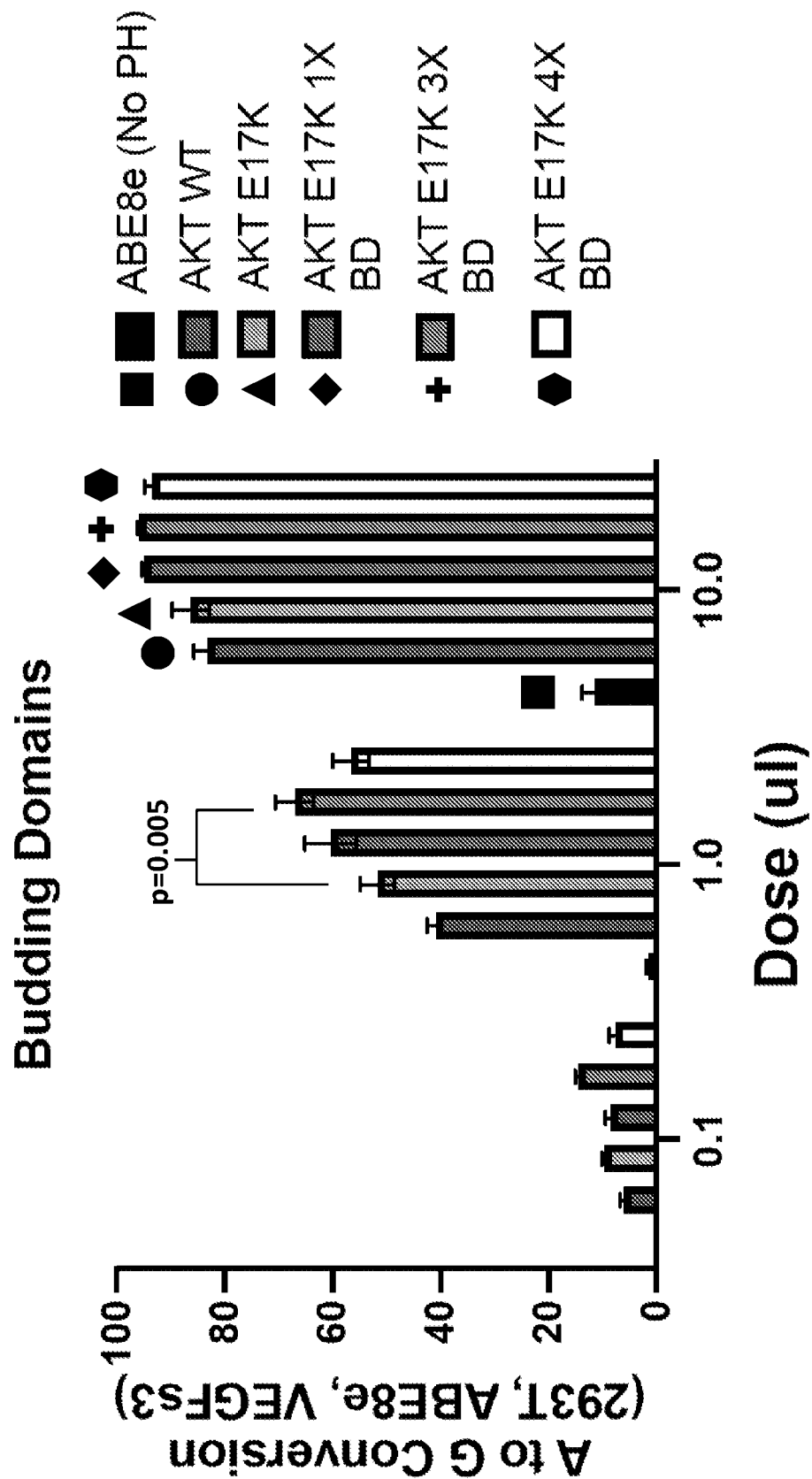


FIG. 8B

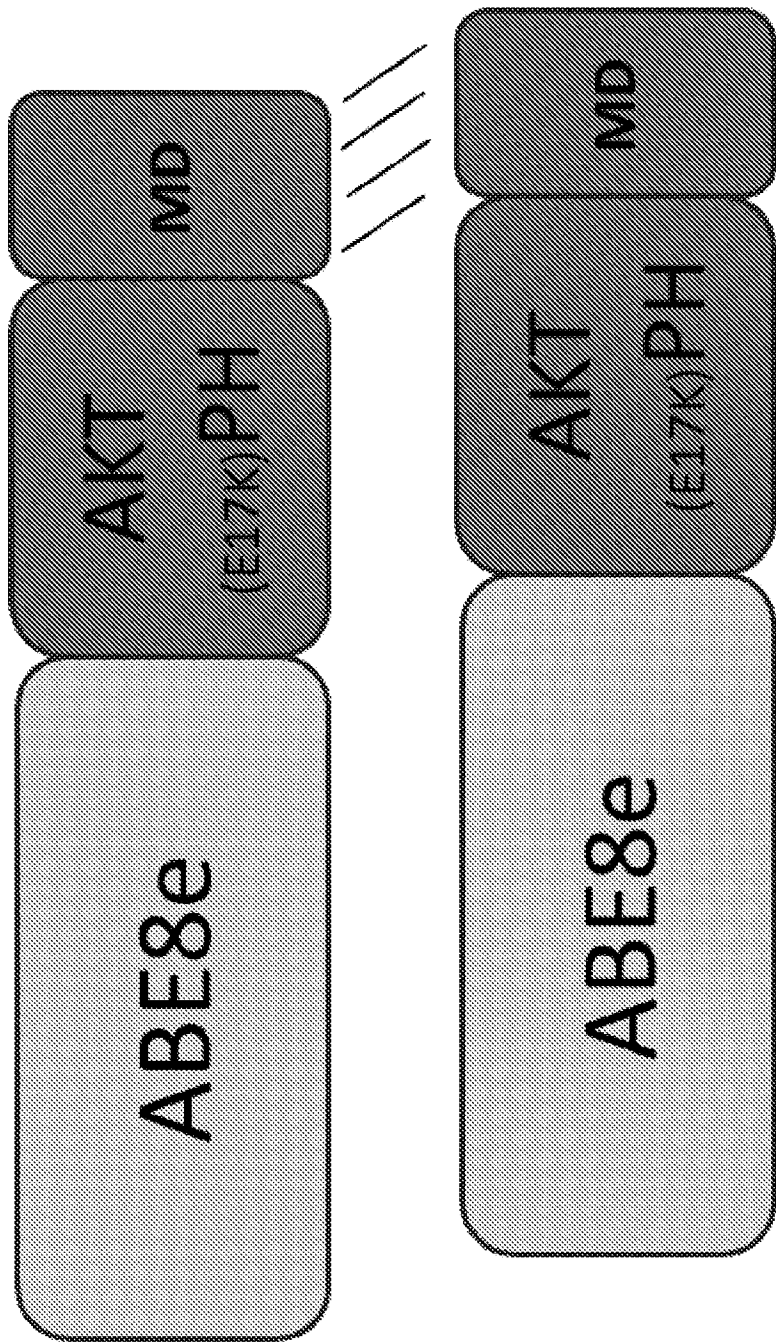


FIG. 9

Cargo/Core:

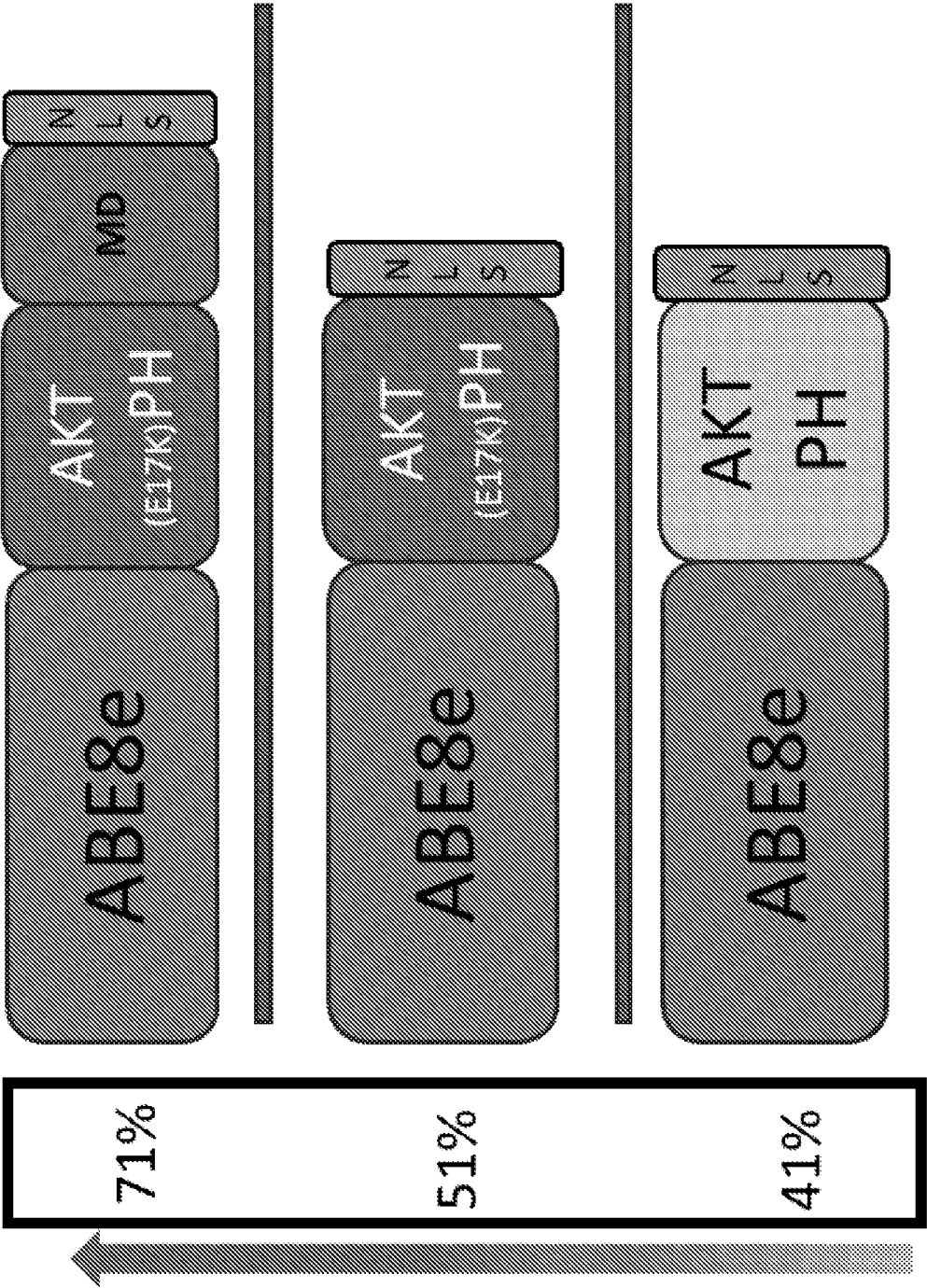


FIG. 10A

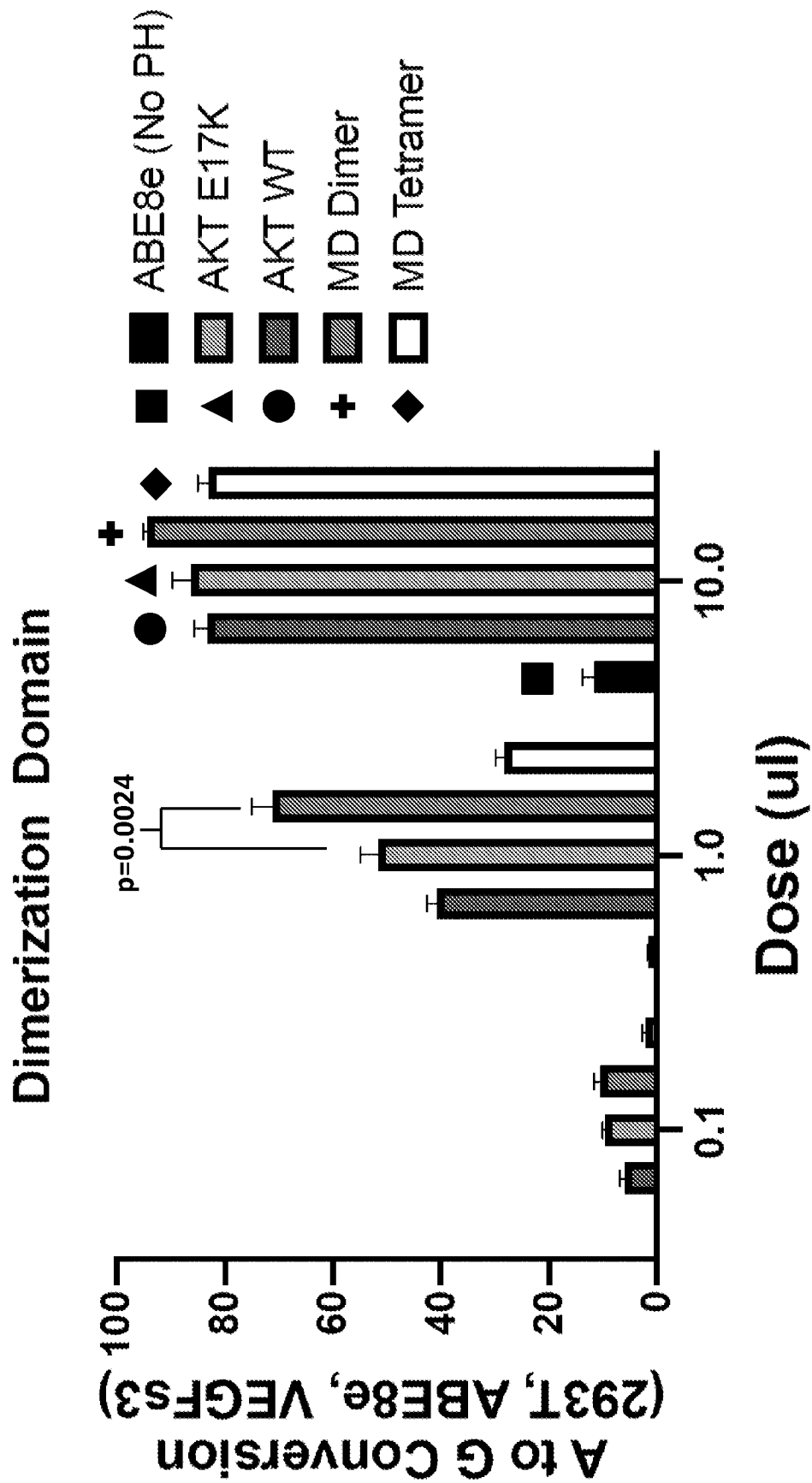


FIG. 10B

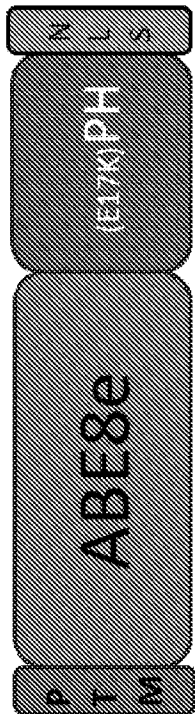


FIG. 11A

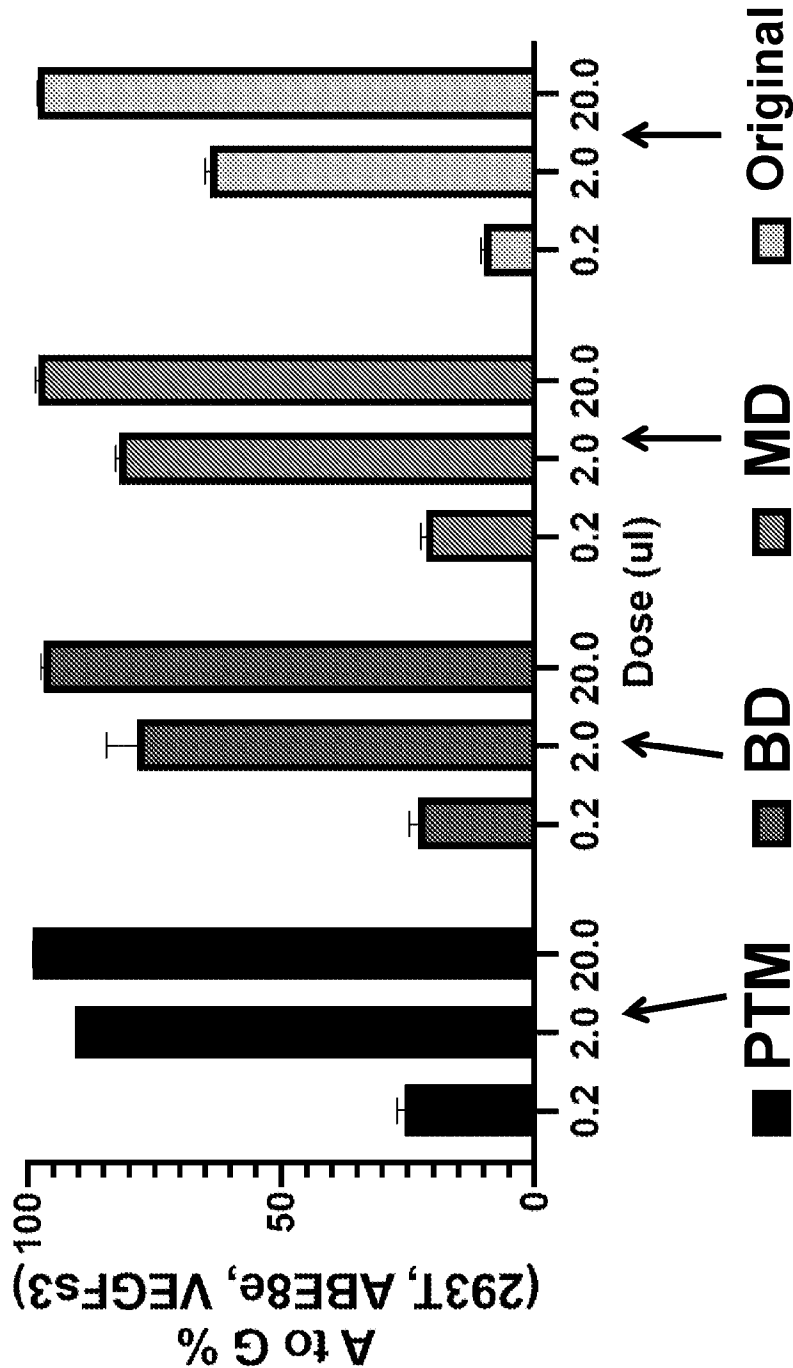


FIG. 11B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033139

A. CLASSIFICATION OF SUBJECT MATTER

IPC: **A61K 47/69** (2024.01); **A61K 9/127** (2024.01); **C07K 14/005** (2024.01); **C07K 14/47** (2024.01); **C07K 19/00** (2024.01); **C12N 15/62** (2024.01); **C12N 15/63** (2024.01); **C12N 5/10** (2024.01); **A61K 39/395** (2024.01)

CPC: **A61K 47/6911**; **C12N 5/10**; **A61K 39/395**; **A61K 9/127**; **C07K 14/005**; **C07K 19/00**; **C07K 14/47**; **C12N 15/62**; **C12N 15/63**; **C07K 2319/03**; **C07K 2319/09**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 11,028,383 B2 (UNIVERSITY OF WASHINGTON et al.) 08 June 2021 (08.06.2021) entire document	1-3, 84-86, 168
A	US 2022/0340889 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 27 October 2022 (27.10.2022) entire document	1-3, 84-86, 168
P, X	WO 2024/107983 A1 (THE GENERAL HOSPITAL CORPORATION) 23 May 2024 (23.05.2024) entire document	1-3, 84-86, 168



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“D” document cited by the applicant in the international application
“E” earlier application or patent but published on or after the international filing date
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
“O” document referring to an oral disclosure, use, exhibition or other means
“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

19 August 2024 (19.08.2024)

Date of mailing of the international search report

03 September 2024 (03.09.2024)

Name and mailing address of the ISA/US

**Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450**

Facsimile No. **571-273-8300**

Authorized officer

**MATOS
TAINA**

Telephone No. **571-272-4300**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033139

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed.
 - b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/033139

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: **4-83, 87-114, 116-167, 169-176**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).