(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 February 2024 (08.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/030976 A2

(51) International Patent Classification:
C07K 9/00 (2006.01) *C07K 19/00* (2006.01)
C07K 16/00 (2006.01)

(21) International Application Number:
PCT/US2023/071544

(22) International Filing Date:
02 August 2023 (02.08.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/394,849 03 August 2022 (03.08.2022) US
63/471,167 05 June 2023 (05.06.2023) US

(71) Applicant: **VOYAGER THERAPEUTICS, INC.** [US/US]; 64 Sidney Street, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **NONNENMACHER, Mathieu Emmanuel**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **MOYER, Tyler Christopher**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **LI, Jiangyu**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **LAKS, Dan Richard**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **HOFFMAN, Brett**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **SHAH, Ishan Sanjeev**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **HEFFNER, Garrett Collins**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **KNOX, Tatiana**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **REN, Xiaoqin**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **PLACIDI, Matteo Placido**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **KNOLL, Elisabeth**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US). **EZELL, Brian Coleman**; c/o Voyager Therapeutics, Inc., 64 Sidney Street, Cambridge, Massachusetts 02139 (US).

(74) Agent: **POTI, Kristin** et al.; K&L Gates LLP, 1 Congress Street, Suite 2900, Boston, Massachusetts 02114 (US).

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

(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:
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: COMPOSITIONS AND METHODS FOR CROSSING THE BLOOD BRAIN BARRIER

(57) Abstract: The disclosure relates to compositions and methods for the preparation, use, and/or formulation of active agents conjugated to ligands for increased crossing of the blood brain barrier.

WO 2024/030976 A2

COMPOSITIONS AND METHODS FOR CROSSING THE BLOOD BRAIN BARRIER

RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 63/394,849 filed on August 3, 2022 and U.S. Provisional Application No. 63/471,167 filed on June 5, 2023; the entire contents of each of which are hereby incorporated by reference in their entirety.

SEQUENCE LISTING

[0001] The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on August 2, 2023, is named V2071-3006PCT_SL.xml and is 4,777,611 bytes in size.

FIELD OF THE DISCLOSURE

[0002] The disclosure relates to compositions and methods for the preparation, use, and/or formulation of active agents conjugated to ligands for increased crossing of the blood brain barrier.

BACKGROUND

[0003] Administering active agents such as therapeutic agents and diagnostic agents to the adult central nervous system (CNS) remains a significant challenge. Engineered compositions comprising said active agents fused or coupled with ligands capable of binding receptors on cells present in the blood brain barrier represent an attractive solution to the limitations of CNS delivery.

[0004] Attempts at providing compositions with improved ability to cross the blood brain barrier, have met with limited success. As such, there is a need for improved methods of producing and delivering active agents of interest to a target cell or tissue, e.g., a CNS cell or tissue.

SUMMARY OF THE DISCLOSURE

[0005] The present disclosure pertains at least in part, to compositions and methods for the production and use of a composition comprising a ligand capable of binding a receptor present on a cell in the blood brain barrier. In some embodiments, the ligand is fused or coupled, e.g., covalently or non-covalently, to an active agent, e.g., a therapeutic agent or a diagnostic agent. Said compositions can be useful for delivery of an active agent, e.g., a therapeutic agent or a diagnostic agent described herein, to a cell or tissue, e.g., a CNS cell or tissue, for the treatment of a disorder, e.g., a neurological or a neurodegenerative disorder, a muscular or a neuromuscular disorder, or a neuro-oncological disorder.

[0006] Accordingly, in one aspect, the present disclosure provides a composition e.g., a fusion molecule or a conjugate molecule, comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an

active agent, e.g., a therapeutic agent or a diagnostic agent, wherein the ligand is fused or coupled, e.g., covalently or non-covalently, to the active agent.

[0007] In another aspect, the present disclosure provides multispecific antibody molecule comprising a first binding domain that binds to ALPL (e.g., an anti-ALPL binding domain) and a second binding domain that binds to a therapeutic target.

[0008] In yet another aspect, the present disclosure provides a method of making a composition described herein, the method comprises (i) providing the ligand that binds to the GPI anchored protein, e.g., ALPL, and the active agent; and (ii) incubating the ligand and active agent under conditions suitable to fuse or couple the ligand to the active agent, thereby generating the composition.

[0009] In yet another aspect, the present disclosure provides a method of delivering an active agent, e.g., a therapeutic agent or a diagnostic agent, to a cell or tissue (e.g., a CNS cell or a CNS tissue). The method comprising administering to the subject an effective amount of a composition comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an active agent, described herein.

[0010] In yet another aspect, the present disclosure provides a method of increasing central nervous system transduction (e.g., increased crossing of the blood brain barrier) in a subject. The method comprising administering to the subject an effective amount of a composition comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an active agent, described herein.

[0011] In yet another aspect, the present disclosure provides a method of treating a subject having or diagnosed with having a genetic disorder, e.g., a monogenic disorder or a polygenic disorder. The method comprising administering to the subject an effective amount of a composition comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an active agent, described herein.

[0012] In yet another aspect, the present disclosure provides a method of treating a subject having or diagnosed with having neurological, e.g., a neurodegenerative, disorder. The method comprising administering an effective amount of a composition comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an active agent, described herein.

[0013] In yet another aspect, the present disclosure provides a method of treating a subject having or diagnosed with having a neuro-oncological disorder. The method comprising administering an effective amount of a composition comprising: (i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and (ii) an active agent, described herein.

[0014] Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following enumerated embodiments.

Enumerated Embodiments

1. A composition, e.g., a fusion molecule or a conjugate molecule, comprising:

(i) a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and

(ii) an active agent, e.g., a therapeutic agent or a diagnostic agent,

wherein the ligand is fused or coupled, e.g., covalently or non-covalently, to the active agent;

optionally wherein the ligand is capable of binding the GPI anchored protein, e.g., ALPL, at a K_D of at least about 10-250 nM, 10-150 nM (e.g., at least 10 nM, 15 nM, 20 nM, 30 nM, 32 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, 110 nM, 120 nM, 130 nM, 140 nM, 150 nM, 175 nM, 200 nM, 215 nM, or 250 nM), e.g., when measured by an SPR assay, e.g., as described in Example 8.

2. The composition of claim 1, wherein the ligand is capable of binding the GPI anchored protein, e.g., ALPL, at a K_D of:

(a) at least about 10-250 nM;

(b) at least about 10-150 nM (e.g., at least 10 nM, 15 nM, 20 nM, 30 nM, 32 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, 110 nM, 120 nM, 130 nM, 140 nM, 150 nM), e.g., wherein the ligand is a viral particle or a peptide;

(c) at least about 10-55 nM, 15-30 nM, 20-30 nM, 25-50 nM, or 30-50 nM (e.g., at least 10 nM, 15 nM, 20 nM, 30 nM, 32 nM, 50 nM, or 55 nM), e.g., wherein the ligand is a viral particle (e.g., an AAV viral particle) or a peptide; or

(c) at least about 150-250 nM, 150-225 nM, 175-250 nM, 175-225 nM, 200-225 nM, 200-250 nM (e.g., 150 nM, 175 nM, 200 nM, 215 nM, or 250 nM), e.g., wherein the ligand is an antibody molecule;

optionally, when (a), (b), (c), and (d) are measured by an SPR assay, e.g., as described in Example 8 or 13.

3. The composition of claim 1 or 2, wherein the ligand is capable of binding the GPI anchored protein e.g., ALPL, in a pH dependent manner, optionally wherein the ligand binds to ALPL at physiological pH (e.g., at a pH of at least about 6.5-8.0, 7.0-8.0, 6.5-7.5, 7.0-7.5, 7.0, 7.1, 7.2, 7.3, or 7.4) and/or does not substantially bind ALPL at an acidic pH (e.g., at a pH of at least about 1.0-5.7, 1.0-5.5, 2.0-5.7, 2.5-5.5, 2.5-5.7, 3.0-5.7, 3.0-5.5, 3.5-5.7, 3.5-5.5, 4.0-5.7, 4.0-5.5, 4.5-5.7, 4.5-5.5, 5.0-5.7, 5.5-5.7, 5.0, 5.1, 5.2, 5.3, 5.4, or 5.5), e.g., as measured by an assay e.g., an SPR or Biacore assay, e.g., as described in Example 8 or 13.

4. The composition of any one of embodiments 1-3, wherein the ligand is or comprises a peptide, a protein, an antibody molecule, a nucleic acid molecule (e.g., an aptamer), or a small molecule.

5. The composition of any one of embodiments 1-3, wherein the ligand comprises a linear peptide or a circular peptide.

6. The composition of any one of embodiments 1-5, wherein the active agent is or comprises a therapeutic agent chosen from a protein (e.g., an enzyme), an antibody molecule, a nucleic acid molecule (e.g., an RNAi agent), or a small molecule.

7. The composition of any one of embodiments 1-5, wherein the active agent is or comprises a ribonucleic acid complex (e.g., a Cas9/gRNA complex), a plasmid, a closed-end DNA, a circ-RNA, or an mRNA.

8. The composition of any one of embodiments 1-5, wherein the active agent is a diagnostic agent is or comprises an imaging agent (e.g., a protein or small molecule compound coupled to a detectable moiety).

9. The composition of any one of embodiments 1-8, wherein the ligand is covalently linked to the active agent.

10. The composition of any one of embodiments 1-9, wherein the ligand is conjugated to the active agent.

11. The composition of any one of embodiments 1-8, wherein the ligand is fused to the active agent, e.g., as part of a fusion peptide or protein.

12. The composition of any one of embodiments 1-11, wherein the ligand is not a component of a viral particle, e.g., an adeno-associated viral (AAV) particle.

13. The composition of any one of embodiments 1-12, wherein the ligand is not a component of a capsid protein, e.g., an AAV capsid protein.

14. The composition of embodiment 13, wherein the ligand is not a component of an AAV9 capsid or variant thereof.

15. The composition of any one of embodiments 1-14, wherein the GPI anchored protein is conserved in at least two to three species, e.g., at least three species (e.g., mice, NHPs (e.g., *Macaca fascicularis*), and/or humans).

16. The composition of embodiment of embodiment 15, wherein the at least two GPI anchored proteins are at least 80%, 85%, 90%, 95%, 99%, or 100% identical to each other.

17. The composition of any one of embodiments 1-16, wherein the GPI anchored protein is present on the surface of a cell in the blood brain barrier.

18. The composition of any one of embodiments 1-17, wherein the GPI anchored protein is ALPL, CD59, LY6E, CA4, GPC5, NTM, HYAL2, LSAMP, BST2, EMP2, ALPL, CPM, NCAM1, EFNA1, PIBF1, SEC24B, PRNP, TFPI, OPCML, CD109, DPM3, CNTN4, PIGN, HBP1, CNTN2, CD55, NEGR1, EFNA5, RECK, NRN1, CNTN1, GPAA1, PGAP1, PIGF, PIGK, MDGA2, DPM1, SVIP, NTNG1, CNTN5, GPC6, PIGG, TMEM8A, THY1, GPIHBP1, PIGT, PIGL, ZFAND2B, PLAUR, DPM2, or GPC1.

19. The composition of any one of embodiments 1-18, wherein the GPI anchored protein is ALPL.

20. The composition of any one of embodiments 1-19, wherein the ligand binds human, cynomolgus, or murine ALPL.

21. The composition of any one of embodiments 1-20, wherein ligand is fused or coupled to a therapeutic agent or diagnostic agent.

22. The composition of any one of embodiments 1-21, wherein the ligand is covalently linked, e.g., directly or indirectly via a linker, to the active agent.

23. The composition of embodiment 22, wherein the ligand is covalently linked to the active agent via a linker.

24. The composition of any one of embodiments 1-23, wherein the ligand is conjugated, e.g., directly or indirectly via linker, to the active agent.

25. The composition of 24, wherein the ligand is conjugated to the active agent via a linker.

26. The composition of any one of embodiments 22-25, wherein the linker is a cleavable linker or a non-cleavable linker.

27. The composition of embodiment 26, wherein the cleavable linker is a pH sensitive linker or an enzyme sensitive linker.

28. The composition of embodiment 27, wherein the pH sensitive linker comprises a hydrazine/hydrazone linker or a disulfide linker.
29. The composition of embodiment 28, wherein the enzyme sensitive linker comprises a peptide based linker, e.g., a peptide linker sensitive to a protease (e.g., a lysosomal protease); or a beta-glucuronide linker.

30. The composition of embodiment 26, wherein the non-cleavable linker is a linker comprising a thioether group or a maleimidocaproyl group.

31. The composition of any one of embodiments 1-23, wherein the ligand is fused, e.g., directly or indirectly via linker, to the active agent, e.g., as part of a fusion peptide or protein.

32. The composition of any one of embodiments 1-31, wherein the ligand and the active agent are fused or coupled post-translationally, e.g., using click chemistry.

33. The composition of any one of embodiments 1-32, wherein the ligand and the active agent are fused or couple via chemically induced dimerization.

34. The composition of any one of embodiments 1-33, wherein the ligand is present N-terminal relative to the active agent.

35. The composition of any one of embodiments 1-33, wherein the ligand is present C-terminal relative to the active agent.

36. The composition of any one of embodiments 1-33, wherein the ligand is fused or coupled at or near the C-terminus of the active agent, wherein the active agent is a therapeutic protein, enzyme, or antibody molecule.

37. The composition of embodiment 36, wherein the ligand is fused or coupled within 20, 30, 40, 50, 60, 70, 80, 90, 100, or more amino acids from the C-terminus of the therapeutic protein, enzyme, or antibody molecule.

38. The composition of any one of embodiments 1-36, wherein the ligand is or comprises a protein or a peptide comprising an amino acid sequence having the following formula: [N1]-[N2]-[N3], wherein:

(i) optionally [N1] comprises X1, X2, and X3, wherein at least one of X1, X2, or X3 is G;

(ii) [N2] comprises the amino acid sequence of SPH, optionally wherein S comprises a modification, e.g., comprises a phosphate group;

(ii) [N3] comprises X4, X5, and X6, wherein at least one of X4, X5, or X6 is a basic amino acid, e.g., a K or R.

39. The composition of embodiment 38, wherein X4, X5, or both of [N3] is a K.

40. The composition of embodiment 38 or 39, wherein X4, X5, or X6 of [N3] is an R.

41. The composition of any one of embodiments 38-40, wherein:

(a) position X4 of [N3] is independently chosen from: K, S, A, V, T, G, F, W, V, N, or R;

(b) position X5 of [N3] is independently chosen from: S, K, T, F, I, L, Y, H, M, or R; and/or

(c) position X6 of [N3] is independently chosen from: G, A, R, M, I, N, T, Y, D, P, V, L, E, W, N, Q, K, or S;

optionally, wherein the protein or peptide comprises an amino acid modification, e.g., a conservative substitution, of any of the aforesaid amino acids in (a)-(c).

42. The composition of any one of embodiments 38-41, wherein [N3] comprises SK, KA, KS, AR, RM, VK, AS, SR, VK, KR, KK, KN, VR, RS, RK, KT, TS, KF, FG, KI, IG, KL, LG, TT, TY, KY, YG, KD, KP, TR, RG, VR, GA, SL, SS, FL, WK, SA, RA, LR, KW, RR, GK, TK, NK, AK, KV, KG, KH, KM, TG, SE, SV, SW, SN, HG, SQ, LW, MG, MA, or SG.

43. The composition of any one of embodiments 38-42, wherein [N3] is SKA, KSG, ARM, VKS, ASR, VKI, KKN, VRM, RKA, KTS, KFG, KIG, KLG, KTT, KTY, KYG, SKD, SKP, TRG, VRG, KRG, GAR, KSA, KSR, SKL, SRA, SKR, SLR, SRG, SSR, FLR, SKW, SKS, WKA, VRR, SKV, SKT, SKG, GKA, TKA, NKA, SKL, SKN, AKA, KTG, KSL, KSE, KSV, KSW, KSN, KHG, KSQ, KSK, KLW, WKG, KMG, KMA, or RSG.

44. The composition of any one of embodiments 38-43, wherein [N2]-[N3] comprises SPHSK (SEQ ID NO: 4701), SPHKS (SEQ ID NO: 4704), SPHAR (SEQ ID NO: 4705), SPHVK (SEQ ID NO: 4706), SPHAS (SEQ ID NO: 4707), SPHKK (SEQ ID NO: 4708), SPHVR (SEQ ID NO: 4709), SPHRK (SEQ ID NO: 4710), SPHKT (SEQ ID NO: 4711), SPHKF (SEQ ID NO: 4712), SPHKI (SEQ ID NO: 4713), SPHKL (SEQ ID NO: 4714), SPHKY (SEQ ID NO: 4715), SPHTR (SEQ ID NO: 4716), SPHKR (SEQ ID NO: 4717), SPHGA (SEQ ID NO: 4718), SPHSR (SEQ ID NO: 4719), SPHSL (SEQ ID NO: 4720), SPHSS (SEQ ID NO: 4721), SPHFL (SEQ ID NO: 4722), SPHWK

(SEQ ID NO: 4723), SPHGK (SEQ ID NO: 4724), SPHTK (SEQ ID NO: 4725), SPHNK (SEQ ID NO: 4726), SPHAK (SEQ ID NO: 4727), SPHKH (SEQ ID NO: 4728), SPHKM (SEQ ID NO: 4729), or SPHRS (SEQ ID NO: 4730).

45. The composition of any one of embodiments 38-44, wherein [N2]-[N3] is or comprises:

(i) SPHSKA (SEQ ID NO: 941), SPHKSG (SEQ ID NO: 946), SPHARM (SEQ ID NO: 947), SPHVKS (SEQ ID NO: 948), SPHASR (SEQ ID NO: 949), SPHVKI (SEQ ID NO: 950), SPHKKN (SEQ ID NO: 954), SPHVRM (SEQ ID NO: 955), SPHRKA (SEQ ID NO: 956), SPHKFG (SEQ ID NO: 957), SPHKIG (SEQ ID NO: 958), SPHKLG (SEQ ID NO: 959), SPHKTS (SEQ ID NO: 963), SPHKTT (SEQ ID NO: 964), SPHKTY (SEQ ID NO: 965), SPHKYG (SEQ ID NO: 966), SPHSKD (SEQ ID NO: 967), SPHSKP (SEQ ID NO: 968), SPHTRG (SEQ ID NO: 972), SPHVRG (SEQ ID NO: 973), SPHKRG (SEQ ID NO: 974), SPHGAR (SEQ ID NO: 975), SPHKSA (SEQ ID NO: 977), SPHKSR (SEQ ID NO: 951), SPHSKL (SEQ ID NO: 960), SPHSRA (SEQ ID NO: 969), SPHSKR (SEQ ID NO: 978), SPHSLR (SEQ ID NO: 952), SPHSRG (SEQ ID NO: 961), SPHSSR (SEQ ID NO: 970), SPHFLR (SEQ ID NO: 979), SPHSKW (SEQ ID NO: 953), SPHSKS (SEQ ID NO: 962), SPHWKA (SEQ ID NO: 971), SPHVRR (SEQ ID NO: 980), SPHSKT (SEQ ID NO: 4731), SPHSKG (SEQ ID NO: 4732), SPHGKA (SEQ ID NO: 4733), SPHNKA (SEQ ID NO: 4734), SPHSKN (SEQ ID NO: 4735), SPHAKA (SEQ ID NO: 4736), SPHSKV (SEQ ID NO: 4737), SPHKTG (SEQ ID NO: 4738), SPHTKA (SEQ ID NO: 4739), SPHKSL (SEQ ID NO: 4740), SPHKSE (SEQ ID NO: 4741), SPHKSV (SEQ ID NO: 4742), SPHKSW (SEQ ID NO: 4743), SPHKSN (SEQ ID NO: 4744), SPHKHG (SEQ ID NO: 4745), SPHKSQ (SEQ ID NO: 4746), SPHKSK (SEQ ID NO: 4747), SPHKLW (SEQ ID NO: 4748), SPHWKG (SEQ ID NO: 4749), SPHKMG (SEQ ID NO: 4750), SPHKMA (SEQ ID NO: 4751), or SPHRSG (SEQ ID NO: 976);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, or 5 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

46. The composition of any one of embodiments 38-45, wherein [N1] comprises X1, X2, and X3, wherein at least one of X1, X2, or X3 is G.

47. The composition of any one of embodiments 38-46, wherein:

(a) position X1 of [N1] is independently chosen from: G, V, R, D, E, M, T, I, S, A, N, L, K, H, P, W, or C;

(b) position X2 of [N1] is independently chosen from: S, V, L, N, D, H, R, P, G, T, I, A, E, Y, M, or Q; and/or

(c) position X3 of [N1] is independently chosen from: G, C, L, D, E, Y, H, V, A, N, P, or S;

optionally wherein the protein or peptide comprises an amino acid modification, e.g., a conservative substitution, of any of the aforesaid amino acids in (a)-(c).

48. The ligand of any one of embodiments 38-47, wherein [N1] comprises GS, SG, GH, HD, GQ, QD, VS, CS, GR, RG, QS, SH, MS, RN, TS, IS, GP, ES, SS, GN, AS, NS, LS, GG, KS, GT, PS, RS, GI, WS, DS, ID, GL, DA, DG, ME, EN, KN, KE, AI, NG, PG, TG, SV, IG, LG, AG, EG, SA, YD, HE, HG, RD, ND, PD, MG, QV, DD, HN, HP, GY, GM, GD, or HS.

49. The composition of any one of embodiments 38-48, wherein [N1] is or comprises GSG, GHD, GQD, VSG, CSG, GRG, CSH, GQS, GSH, RVG, GSC, GLL, GDD, GHE, GNY, MSG, RNG, TSG, ISG, GPG, ESG, SSG, GNG, ASG, NSG, LSG, GGG, KSG, HSG, GTG, PSG, GSV, RSG, GIG, WSG, DSG, IDG, GLG, DAG, DGG, MEG, ENG, GSA, KNG, KEG, AIG, GYD, GHG, GRD, GND, GPD, GMG, GQV, GHN, GHP, or GHS.

50. The composition of any one of embodiments 38-49, wherein [N1]-[N2] comprises:

(i) SGSPH (SEQ ID NO: 4752), HDSPH (SEQ ID NO: 4703), QDSPH (SEQ ID NO: 4753), RGSPH (SEQ ID NO: 4754), SHSPH (SEQ ID NO: 4755), QSSPH (SEQ ID NO: 4756), DDSPH (SEQ ID NO: 4757), HESPH (SEQ ID NO: 4758), NYSPH (SEQ ID NO: 4759), VGSPH (SEQ ID NO: 4760), SCSPH (SEQ ID NO: 4761), LLSPH (SEQ ID NO: 4762), NGSPH (SEQ ID NO: 4763), PGSPH (SEQ ID NO: 4764), GGSPH (SEQ ID NO: 4765), TGSPH (SEQ ID NO: 4766), SVSPH (SEQ ID NO: 4767), IGSPH (SEQ ID NO: 4768), DGSPH (SEQ ID NO: 4769), LGSPH (SEQ ID NO: 4770), AGSPH (SEQ ID NO: 4771), EGSPH (SEQ ID NO: 4772), SASPH (SEQ ID NO: 4773), YDSPH (SEQ ID NO: 4774), HGSPH (SEQ ID NO: 4775), RDSPH (SEQ ID NO: 4776), NDSPH (SEQ ID NO: 4777), PDSPH (SEQ ID NO: 4778), MGSPH (SEQ ID NO: 4779), QVSPH (SEQ ID NO: 4780), HNSPH (SEQ ID NO: 4781), HPSPH (SEQ ID NO: 4782), or HSSPH (SEQ ID NO: 4783);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, or 4 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

51. The composition of any one of embodiments 38-50, wherein [N1]-[N2] is or comprises:

(i) GSGSPH (SEQ ID NO: 4695), GHDSPH (SEQ ID NO: 4784), GQDSPH (SEQ ID NO: 4785), VSGSPH (SEQ ID NO: 4786), CSGSPH (SEQ ID NO: 4787), GRGSPH (SEQ ID NO: 4788), CSHSPH (SEQ ID NO: 4789), GQSSPH (SEQ ID NO: 4790), GSHSPH (SEQ ID NO: 4791), GDDSPH (SEQ ID NO: 4792), GHESPH (SEQ ID NO: 4793), GNYSPH (SEQ ID NO: 4794), RVGSPH (SEQ ID NO: 4795), GSCSPH (SEQ ID NO: 4796), GLLSPH (SEQ ID NO: 4797), MSGSPH (SEQ ID NO: 4798), RNGSPH (SEQ ID NO: 4799), TSGSPH (SEQ ID NO: 4800), ISGSPH (SEQ ID NO: 4801), GPGSPH (SEQ ID NO: 4802), ESGSPH (SEQ ID NO: 4803), SSGSPH (SEQ ID NO: 4804), GNGSPH (SEQ ID NO: 4805), ASGSPH (SEQ ID NO: 4806), NSGSPH (SEQ ID NO: 4807), LSGSPH (SEQ ID NO: 4808), GGGSPH (SEQ ID NO: 4809), KSGSPH (SEQ ID NO: 4810), HSGSPH (SEQ ID NO: 4811), GTGSPH (SEQ ID NO: 4812), PSGSPH (SEQ ID NO: 4813), GSVSPH (SEQ ID NO: 4814), RSGSPH (SEQ ID NO: 4815), GIGSPH (SEQ ID NO: 4816), WSGSPH (SEQ ID NO: 4817), DSGSPH (SEQ ID NO: 4818), IDGSPH (SEQ ID NO: 4819), GLGSPH (SEQ ID NO: 4820), DAGSPH (SEQ ID NO: 4821), DGGSPH (SEQ ID NO: 4822), MEGSPH (SEQ ID NO: 4823), ENGSPH (SEQ ID NO: 4824), GSASPH (SEQ ID NO: 4825), KNGSPH (SEQ ID NO: 4826), KEGSPH (SEQ ID NO: 4827), AIGSPH (SEQ ID NO: 4828), GYDSPH (SEQ ID NO: 4829), GHGSPH (SEQ ID NO: 4830), GRDSPH (SEQ ID NO: 4831), GNDSPH (SEQ ID NO: 4832), GPDSPH (SEQ ID NO: 4833), GMGSPH (SEQ ID NO: 4834), GQVSPH (SEQ ID NO: 4835), GHNSPH (SEQ ID NO: 4836), GHPSPH (SEQ ID NO: 4837), or GHSSPH (SEQ ID NO: 4838);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, or 5 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

52. The composition of any one of embodiments 38-51, wherein [N1]-[N2]-[N3] comprises:

(i) SGSPHSK (SEQ ID NO: 4839), HDSPHKS (SEQ ID NO: 4840), SGSPHAR (SEQ ID NO: 4841), SGSPHVK (SEQ ID NO: 4842), QDSPHKS (SEQ ID NO: 4843), SGSPHKK (SEQ ID NO: 4844), SGSPHVR (SEQ ID NO: 4845), SGSPHAS (SEQ ID NO: 4846), SGSPHRK (SEQ ID NO: 4847), SGSPHKT (SEQ ID NO: 4848), SHSPHKS (SEQ ID NO: 4849), QSSPHRS (SEQ ID NO: 4850), RGSPHAS (SEQ ID NO: 4851), RGSPHSK (SEQ ID NO: 4852), SGSPHKF (SEQ ID NO: 4853), SGSPHKI (SEQ ID NO: 4854), SGSPHKL (SEQ ID NO: 4855), SGSPHKY (SEQ ID NO: 4856), SGSPHTR (SEQ ID NO: 4857), SHSPHKR (SEQ ID NO: 4858), SGSPHGA (SEQ ID NO: 4859), HDSPHKR (SEQ ID NO: 4860), DDSPHKS (SEQ ID NO: 4861), HESPHKS (SEQ ID

NO: 4862), NYSPHKI (SEQ ID NO: 4863), SGSPHSR (SEQ ID NO: 4864), SGSPHSL (SEQ ID NO: 4865), SGSPHSS (SEQ ID NO: 4866), VGSPHSK (SEQ ID NO: 4867), SCSPHRK (SEQ ID NO: 4868), SGSPHFL (SEQ ID NO: 4869), LLSPHWK (SEQ ID NO: 4870), NGSPHSK (SEQ ID NO: 4871), PGSPHSK (SEQ ID NO: 4872), GGSPHSK (SEQ ID NO: 4873), TGSPHSK (SEQ ID NO: 4874), SVSPHGK (SEQ ID NO: 4875), SGSPHTK (SEQ ID NO: 4876), IGSPHSK (SEQ ID NO: 4877), DGSPHSK (SEQ ID NO: 4878), SGSPHNK (SEQ ID NO: 4879), LGSPHSK (SEQ ID NO: 4880), AGSPHSK (SEQ ID NO: 4881), EGSPHSK (SEQ ID NO: 4882), SASPHSK (SEQ ID NO: 4883), SGSPHAK (SEQ ID NO: 4884), HDSPHKI (SEQ ID NO: 4885), YDSPHKS (SEQ ID NO: 4886), HDSPHKT (SEQ ID NO: 4887), RGSPHKR (SEQ ID NO: 4888), HGSPHSK (SEQ ID NO: 4889), RDSPHKS (SEQ ID NO: 4890), NDSPHKS (SEQ ID NO: 4891), QDSPHKI (SEQ ID NO: 4892), PDSPHKI (SEQ ID NO: 4893), PDSPHKS (SEQ ID NO: 4894), MGSPHSK (SEQ ID NO: 4895), HDSPHKH (SEQ ID NO: 4896), QVSPHKS (SEQ ID NO: 4897), HNSPHKS (SEQ ID NO: 4898), NGSPHKR (SEQ ID NO: 4899), HDSPHKY (SEQ ID NO: 4900), NDSPHKI (SEQ ID NO: 4901), HDSPHKL (SEQ ID NO: 4902), HPSPHWK (SEQ ID NO: 4903), HDSPHKM (SEQ ID NO: 4904), or HSSPHRS (SEQ ID NO: 4905);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, 5, or 6 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

53. The composition of any one of embodiments 38-52, wherein [N1]-[N2]-[N3] is or comprises:

(i) GSGSPHSKA (SEQ ID NO: 4697), GHDSPHKSG (SEQ ID NO: 4698), GSGSPHARM (SEQ ID NO: 4906), GSGSPHVKS (SEQ ID NO: 4907), GQDSPHKSG (SEQ ID NO: 4908), GSGSPHASR (SEQ ID NO: 4909), GSGSPHVKI (SEQ ID NO: 4910), GSGSPHKKN (SEQ ID NO: 4911), GSGSPHVRM (SEQ ID NO: 4912), VSGSPHSKA (SEQ ID NO: 4913), CSGSPHSKA (SEQ ID NO: 4914), GSGSPHRKA (SEQ ID NO: 4915), CSGSPHKTS (SEQ ID NO: 4916), CSHSPHKSG (SEQ ID NO: 4917), GQSSPHRSG (SEQ ID NO: 4918), GRGSPHASR (SEQ ID NO: 4919), GRGSPHSKA (SEQ ID NO: 4920), GSGSPHKFG (SEQ ID NO: 4921), GSGSPHKIG (SEQ ID NO: 4922), GSGSPHKLG (SEQ ID NO: 4923), GSGSPHKTS (SEQ ID NO: 4924), GSGSPHKTT (SEQ ID NO: 4925), GSGSPHKTY (SEQ ID NO: 4926), GSGSPHKYG (SEQ ID NO: 4927), GSGSPHSKD (SEQ ID NO: 4928), GSGSPHSKP (SEQ ID NO: 4929), GSGSPHTRG (SEQ ID NO: 4930), GSGSPHVRG (SEQ ID NO: 4931), GSHSPHKRG (SEQ ID NO: 4932), GSHSPHKSG (SEQ ID NO: 4933), VSGSPHASR (SEQ ID NO: 4934), VSGSPHGAR (SEQ ID NO: 4935), VSGSPHKFG (SEQ ID NO: 4936), GHDSPHKRG (SEQ ID NO: 4937), GDDSPHKSG (SEQ

ID NO: 4938), GHESPHKSA (SEQ ID NO: 4939), GHDSPHKSA (SEQ ID NO: 4940), GNYSPHKIG (SEQ ID NO: 4941), GHDSPHKSR (SEQ ID NO: 4942), GSGSPHSKL (SEQ ID NO: 4943), GSGSPHSRA (SEQ ID NO: 4944), GSGSPHSKR (SEQ ID NO: 4945), GSGSPHSLR (SEQ ID NO: 4946), GSGSPHSRG (SEQ ID NO: 4947), GSGSPHSSR (SEQ ID NO: 4948), RVGSPHSKA (SEQ ID NO: 4949), GSCSPHRKA (SEQ ID NO: 4950), GSGSPHFLR (SEQ ID NO: 4951), GSGSPHSKW (SEQ ID NO: 4952), GSGSPHSKS (SEQ ID NO: 4953), GLLSPHWKA (SEQ ID NO: 4954), GSGSPHVRR (SEQ ID NO: 4955), GSGSPHSKV (SEQ ID NO: 4956), MSGSPHSKA (SEQ ID NO: 4957), RNGSPHSKA (SEQ ID NO: 4958), TSGSPHSKA (SEQ ID NO: 4959), ISGSPHSKA (SEQ ID NO: 4960), GPGSPHSKA (SEQ ID NO: 4961), GSGSPHSKT (SEQ ID NO: 4962), ESGSPHSKA (SEQ ID NO: 4963), SSGSPHSKA (SEQ ID NO: 4964), GNGSPHSKA (SEQ ID NO: 4965), ASGSPHSKA (SEQ ID NO: 4966), NSGSPHSKA (SEQ ID NO: 4967), LSGSPHSKA (SEQ ID NO: 4968), GGGSPHSKA (SEQ ID NO: 4969), KSGSPHSKA (SEQ ID NO: 4970), GGGSPHSKS (SEQ ID NO: 4971), GSGSPHSKG (SEQ ID NO: 4972), HSGSPHSKA (SEQ ID NO: 4973), GTGSPHSKA (SEQ ID NO: 4974), PSGSPHSKA (SEQ ID NO: 4975), GSVSPHGKA (SEQ ID NO: 4976), RSGSPHSKA (SEQ ID NO: 4977), GSGSPHTKA (SEQ ID NO: 4978), GIGSPHSKA (SEQ ID NO: 4979), WSGSPHSKA (SEQ ID NO: 4980), DSGSPHSKA (SEQ ID NO: 4981), IDGSPHSKA (SEQ ID NO: 4982), GSGSPHNKA (SEQ ID NO: 4983), GLGSPHSKS (SEQ ID NO: 4984), DAGSPHSKA (SEQ ID NO: 4985), DGGSPHSKA (SEQ ID NO: 4986), MEGSPHSKA (SEQ ID NO: 4987), ENGSPHSKA (SEQ ID NO: 4988), GSASPHSKA (SEQ ID NO: 4989), GNGSPHSKS (SEQ ID NO: 4990), KNGSPHSKA (SEQ ID NO: 4991), KEGSPHSKA (SEQ ID NO: 4992), AIGSPHSKA (SEQ ID NO: 4993), GSGSPHSKN (SEQ ID NO: 4994), GSGSPHAKA (SEQ ID NO: 4995), GHDSPHKIG (SEQ ID NO: 4996), GYDSPHKSG (SEQ ID NO: 4997), GHESPHKSG (SEQ ID NO: 4998), GHDSPHKTG (SEQ ID NO: 4999), GRGSPHKRG (SEQ ID NO: 5000), GQDSPHKSG (SEQ ID NO: 4908), GHDSPHKSL (SEQ ID NO: 5001), GHGSPHSKA (SEQ ID NO: 5002), GHDSPHKSE (SEQ ID NO: 5003), VSGSPHSKA (SEQ ID NO: 4913), GRDSPHKSG (SEQ ID NO: 5004), GNDSPHKSV (SEQ ID NO: 5005), GQDSPHKIG (SEQ ID NO: 5006), GHDSPHKSV (SEQ ID NO: 5007), GPDSPHKIG (SEQ ID NO: 5008), GPDSPHKSG (SEQ ID NO: 5009), GHDSPHKSW (SEQ ID NO: 5010), GHDSPHKSN (SEQ ID NO: 5011), GMGSPHSKT (SEQ ID NO: 5012), GHDSPHKHG (SEQ ID NO: 5013), GQVSPHKSG (SEQ ID NO: 5014), GDDSPHKSV (SEQ ID NO: 5015), GHNSPHKSG (SEQ ID NO: 5016), GNGSPHKRG (SEQ ID NO: 5017), GHDSPHKYG (SEQ ID NO: 5018), GHDSPHKSQ (SEQ ID NO: 5019), GNDSPHKIG (SEQ ID NO: 5020), GHDSPHKSK (SEQ ID NO: 5021), GHDSPHKLW (SEQ ID NO: 5022), GHPSPHWKG (SEQ ID NO: 5023), GHDSPHKMG (SEQ ID NO: 5024), GHDSPHKMA (SEQ ID NO: 5025), or GHSSPHRSG (SEQ ID NO: 5026);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, 5, 6, 7, or 8 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

54. The composition of any one of embodiments 38-53, wherein [N3] comprises SK, KA, KS, or SG.

55. The composition of any one of embodiments 38-54, wherein [N3] is or comprises SKA, KSG, or KYG.

56. The composition of any one of embodiments 38-55, wherein [N2]-[N3] comprises SPHSK (SEQ ID NO: 4701), SPHKS (SEQ ID NO: 4704), or SPHKY (SEQ ID NO: 4715).

57. The composition of any one of embodiments 38-56, wherein [N2]-[N3] is or comprises SPHSKA (SEQ ID NO: 941).

58. The composition of any one of embodiments 38-56, wherein [N2]-[N3] is or comprises SPHKSG (SEQ ID NO: 946).

59. The composition of any one of embodiments 38-56, wherein [N2]-[N3] is or comprises SPHKYG (SEQ ID NO: 966).

60. The composition of any one of embodiments 38-59, wherein [N1] comprises GS, SG, GH, or HD.

61. The composition of any one of embodiments 38-60, wherein [N1] is or comprises GSG.

62. The composition of any one of embodiments 38-60, wherein [N1] is or comprises GHD.

63. The composition of any one of embodiments 38-57, 60, or 61, wherein [N1]-[N2]-[N3] comprises SGSPHSK (SEQ ID NO: 4839).

64. The composition of any one of embodiments 38-56, 58, 60, or 62, wherein [N1]-[N2]-[N3] comprises HDSPHKS (SEQ ID NO: 4840).

65. The composition of any one of embodiments 38-56 or 59-61, wherein [N1]-[N2]-[N3] comprises SGSPHKYG (SEQ ID NO: 5027).

66. The composition of any one of embodiments 38-57, 60, 61, or 63, wherein [N1]-[N2]-[N3] is or comprises GSGSPHSKA (SEQ ID NO: 4697).

67. The composition of any one of embodiments 38-56, 58, 60, 62, or 54, wherein [N1]-[N2]-[N3] is or comprises GHDSPHKSG (SEQ ID NO: 4698).

68. The composition of any one of embodiments 38-56, 59-61, or 65, wherein [N1]-[N2]-[N3] is or comprises GSGSPHKYG (SEQ ID NO: 4927).

69. The composition of any one of embodiments 38-68, which further comprises [N4], wherein [N4] comprises X7 X8 X9 X10, and wherein:

(a) position X7 is independently chosen from Q, W, K, R, G, L, V, S, P, H, K, I, M, A, E, or F;

(b) position X8 is independently chosen from N, Y, C, K, T, H, R, D, V, S, P, G, W, E, F, A, I, M, Q, or L;

(c) position X9 is independently chosen from Q, G, K, H, R, T, L, D, A, P, I, F, V, M, W, Y, S, E, N, or Y; and

(d) position X10 is independently chosen from Q, H, L, R, W, K, A, P, E, M, I, S, G, N, Y, C, V, T, D, or V;

optionally wherein the protein comprises an amino acid modification, e.g., a conservative substitution, of any of the aforesaid amino acids in (a)-(d).

70. The composition of embodiment 69, wherein:

(a) position X7 of [N4] is Q or R;

(b) position X8 of [N4] is N or R;

(c) position X9 of [N4] is Q or R; and

(d) position X10 of [N4] is Q, L, or R.

71. The composition of embodiment 69 or 70, wherein [N4] is or comprises:

(i) QNQQ (SEQ ID NO: 5028), WNQQ (SEQ ID NO: 5029), QYYV (SEQ ID NO: 5030), RRQQ (SEQ ID NO: 5031), GCGQ (SEQ ID NO: 5032), LRQQ (SEQ ID NO: 5033), RNQQ (SEQ ID NO: 5034), VNQQ (SEQ ID NO: 5035), FRLQ (SEQ ID NO: 5036), FNQQ (SEQ ID NO: 5037), LLQQ (SEQ ID NO: 5038), SNQQ (SEQ ID NO: 5039), RLQQ (SEQ ID NO: 5040), LNQQ (SEQ ID NO: 5041), QRKL (SEQ ID NO: 5042), LRRQ (SEQ ID NO: 5043), QRLR (SEQ ID NO: 5044), QRRL (SEQ ID NO: 5045), RRLQ (SEQ ID NO: 5046), RLRQ (SEQ ID NO: 5047), SKRQ (SEQ ID NO: 5048), QLYR (SEQ ID NO: 5049), QLTV (SEQ ID NO: 5050), QNKQ (SEQ ID NO: 5051), KNQQ (SEQ ID NO: 5052), QKQQ (SEQ ID NO: 5053), QTQQ (SEQ ID NO: 5054), QNHQ (SEQ

ID NO: 5055), QHQQ (SEQ ID NO: 5056), QNQH (SEQ ID NO: 5057), QHRQ (SEQ ID NO: 5058), LTQQ (SEQ ID NO: 5059), QNQW (SEQ ID NO: 5060), QNTH (SEQ ID NO: 5061), RRRQ (SEQ ID NO: 5062), QYQQ (SEQ ID NO: 5063), QNDQ (SEQ ID NO: 5064), QNRH (SEQ ID NO: 5065), RDQQ (SEQ ID NO: 5066), PNLQ (SEQ ID NO: 5067), HVRQ (SEQ ID NO: 5068), PNQH (SEQ ID NO: 5069), HNQQ (SEQ ID NO: 5070), QSQQ (SEQ ID NO: 5071), QPAK (SEQ ID NO: 5072), QNLA (SEQ ID NO: 5073), QNQL (SEQ ID NO: 5074), QGQQ (SEQ ID NO: 5075), LNRQ (SEQ ID NO: 5076), QNPP (SEQ ID NO: 5077), QNLQ (SEQ ID NO: 5078), QDQE (SEQ ID NO: 5079), QDQQ (SEQ ID NO: 5080), HWQQ (SEQ ID NO: 5081), PNQQ (SEQ ID NO: 5082), PEQQ (SEQ ID NO: 5083), QRTM (SEQ ID NO: 5084), LHQH (SEQ ID NO: 5085), QHRI (SEQ ID NO: 5086), QYIH (SEQ ID NO: 5087), QKFE (SEQ ID NO: 5088), QFPS (SEQ ID NO: 5089), QNPL (SEQ ID NO: 5090), QAIK (SEQ ID NO: 5091), QNRQ (SEQ ID NO: 5092), QYQH (SEQ ID NO: 5093), QNPQ (SEQ ID NO: 5094), QHQL (SEQ ID NO: 5095), QSPP (SEQ ID NO: 5096), QAKL (SEQ ID NO: 5097), KSQQ (SEQ ID NO: 5098), QDRP (SEQ ID NO: 5099), QNLG (SEQ ID NO: 5100), QAFH (SEQ ID NO: 5101), QNAQ (SEQ ID NO: 5102), HNQL (SEQ ID NO: 5103), QKLN (SEQ ID NO: 5104), QNVQ (SEQ ID NO: 5105), QAQQ (SEQ ID NO: 5106), QTPP (SEQ ID NO: 5107), QPPA (SEQ ID NO: 5108), QERP (SEQ ID NO: 5109), QDLQ (SEQ ID NO: 5110), QAMH (SEQ ID NO: 5111), QHPS (SEQ ID NO: 5112), PGLQ (SEQ ID NO: 5113), QGIR (SEQ ID NO: 5114), QAPA (SEQ ID NO: 5115), QIPP (SEQ ID NO: 5116), QTQL (SEQ ID NO: 5117), QAPS (SEQ ID NO: 5118), QNTY (SEQ ID NO: 5119), QDKQ (SEQ ID NO: 5120), QNHL (SEQ ID NO: 5121), QIGM (SEQ ID NO: 5122), LNKQ (SEQ ID NO: 5123), PNQL (SEQ ID NO: 5124), QLQQ (SEQ ID NO: 5125), QRMS (SEQ ID NO: 5126), QGIL (SEQ ID NO: 5127), QDRQ (SEQ ID NO: 5128), RDWQ (SEQ ID NO: 5129), QERS (SEQ ID NO: 5130), QNYQ (SEQ ID NO: 5131), QRTC (SEQ ID NO: 5132), QIGH (SEQ ID NO: 5133), QGAI (SEQ ID NO: 5134), QVPP (SEQ ID NO: 5135), QVQQ (SEQ ID NO: 5136), LMRQ (SEQ ID NO: 5137), QYSV (SEQ ID NO: 5138), QAIT (SEQ ID NO: 5139), QKTL (SEQ ID NO: 5140), QLHH (SEQ ID NO: 5141), QNII (SEQ ID NO: 5142), QGHH (SEQ ID NO: 5143), QSKV (SEQ ID NO: 5144), QLPS (SEQ ID NO: 5145), IGKQ (SEQ ID NO: 5146), QAIH (SEQ ID NO: 5147), QHGL (SEQ ID NO: 5148), QFMC (SEQ ID NO: 5149), QNQM (SEQ ID NO: 5150), QHLQ (SEQ ID NO: 5151), QPAR (SEQ ID NO: 5152), QSLQ (SEQ ID NO: 5153), QSQL (SEQ ID NO: 5154), HSQQ (SEQ ID NO: 5155), QMPS (SEQ ID NO: 5156), QGSL (SEQ ID NO: 5157), QVPA (SEQ ID NO: 5158), HYQQ (SEQ ID NO: 5159), QVPS (SEQ ID NO: 5160), RGEQ (SEQ ID NO: 5161), PGQQ (SEQ ID NO: 5162), LEQQ (SEQ ID NO: 5163), QNQS (SEQ ID NO: 5164), QKVI (SEQ ID NO: 5165), QNND (SEQ ID NO: 5166), QSVH (SEQ ID NO: 5167), QPLG (SEQ ID NO: 5168), HNQE (SEQ ID NO: 5169), QIQQ (SEQ ID NO: 5170), QVRN (SEQ ID NO: 5171), PSNQ (SEQ ID NO: 5172), QVGH (SEQ ID NO: 5173), QRDI (SEQ ID NO: 5174), QMPN (SEQ ID NO: 5175), RGLQ (SEQ ID NO: 5176), PSLQ (SEQ ID NO: 5177), QRDQ (SEQ ID NO: 5178), QAKG (SEQ ID NO: 5179), QSAH (SEQ ID NO: 5180), QSTM (SEQ ID NO: 5181), QREM (SEQ ID NO: 5182), QYRA (SEQ ID NO: 5183), QRQQ (SEQ ID NO: 5184),

QWQQ (SEQ ID NO: 5185), QRMN (SEQ ID NO: 5186), GDSQ (SEQ ID NO: 5187), QKIS (SEQ ID NO: 5188), PSMQ (SEQ ID NO: 5189), SPRQ (SEQ ID NO: 5190), MEQQ (SEQ ID NO: 5191), QYQN (SEQ ID NO: 5192), QIRQ (SEQ ID NO: 5193), QSVQ (SEQ ID NO: 5194), RSQQ (SEQ ID NO: 5195), QNKL (SEQ ID NO: 5196), QIQH (SEQ ID NO: 5197), PRQQ (SEQ ID NO: 5198), HTQQ (SEQ ID NO: 5199), QRQH (SEQ ID NO: 5200), RNQE (SEQ ID NO: 5201), QSKQ (SEQ ID NO: 5202), QNQP (SEQ ID NO: 5203), QSPQ (SEQ ID NO: 5204), QTRQ (SEQ ID NO: 5205), QNLH (SEQ ID NO: 5206), QNQE (SEQ ID NO: 5207), LNQP (SEQ ID NO: 5208), QNQD (SEQ ID NO: 5209), QNLL (SEQ ID NO: 5210), QLVI (SEQ ID NO: 5211), RTQE (SEQ ID NO: 5212), QTHQ (SEQ ID NO: 5213), QDQH (SEQ ID NO: 5214), QSQH (SEQ ID NO: 5215), VRQQ (SEQ ID NO: 5216), AWQQ (SEQ ID NO: 5217), QSVP (SEQ ID NO: 5218), QNIQ (SEQ ID NO: 5219), LDQQ (SEQ ID NO: 5220), PDQQ (SEQ ID NO: 5221), ESQQ (SEQ ID NO: 5222), QRQL (SEQ ID NO: 5223), QIIV (SEQ ID NO: 5224), QKQS (SEQ ID NO: 5225), QSHQ (SEQ ID NO: 5226), QFVV (SEQ ID NO: 5227), QSQP (SEQ ID NO: 5228), QNEQ (SEQ ID NO: 5229), INQQ (SEQ ID NO: 5230), RNRQ (SEQ ID NO: 5231), RDQK (SEQ ID NO: 5232), QWKR (SEQ ID NO: 5233), ENRQ (SEQ ID NO: 5234), QTQP (SEQ ID NO: 5235), QKQL (SEQ ID NO: 5236), RNQL (SEQ ID NO: 5237), ISIQ (SEQ ID NO: 5238), QTVC (SEQ ID NO: 5239), QQIM (SEQ ID NO: 5240), LNHQ (SEQ ID NO: 5241), QNQA (SEQ ID NO: 5242), QMIH (SEQ ID NO: 5243), RNHQ (SEQ ID NO: 5244), or QKMN (SEQ ID NO: 5245);

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, or 3 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

72. The composition of any one of embodiments 69-71, wherein [N1]-[N2]-[N3]-[N4] is or comprises:

(i) the amino acid sequence of any of SEQ ID NOs: 1800-2241;

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

73. The composition of any one of embodiments 69-72, wherein [N1]-[N2]-[N3]-[N4] is or comprises GSGSPHSKAQNQQ (SEQ ID NO: 1801).

74. The composition of any one of embodiments 69-72, wherein [N1]-[N2]-[N3]-[N4] is or comprises GHDSPHKSGQNQQ (SEQ ID NO: 1800).

75. The composition of any one of embodiments 69-72, wherein [N1]-[N2]-[N3]-[N4] is or comprises GSGSPHKYGQNQQT (SEQ ID NO: 910).

76. The composition of any one of embodiments 38-75, which further comprises [N0], wherein [N0] comprises XA XB and XC, and wherein:

(a) position XA is independently chosen from T, S, Y, M, A, C, I, R, L, D, F, V, Q, N, H, E, or G;

(b) position XB is independently chosen from I, M, P, E, N, D, S, A, T, G, Q, F, V, L, C, H, R, W, or L; and

(c) position XC is independently chosen from N, M, E, G, Y, W, T, I, Q, F, V, A, L, I, P, K, R, H, S, D, or S; and

optionally wherein the protein or peptide comprises an amino acid modification, e.g., a conservative substitution, of any of the aforesaid amino acids in (a)-(c).

77. The composition of embodiment 76, wherein [N0] is or comprises TIN, SMN, TIM, YLS, GLS, MPE, MEG, MEY, AEW, CEW, ANN, IPE, ADM, IEY, ADY, IET, MEW, CEY, RIN, MEI, LEY, ADW, IEI, DIM, FEQ, MEF, CDQ, LPE, IEN, MES, AEI, VEY, IIN, TSN, IEV, MEM, AEV, MDA, VEW, AEQ, LEW, MEL, MET, MEA, IES, MEV, CEI, ATN, MDG, QEV, ADQ, NMN, IEM, ISN, TGN, QQQ, HDW, IEG, TII, TFP, TEK, EIN, TVN, TFN, SIN, TER, TSY, ELH, AIN, SVN, TDN, TFH, TVH, TEN, TSS, TID, TCN, NIN, TEH, AEM, AIK, TDK, TFK, SDQ, TEI, NTN, TET, SIK, TEL, TEA, TAN, TIY, TFS, TES, TTN, TED, TNN, EVH, TIS, TVR, TDR, TIK, NHI, TIP, ESD, TDL, TVP, TVI, AEH, NCL, TVK, NAD, TIT, NCV, TIR, NAL, VIN, TIQ, TEF, TRE, QGE, SEK, NVN, GGE, EFV, SDK, TEQ, EVQ, TEY, NCW, TDV, SDI, NSI, NSL, EVV, TEP, SEL, TWQ, TEV, AVN, GVL, TLN, TEG, TRD, NAI, AEN, AET, ETA, NNL, or any dipeptide thereof.

78. The composition of embodiment 76 or 77, wherein [N0]-[N1]-[N2]-[N3]-[N4] is or comprises:

(i) the amino acid sequence of any one of SEQ ID NOs: 2242-2886;

(ii) an amino acid sequence comprising any portion of an amino acid sequence in (i), e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids, e.g., consecutive amino acids, thereof;

(iii) an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (i); or

(iv) an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the amino acid sequences in (i).

79. The composition of any one of embodiments 76-78, wherein [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGSGSPHSKAQNQQ (SEQ ID NO: 2242).

80. The composition of any one of embodiments 76-78, wherein [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGHDSPHKSGQNQQ (SEQ ID NO: 2243).

81. The composition of any one of embodiments 76-78, wherein [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGSGSPHKYGQNQQT (SEQ ID NO: 5246).

82. The composition of any one of embodiments 38-81, wherein [N3] is present immediately subsequent to [N2].

83. The composition of any one of embodiments 38-82, which comprises from N-terminus to C-terminus, [N2]-[N3].

84. The composition of any one of embodiments 38-83, which comprises from N-terminus to C-terminus, [N1]-[N2]-[N3].

85. The composition of any one of embodiments 76-84, which comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3].

86. The composition of any one of embodiments 69-85, which comprises from N-terminus to C-terminus, [N1]-[N2]-[N3]-[N4].

87. The composition of any one of embodiments 76-86, which comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]-[N4].

88. The composition of any one of embodiments 1-87, wherein the ligand comprises at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides according to any one of embodiments 35-84.

89. The composition of embodiment 88, wherein the at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides comprise the same amino acid sequence.

90. The composition of embodiment 88, wherein the at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides comprise different amino acid sequences.

91. The composition of any one of embodiments 88-90, wherein the at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides are present in tandem (e.g., connected directly or indirectly via a linker) or in a multimeric configuration.

92. The composition of any one of embodiments 38-91, wherein the protein or peptide comprises an amino acid sequence of at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 20, 25, 30, or 35 amino acids in length.

93. The composition of embodiment of 92, wherein the protein or peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or all of the amino acids TLKFSVAGPSNMAVQG (SEQ ID NO: 4694), optionally wherein the at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the amino acids LKFSVAGPSNMAVQG (SEQ ID NO: 21) is present C-terminal relative to [N4].

94. The composition of any one of embodiments 5-93, wherein the peptide comprises the amino acid sequence of SPH, wherein the S comprises a modification, e.g., comprises a phosphate group.

95. The composition of any one of embodiments 5-94, wherein the peptide comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), optionally wherein the S at position 1, numbered according to SEQ ID NO: 941, comprises a modification, e.g., comprises a phosphate group.

96. The composition of any one of embodiments 2-94, wherein the peptide comprises the amino acid sequence of SPHK (SEQ ID NO: 6398), optionally wherein S comprises a modification, e.g., comprises a phosphate group.

97. The composition of any one of embodiments 2-94 or 96, wherein the peptide comprises the amino acid sequence of HDSPHK (SEQ ID NO: 2), optionally wherein S comprises a modification, e.g., comprises a phosphate group.

98. The composition of any one of embodiments 38-97, wherein the modification comprises a phosphate group.

99. The composition of any one of embodiments 38-98, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids present N-terminal relative the amino acid sequence of SPH.

100. The composition of any one of embodiments 96-99, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids present N-terminal relative the amino acid sequence of HDSPHK (SEQ ID NO: 2).

101. The composition of any one of embodiments 96-100, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids present C-terminal relative the amino acid sequence of HDSPHK (SEQ ID NO: 2).

102. The composition of any one of embodiments 94, 95, or 98, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids present N-terminal relative the amino acid sequence of SPHSKA (SEQ ID NO: 941).

103. The composition of any one of embodiments 94, 95, 98, or 102, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids present C-terminal relative the amino acid sequence of SPHSKA (SEQ ID NO: 941).

104. The composition of any one of embodiments 76-94, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 amino acids present N-terminal relative the amino acid sequence of [N0]-[N2]-[N3]-[N4].

105. The composition of any one of embodiments 76-94 or 104, wherein the peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 amino acids present C-terminal relative the amino acid sequence of [N0]-[N2]-[N3]-[N4].

106. The composition of any one of embodiments 5-105, wherein the peptide comprises the amino acid sequence of:

(i) GHDSPHKS (SEQ ID NO: 4487), optionally wherein the S at position 4 of SEQ ID NO: 4487, comprises a modification, e.g., comprises a phosphate group;

(ii) NGHDSPHKSG (SEQ ID NO: 4489), optionally wherein the S at position 5 of SEQ ID NO: 4489, comprises a modification, e.g., comprises a phosphate group;

(iii) INGHDSPHKSGQ (SEQ ID NO: 4490), optionally wherein the S at position 6 of SEQ ID NO: 4490, comprises a modification, e.g., comprises a phosphate group;

(iv) TINGHDSPHKSGQN (SEQ ID NO: 4491), optionally wherein the S at position 7 of SEQ ID NO: 4491, comprises a modification, e.g., comprises a phosphate group;

(v) KTINGHDSPHKSGQNQ (SEQ ID NO: 4492), optionally wherein the S at position 8 of SEQ ID NO: 4492, comprises a modification, e.g., comprises a phosphate group;

(vi) LYYLSKTINGHDSPHKSGQNQQTLKF (SEQ ID NO: 4518), optionally wherein the S at position 13 of SEQ ID NO: 4518, comprises a modification, e.g., comprises a phosphate group;

(vii) RLMNPLIDQYLYYLSKTINGHDSPHKSGQNQQTLKFSVAGPSNMAV (SEQ ID NO: 4519), optionally wherein the S at position 23 of SEQ ID NO: 4519, comprises a modification, e.g., comprises a phosphate group;

(viii) GSPHSKAQ (SEQ ID NO: 4493), optionally wherein the S at position 2 of SEQ ID NO: 4493, comprises a modification, e.g., comprises a phosphate group;

(ix) SGSPHSKAQN (SEQ ID NO: 4494), optionally wherein the S at position 3 of SEQ ID NO: 4494, comprises a modification, e.g., comprises a phosphate group;

(x) GSGSPHSKAQNQ (SEQ ID NO: 4495), optionally wherein the S at position 4 of SEQ ID NO: 4495, comprises a modification, e.g., comprises a phosphate group;

(xi) NGSGSPHSKAQNQQ (SEQ ID NO: 4496), optionally wherein the S at position 5 of SEQ ID NO: 4496, comprises a modification, e.g., comprises a phosphate group; or

(xii) INGSGSPHSKAQNQQT (SEQ ID NO: 4497), optionally wherein the S at position 6 of SEQ ID NO: 4497, comprises a modification, e.g., comprises a phosphate group.

107. The composition of any one of embodiments 5-94 or 96-106, wherein the peptide comprises the amino acid sequence of NGHDpSPHKSG (SEQ ID NO: 4515).

108. The composition of any one of embodiments 5-94 or 96-107, wherein the peptide comprises the amino acid sequence of KTINGHDpSPHKSGQNQ (SEQ ID NO: 4516).

109. The composition of any one of embodiments 5-94 or 96-107, wherein the peptide comprises the amino acid sequence of YLSKTINGHDpSPHKSGQNQQTLKFS (SEQ ID NO: 4517).

110. The composition of embodiment 1-109, wherein the ligand is a conjugate comprising at least 2-5, e.g., at least 2, 3, 4, or 5, proteins or peptides according to any one of embodiments 38-109, wherein the conjugate comprises a chemical linkage, e.g., succinimidyl ester or biotin.

111. The composition of any one of embodiments 1-110, wherein the ligand is a fusion protein comprising at least 2-5, e.g., at least 2, 3, 4, or 5, proteins or peptides according to any one of embodiments 3-108, wherein each protein or peptide of the fusion protein is connected directly or via a linker.

112. The composition of any one of embodiments 1-111, wherein the peptide or protein is identified using phage-display.

113. The composition of any one of embodiments 1-37, wherein the ligand is or comprises an aptamer.

114. The composition of embodiment 113, wherein the aptamer binds to human, murine, or NHP ALPL.

115. The composition of embodiment 113 or 114, wherein the aptamer is or comprises DNA, RNA, modified DNA, modified RNA, or a combination thereof.

116. The composition of any one of embodiment 114-115, wherein the aptamer is fused or coupled to a therapeutic agent chosen from a protein (e.g., an enzyme), an antibody molecule, a nucleic acid molecule (e.g., an RNAi agent), or a small molecule.

117. The composition of any one of embodiments 1-37, wherein the ligand is or comprises an antibody molecule that binds to the GPI anchored protein, e.g., ALPL.

118. The composition of embodiment 117, wherein the antibody molecule comprises a full antibody or an antigen binding fragment.

119. The composition of embodiment 117 or 118, wherein the antigen binding fragment is a Fab or a Fab fragment, a F(ab)2 fragment, an Fv fragment, dAb fragment, a single chain antibody (scFv) or a scFv fragment, an antibody variable region, a diabody, a VHH, a camelid antibody, a single domain antibody or a nanobody.

120. The composition of any one of embodiments 117-119, wherein the antibody molecule is a monospecific antibody, a multispecific antibody, e.g., a bispecific or biparatopic antibody.

121. The composition of any one of embodiments 117-120, wherein the antibody molecule is a human antibody, a humanized antibody, a chimeric antibody, a phage display antibody, a recombinant antibody, a murine antibody.

122. The composition of any one of embodiments 117-121, wherein the antibody molecule comprises a half-life extender.

123. The composition of any one of embodiments 117-122, wherein the variable domain of the antibody molecule binds to ALPL, e.g., human ALPL.

124. The composition of any one of embodiments 117-123, wherein the antibody molecule is an antibody as provided in Table 40 (e.g., Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, 2F4, or a variant thereof.

125. The composition of any of any one of embodiments 117-124, wherein the antibody molecule binds the same or substantially the same epitope as any one of an antibody as provided in Table 40 (e.g., Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, or 2F4.

126. The composition of any of any one of embodiments 117-125, wherein the antibody molecule competes for binding with any one of an antibody as provided in Table 40 (e.g., Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, or 2F4.

127. The composition of any one of 117-126, which further comprises a therapeutic antibody molecule, e.g., a multispecific antibody comprising a first binding domain that binds to ALPL (e.g., an anti-ALPL binding domain) and a second binding domain that binds to a therapeutic target.

128. A multispecific antibody molecule comprising a first binding domain that binds to ALPL (e.g., an anti-ALPL binding domain) and a second binding domain that binds to a therapeutic target.

129. The multispecific antibody molecule of embodiment 128, wherein the first and/or second binding domain is a full-length antibody, or an antigen-binding fragment (e.g., a Fab, F(ab')2, Fv, a single chain Fv (scFv), a single domain antibody, a half-arm antibody, a diabody (dAb), a bivalent antibody, a bispecific antibody or fragment thereof, a single domain variant thereof, or a camelid antibody).

130. The multispecific antibody molecule of embodiment 128 or 129, wherein:

(i) the anti-ALPL binding domain is an Fab and the second binding domain is an scFv;

(ii) the anti-ALPL binding domain is an Fab and the second binding domain is an Fab;

(iii) the anti-ALPL binding domain is an scFv and the second binding domain is an scFv; or

(iv) the anti-ALPL binding domain is an scFv and the second binding domain is an Fab.

131. The multispecific antibody molecule of any one of embodiments 128-130, wherein the multispecific antibody molecule comprises an immunoglobulin constant region (e.g., an Fc region).

132. The multispecific antibody molecule of embodiment 131, wherein the immunoglobulin constant region (e.g., an Fc region) is linked (e.g., covalently linked) to the first and/or the second binding domain.

133. The multispecific antibody molecule of any one of embodiments 128-132, wherein the first and/or second binding domain comprises a light chain constant region chosen from the light chain constant region of kappa or lambda, or a fragment thereof.

134. The multispecific antibody molecule of any one of embodiments 128-133, wherein the first binding domain and the second binding domain comprise a common light chain variable region.

135. The multispecific antibody molecule of any one of embodiments 128-134, comprising a dimerization domain, e.g., an interface of a first and second immunoglobulin chain constant regions (e.g., Fc regions).

136. The multispecific antibody molecule of embodiment 135, wherein the dimerization domain is engineered, e.g., mutated, to increase or decrease dimerization, e.g., relative to a non-engineered interface.

137. The multispecific antibody molecule of embodiment 136, wherein the dimerization of the immunoglobulin chain constant regions (e.g., Fc regions) is enhanced by providing an Fc interface of a first and a second Fc regions with one or more of: a paired cavity-protuberance ("knob-in-a hole"), an electrostatic interaction, or a strand-exchange, such that a greater ratio of heteromultimer:homomultimer forms, e.g., relative to a non-engineered interface.

138. The multispecific antibody molecule of any one of embodiments 135-137, wherein the immunoglobulin chain constant region (e.g., Fc region) comprises an amino acid substitution at a position chosen from one or more of 347, 349, 350, 351, 366, 368, 370, 392, 394, 395, 397, 398, 399, 405, 407, or 409, e.g., of the Fc region of human IgG1.

139. The multispecific antibody molecule of any one of embodiments 135-138, wherein the immunoglobulin chain constant region (e.g., Fc region) comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole), or T366W (e.g., corresponding to a protuberance or knob), or a combination thereof.

140. The multispecific antibody molecule of any one of embodiments 128-139, wherein: the anti-ALPL binding domain comprises a first polypeptide and a second polypeptide, and the second binding domain comprises a third polypeptide and a fourth polypeptide, wherein:

(i) the first polypeptide comprises, e.g., from N-terminal to C-terminal: a first heavy chain variable region (VH), a first heavy chain constant region 1 (CH1), and a first Fc region that promotes association between the first and third polypeptides, wherein the first Fc region comprises a first heavy chain constant region 2 (CH2) and a first heavy chain constant region 3 (CH3);

(ii) the second polypeptide comprises, e.g., from N-terminal to C-terminal: a first light chain variable region (VL) and a first light chain constant region (CL);

(iii) the third polypeptide comprises, e.g., from N-terminal to C-terminal: a second heavy chain variable region (VH), a second heavy chain constant region 1 (CH1), and a second Fc region that promotes association between the first and third polypeptides, wherein the second Fc region comprises a second heavy chain constant region 2 (CH2) and a second heavy chain constant region 3 (CH3); and

(iv) the fourth polypeptide comprises, e.g., from N-terminal to C-terminal: a second light chain variable region (VL) and a second light chain constant region (CL).

141. The multispecific antibody molecule of any one of embodiments 128-139, wherein:

(i) the anti-ALPL binding domain (e.g., an anti-ALPL Fab or scFv), is situated N-terminal relative to the second binding domain that binds to a therapeutic target (e.g., a Fab or scFv); or

(ii) the second binding domain that binds to a therapeutic target (e.g., a Fab or scFv) is situated N-terminal relative to the anti-ALPL binding domain (e.g., an anti-ALPL Fab or scFv), optionally wherein an Fc region is situated between the anti-ALPL binding domain binding domain and the second binding domain that binds to a therapeutic target.

142. The multispecific antibody molecule of any one of embodiments 128-141, wherein the Fc region of the first and/or second binding domain:

(i) has reduced affinity, e.g., ablated, affinity for an Fc receptor, e.g., as compared to a reference, wherein the reference is a wild-type Fc receptor;

(ii) comprises a mutation at one, two, or all of positions I253 (e.g., I253A), H310 (e.g., H310A or H310Q), and/or H435 (e.g., H435A or H435Q), numbered according to the EU index as in Kabat;

(iii) has reduced effector function (e.g., reduced ADCC), compared to a reference wherein the reference is a wild-type Fc receptor;

(iv) comprises a mutation at one, two, three, four, or all of positions L235 (e.g., L235V), F243 (e.g., F243L), R292 (e.g., R292P), Y300 (e.g., Y300L), and P396 (e.g., P396L), numbered according to the EU index as in Kabat.

143. The multispecific antibody molecule of any one of embodiments 128-142, wherein the therapeutic target comprises:

(i) a CNS related target, e.g., an antigen associated with a neurological or neurodegenerative disorder, e.g., β-amyloid, APOE, tau, SOD1, TDP-43, huntingtin (HTT), and/or synuclein;

(ii) a muscular or neuromuscular related target, e.g., an antigen associated with a muscular or neuromuscular disorder; or

(iii) a neuro-oncology related target, e.g., an antigen associated with a neuro-oncological disorder, e.g., HER2, or EGFR (e.g., EGFRvIII).

144. The composition of any one of embodiments 1-37, wherein ligand is or comprises a first Fc polypeptide.

145. The composition of embodiment 144, wherein the first Fc polypeptide is fused or coupled to an active agent comprising a second Fc polypeptide.

146. The composition of embodiment 145, wherein the first Fc polypeptide and the second Fc polypeptide form a dimer.

147. The composition of embodiment 145 or 146, wherein the second Fc polypeptide is fused or coupled (e.g., directly or indirectly via a linker) to a therapeutic protein or variant thereof (e.g., an enzyme).

148. The composition of any one of embodiments 145-147, wherein the second Fc polypeptide is covalently linked to the therapeutic protein or variant thereof.

149. The composition of any one of embodiments 145-148, wherein the second Fc polypeptide is connected to the therapeutic protein or variant thereof indirectly via a linker.

150. The composition of embodiment 149, wherein the linker is a peptide linker (e.g., a flexible peptide linker (e.g., a glycine-serine linker) or a peptide linker sensitive to a protease), a cleavable linker (e.g., a pH sensitive linker or an enzyme sensitive linker), or a non-cleavable linker (e.g., a linker comprising a thioether group or a maleimidocaproyl group).

151. The composition of embodiment 149 or 150, wherein the linker is a glycine-serine linker, e.g., a G4S linker (SEQ ID NO: 6407) or a (G4S)2 linker (SEQ ID NO: 6408).

152. The composition of any one of embodiments 147-151, wherein the therapeutic protein is present at the N-terminus of the second Fc polypeptide.

153. The composition of any one of embodiments 147-151, wherein the therapeutic protein is present at the C-terminus of the second Fc polypeptide.

154. The composition of any one of embodiments 147-153, wherein the therapeutic protein or functional variant thereof is associated with (e.g., aberrantly expressed in) a neurological or neurodegenerative disorder, a muscular or neuromuscular disorder, or a neuro-oncological disorder.

155. The composition of any one of embodiments 147-154, wherein the therapeutic protein or functional variant thereof is chosen from apolipoprotein E (APOE) (e.g., ApoE2, ApoE3 and/or ApoE4); human survival of motor neuron (SMN) 1 or SMN2; glucocerebrosidase (GBA1); aromatic L-amino acid decarboxylase (AADC); aspartoacylase (ASPA); tripeptidyl peptidase I (CLN2); beta-galactosidase (GLB1); N-sulphoglucosamine sulphohydrolase (SGSH); N-acetyl-alpha-glucosaminidase (NAGLU); iduronate 2-sulfatase (IDS); intracellular cholesterol transporter (NPC1); or gigaxonin (GAN).

156. The composition of any one of embodiments 144-155, wherein the first Fc polypeptide is fused or coupled to a second therapeutic protein or variant thereof, e.g., an enzyme, optionally wherein the therapeutic protein or variant thereof is fused or coupled to the N-terminus or the C-terminus of the first Fc polypeptide.

157. The composition of any one of embodiments 145-156, wherein the first Fc polypeptide and the second Fc polypeptide comprise a dimerization domain, e.g., an interface of a first and second Fc polypeptides.

158. The composition of embodiment 157, wherein the dimerization domain is engineered, e.g., mutated, to increase or decrease dimerization, e.g., relative to a non-engineered interface.

159. The composition of embodiment 158, wherein the dimerization of the first Fc polypeptide and the second Fc polypeptide is enhanced by providing an Fc interface of the first and a second Fc polypeptides with one or more of: a paired cavity-protuberance ("knob-in-a hole"), an electrostatic interaction, or a strand-exchange, such that a greater ratio of heteromultimer:homomultimer forms, e.g., relative to a non-engineered interface.

160. The composition of any one of embodiments 145-159, wherein the first Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof).

161. The composition of any one of embodiments 145-160, wherein the second Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob).

162. The composition of any one of embodiment 145-161, wherein the first Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof); and the second Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob).

163. The composition of any one of embodiments 145-162, wherein the second Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof).

164. The composition of any one of embodiments 145-159 or 163, wherein the first Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob).

165. The composition of any one of embodiment 145-159, 163, or 164, wherein the second Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof); and the first Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob).

166. The composition of any one of embodiments 145-165, wherein the first Fc polypeptide, the second Fc polypeptide, or both:

(i) has reduced affinity, e.g., ablated, affinity for an Fc receptor, e.g., as compared to a reference, wherein the reference is a wild-type Fc receptor;

(ii) comprises a mutation at one, two, or all of positions I253 (e.g., I253A), H310 (e.g., H310A or H310Q), and/or H435 (e.g., H435A or H435Q), numbered according to the EU index as in Kabat;

(iii) has reduced effector function (e.g., reduced ADCC), compared to a reference wherein the reference is a wild-type Fc receptor;

(iv) comprises a mutation at one, two, three, four, or all of positions L235 (e.g., L235V), F243 (e.g., F243L), R292 (e.g., R292P), Y300 (e.g., Y300L), and P396 (e.g., P396L), numbered according to the EU index as in Kabat.

167. The composition of any one of embodiments 145-166, wherein the first Fc polypeptide, the second Fc polypeptide, or both comprises a half-life extender or an amino acid modification that increases serum half-life (e.g., (i) a Leu at position 428 and a Ser at position 434, or (ii) a Ser or Ala at position 434, according to EU numbering).

168. The composition of any one of embodiments 144-167, wherein the first Fc polypeptide comprises the protein or peptide according to any one of embodiments 35-84.

169. The composition of any one of embodiments 168, wherein the protein or peptide is present in the CH3 domain of the first Fc polypeptide.

170. The composition of embodiment 169, wherein the CH3 domain is modified from a human IgGl, IgG2, IgG3, or IgG4 CH3 domain.

171. The composition of embodiment 169 or 170, wherein the CH3 domain comprises one, two, three, four, five, six, seven, eight, nine, ten, or eleven substitutions in a set of amino acid positions comprising 380, 384, 386, 387, 388, 389, 390, 413, 415, 416, and 421, according to EU numbering.

172. The composition of any one of embodiments 168-171, wherein the protein or peptide is present at or near the C-terminus of the first Fc polypeptide (e.g., within 20, 30, 40, 50, 60, 70, 80, 90, 100, or more amino acids from the C-terminus of the therapeutic protein, enzyme, or antibody molecule).

173. The composition of any one of embodiments 145-172, wherein the first Fc polypeptide, the second Fc polypeptide or both the first Fc polypeptide and the second Fc polypeptide does not comprise an immunoglobulin heavy and/or light chain variable region sequence or an antigen-binding portion thereof.

174. The composition of embodiment 1-11 or 15-37, wherein the ligand is a component of a viral particle, e.g., an AAV particle or a lentivirus.

175. The composition of any one of embodiments 1-11, 15-37, or 174, wherein the ligand is a component of a capsid protein, e.g., an AAV capsid protein.

176. The composition of any one of embodiments 1-11, 15-37, 174, or 175, wherein the ligand is a component of an AAV9 capsid or variant thereof.

177. The composition of any one of embodiments 1-11, 15-37, or 174-176, wherein the ligand is an AAV9 capsid variant comprising a modification, e.g., a substitution, insertion, and/or deletion, in loop IV of AAV9.

178. The composition of any one of embodiments 1-11, 15-37, or 174-176, wherein ligand is an AAV9 capsid variant comprising the amino acid sequence of any one of embodiments 35-84.

179. The composition of any one of embodiments 1-11, 15-37, or 174, wherein the ligand is a lentiviral particle, wherein at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80% of the surface of the lentiviral particle comprises at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides, e.g., ALPL-binding peptides, or a peptide or protein according to any one of embodiments 38-109.

180. The composition of embodiment 1-37, wherein the ligand is a small molecule.

181. The composition of embodiment 180, wherein the small molecule is an inhibitor of ALPL, e.g., a small molecule that interferes with ALPL dimerization.

182. The composition of embodiment 180 or 181, wherein the small molecule is an aryl sulfonamide, a phosphonate derivative, a pyrazole, a triazole, or an imidazole, optionally wherein the small molecule is 2,5-Dimethoxy-N-(quinolin-3-yl)benzenesulfonamide (Tissue-Nonspecific Alkaline Phosphatase Inhibitor (TNAPi)) or 5-((5-chloro-2-methoxyphenyl)sulfonamido)nicotinamide (SBI-425).

183. The composition of any one of the preceding embodiments, wherein binding to ALPL results in increased cellular transduction, e.g., as compared to a reference sequence of SEQ ID NO: 138, e.g., when measured by a transduction assay or binding/internalization assay as described (e.g., as described in Example 8).

184. The composition of any one of the preceding embodiments, wherein binding to ALPL results in increased crossing of the blood brain barrier, e.g., as compared to a reference sequence of SEQ ID NO: 138, e.g., when measured by a transduction assay or binding/internalization assay as described (e.g., as described in Example 8).

185. The composition of any one of embodiments 1-127 or 144-184, wherein the therapeutic agent or diagnostic agent is an antibody molecule or an Fc polypeptide.

186. The composition of embodiment 185, wherein the antibody molecule comprises a full antibody or an antigen binding fragment.

187. The composition of embodiment 186, wherein the antigen binding fragment is a Fab or a Fab fragment, a F(ab)2 fragment, an Fv fragment, dAb fragment, a single chain antibody (scFv) or a scFv fragment, an antibody variable region, a diabody, a VHH, a camelid antibody, a single domain antibody or a nanobody.

188. The composition of any one of embodiments 185-187, wherein the antibody molecule is a monospecific antibody, a multispecific antibody, e.g., a bispecific or biparatopic antibody.

189. The composition of any one of embodiments 185-188, wherein the antibody molecule is a human antibody, a humanized antibody, a chimeric antibody, a phage display antibody, a recombinant antibody, a murine antibody.

190. The composition of any one of embodiments 185-189, wherein the antibody molecule is an antibody-drug conjugate.

191. The composition of embodiment 190, wherein the antibody molecule is conjugated to a cytotoxic or cytostatic agent, e.g., a chemotherapeutic agent or an anti-neoplastic drug.

192. The composition of any one of embodiments 185-189, wherein the antibody molecule is coupled to a radioactive isotope, e.g., α-, β-, or γ-emitter, or β-and γ-emitter.

193. The composition of any one of embodiments 185-192, wherein the antibody molecule comprises an Fc region, which comprises amino acid modifications that increase serum half-life.

194. The composition of embodiment 193, wherein the amino acid modifications that increase serum half-life comprise (i) a Leu at position 428 and a Ser at position 434, or (ii) a Ser or Ala at position 434, according to EU numbering.

195. The composition of embodiment 193 or 194, wherein the Fc region of the antibody molecule:

(i) has reduced affinity, e.g., ablated, affinity for an Fc receptor, e.g., as compared to a reference, wherein the reference is a wild-type Fc receptor;

(ii) comprises a mutation at one, two, or all of positions I253 (e.g., I253A), H310 (e.g., H310A or H310Q), and/or H435 (e.g., H435A or H435Q), numbered according to the EU index as in Kabat;

(iii) has reduced effector function (e.g., reduced ADCC), compared to a reference wherein the reference is a wild-type Fc receptor;

(iv) comprises a mutation at one, two, three, four, or all of positions L235 (e.g., L235V), F243 (e.g., F243L), R292 (e.g., R292P), Y300 (e.g., Y300L), and P396 (e.g., P396L), numbered according to the EU index as in Kabat.

196. The composition of any one of embodiments 185-195, wherein the antibody molecule binds:

(i) a CNS related target, e.g., an antigen associated with a neurological or neurodegenerative disorder, e.g., β-amyloid, APOE, tau, SOD1, TDP-43, huntingtin (HTT), and/or synuclein;

(ii) a muscular or neuromuscular related target, e.g., an antigen associated with a muscular or neuromuscular disorder; or

(iii) a neuro-oncology related target, e.g., an antigen associated with a neuro-oncological disorder, e.g., HER2, or EGFR (e.g., EGFRvIII).

197. The composition of any one of embodiments 1-127 or 144-118565, wherein the ligand is present or coupled to a carrier, e.g., an exosome, a microvesicle, or a lipid nanoparticle (LNP).

198. The composition of embodiment 197, wherein the carrier is an exosome or a LNP.

199. The composition of embodiment 197 or 198, wherein the ligand is present on the surface of the carrier.

200. The composition of any one of embodiments 197-199, wherein at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80% of the surface of the carrier comprises at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides according to any one of embodiments 35-84.

201. The composition of any one of embodiments 197-200, wherein the carrier comprises a therapeutic agent.

202. The composition of any one of embodiments 197-201, wherein the carrier comprises an RNAi agent, an mRNA, a ribonucleoprotein complex (e.g., a Cas9/gRNA complex), or a circRNA.

203. The composition of any one of embodiments 197-202, wherein the ligand is conjugated to the surface of the carrier by post-insertion.

204. The composition of any one of embodiments 197-202, wherein the ligand is conjugated to the surface of the carrier via a covalent bond (e.g., using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) chemistry or thiol-maleimide linkage reactions).

205. The composition of any one of embodiments 1-127 or 144-184, wherein the ligand is coupled to an RNAi agent directly or via a linker.

206. The composition of embodiment 205, wherein the RNAi agent is a dsRNA, a siRNA, a shRNA, a pre-miRNA, a pri-miRNA, a miRNA, a stRNA, a lncRNA, a piRNA, an antisense oligonucleotide agent (ASO), or a snoRNA.

207. The composition of embodiment 205 or 206, wherein the RNAi agent is a siRNA or an ASO.

208. The composition of embodiment 206 or 207, wherein the siRNA or the ASO comprises at least one modified nucleotide.

209. The composition of any one of embodiment 206-208, wherein not more than five of the sense strand nucleotides of the siRNA and not more than five of the nucleotides of the antisense strand of the siRNA are unmodified nucleotides.

210. The composition of any one of embodiment 206-209, wherein all of the nucleotides of the sense strand of the siRNA and all of the nucleotides of the antisense strand of the siRNA are modified.

211. The composition of any one of embodiment 206-208, wherein not more than five of the nucleotides of ASO are unmodified nucleotides.

212. The composition of any one of embodiment 206-208 or 211, wherein all of the nucleotides of the ASO are modified.

213. The composition of any one of embodiment 208-312, wherein the modified nucleotides is selected from the group consisting of a deoxy-nucleotide, a 3'-terminal deoxythimidine (dT) nucleotide, a 2'-O-methyl modified nucleotide, a 2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, a locked nucleotide, an unlocked nucleotide, a conformationally restricted nucleotide, a constrained ethyl nucleotide, an abasic nucleotide, a 2'-amino-modified nucleotide, a 2'-O-allyl-modified nucleotide, 2'-C-alkyl-modified nucleotide, a 2'-methoxyethyl modified nucleotide, a 2'-O-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, a non-natural base comprising nucleotide, a tetrahydropyran modified nucleotide, a 1,5-anhydrohexitol modified

nucleotide, a cyclohexenyl modified nucleotide, a nucleotide comprising a phosphorothioate group, a nucleotide comprising a methylphosphonate group, a nucleotide comprising a 5'-phosphate, a nucleotide comprising a 5'-phosphate mimic, a glycol modified nucleotide, and a 2-O-(N-methylacetamide) modified nucleotide; and combinations thereof.

214. The composition of any one of embodiments 205-213, wherein the RNAi agent modulates, e.g., inhibits, expression of, a CNS related gene, mRNA, and/or protein.

215. The composition of embodiment 214, wherein the CNS gene is chosen from SOD1, MAPT, APOE, HTT, C9ORF72, TDP-43, APP, BACE, SNCA, ATXN1, ATXN3, ATXN7, SCN1A-SCN5A, SCN8A-SCN11A, SMN, or a combination thereof.

216. The composition of any one of embodiments 205-215, wherein the ligand comprises the protein or peptide according to any one of embodiments 35-84.

217. The composition of embodiments 205-216, wherein the ligand comprises at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides according to any one of embodiments 35-84.

218. The composition of embodiment 216 or 217, wherein the at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides are present in tandem (e.g., connected directly or indirectly via a linker) or in a multimeric configuration.

219. The composition of any one of embodiments 216-218, wherein the protein or peptide comprises an amino acid sequence of at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 20, 25, 30, or 35 amino acids in length.

220. The composition of embodiment of 219, wherein the protein or peptide further comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the amino acids LKFSVAGPSNMAVQG (SEQ ID NO: 21).

221. The composition of any one of embodiments 205-221, wherein the ligand is covalently linked, e.g., directly or indirectly via a linker to the RNAi agent.

222. The composition of any one of embodiments 205-221, wherein the ligand is conjugated, e.g., directly or indirectly via a linker to the RNAi agent.

223. The composition of any one of embodiments 205-222, wherein the ligand is conjugated to the RNAi agent via a linker, e.g., a crosslinker.

224. The composition of embodiment 223, wherein the crosslinker comprises succinimidyl-4-(N-maleimidomethyl) and/or a saturated or unsaturated hydrocarbon chain (e.g., cyclohexane-1-carboxylate).

225. The composition of embodiment 223 or 224, wherein the crosslinker comprises succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate.

226. The composition of any one of embodiments 205-224, wherein the ligand is conjugated to the RNAi agent via a linker comprising an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate.

227. The composition of any one of embodiments 205-226, wherein the ligand is conjugated, e.g., directly or indirectly via a linker, to the N-terminus of at least one strand of the RNAi agent.

228. The composition of any one of embodiments 205-226, wherein the ligand is conjugated, e.g., directly or indirectly via a linker, to the C-terminus of at least one strand of the RNAi agent.

229. The composition of any one of embodiments 205-226, wherein the ligand is conjugated, e.g., directly or indirectly via a linker, to an internal nucleotide of at least one strand of the RNAi agent.

230. The composition of any one of embodiments 227-229, wherein the at least one strand of the RNAi agent is the sense strand.

231. The composition of any one of embodiments 205-230, wherein the composition further comprises a lipophilic moiety.

232. The composition of embodiment 231, wherein the lipophilic moiety is an aliphatic, alicyclic, or polyalicyclic compound.

233. The composition of embodiment 231 or 232, wherein the lipophilic moiety is selected from the group consisting of lipid, cholesterol, retinoic acid, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-bis-O(hexadecyl)glycerol, geranyloxyhexyanol, hexadecylglycerol, borneol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid, O3-(oleoyl)lithocholic acid, O3-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxazine.

234. The composition of any one of embodiments 231-233, wherein the lipophilic moiety contains a saturated or unsaturated C4-C30 hydrocarbon chain, and an optional functional group selected from the group consisting of hydroxyl, amine, carboxylic acid, sulfonate, phosphate, thiol, azide, and alkyne.

235. The composition of embodiment 234, wherein the lipophilic moiety contains a saturated or unsaturated C6-C18 hydrocarbon chain, e.g., a saturated or unsaturated C16 hydrocarbon chain.

236. The composition of any one of embodiments 231-235, wherein the lipophilic moiety is conjugated via a carrier that replaces one or more nucleotide(s) in the internal position(s) of the iRNA agent, e.g., the siRNA or ASO.

237. The composition of embodiment 236, wherein the carrier is a cyclic group selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl; or is an acyclic moiety based on a serinol backbone or a diethanolamine backbone.

238. The composition of any one of embodiments 231-237, wherein the lipophilic moiety is conjugated to the RNAi agent, e.g., the siRNA or ASO, via a linker containing an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate.

239. The composition of any one of embodiments 231-238, wherein the lipophilic moiety is conjugated to a nucleobase, sugar moiety, or internucleosidic linkage.

240. The composition of any one of embodiments 231-239, wherein the lipophilic moiety is conjugated via a bio-cleavable linker selected from the group consisting of DNA, RNA, disulfide, amide, functionalized monosaccharides or oligosaccharides of galactosamine, glucosamine, glucose, galactose, mannose, and combinations thereof.

241. The composition of any one of embodiments 231-240, wherein the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to the N-terminus of at least one strand of the RNAi agent.

242. The composition of any one of embodiments 231-241, wherein the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to the C-terminus of at least one strand of the RNAi agent.

243. The composition of any one of embodiments 231-242, wherein the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to an internal nucleotide of at least one strand of the RNAi agent.

244. The composition of any one of embodiments 241-243, wherein the at least one strand of the RNAi agent is the sense strand.

245. The composition of any one of embodiments 231-244, wherein the ligand and the lipophilic moiety are present on the same strand, e.g., the sense strand.

246. The composition of any one of embodiments 231-244, wherein the ligand and the lipophilic moiety are present on different strands.

247. The composition of any one of embodiments 206-246, wherein a 3' end of a sense strand of a siRNA agent is protected via an end cap which is a cyclic group having an amine, said cyclic group being selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl.

248. The composition of any one of embodiments 205-247, wherein the composition further comprises an N-acetylgalactosamine (GalNAc) conjugate.

249. The composition of embodiment 248, wherein the GalNAc conjugate is attached through a monovalent linker; or a bivalent, trivalent, or tetravalent branched linker.

250. The composition of any one of embodiments 1-127 or 144-184, wherein the active agent is a diagnostic agent.

251. The composition of embodiment 250, wherein the diagnostic agent is or comprises an imaging agent (e.g., a protein or small molecule compound coupled to a detectable moiety).

252. The composition of embodiment 251, wherein the imaging agent comprises a PET or MRI ligand, or an antibody molecule coupled to a detectable moiety.

253. The composition of embodiment 252, wherein the detectable moiety is or comprises a radiolabel, a fluorophore, a chromophore, or an affinity tag.

254. The composition of embodiment 253, wherein the radiolabel is or comprises tc99m, iodine-123, a spin label, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen- 15, oxygen-17, gadolinium, manganese, or iron.

255. A vector comprising a polynucleotide encoding the ligand of any one of embodiments 1-127 or 144-184.

256. A cell comprising the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the vector of embodiment 255, optionally wherein the cell is a mammalian cell, a cell of the central nervous system, or and/or a cell present in the blood brain barrier.

257. A method of making the composition of any one of embodiments 1-127 or 144-254, comprising:

(i) providing the ligand that binds to the GPI anchored protein, e.g., ALPL, and the active agent; and

(ii) incubating the ligand and active agent under conditions suitable to fuse or couple the ligand to the active agent,

thereby generating the composition.

258. A pharmaceutical composition comprising the composition of any one of embodiments 1-127 or 144-254 or the multispecific antibody molecule of any one of embodiments 128-143, and a pharmaceutically acceptable excipient.

259. A method of delivering an active agent, e.g., a therapeutic agent or a diagnostic agent, to a cell or tissue (e.g., a CNS cell or a CNS tissue), comprising administering the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

260. The method of embodiment 259, wherein the cell is a cell of a brain region or a spinal cord region, optionally a cell of the frontal cortex, sensory cortex, motor cortex, caudate, cerebellar cortex, cerebral cortex, brain stem, hippocampus, or thalamus.

261. The method of embodiment 259 or 260, wherein the cell or tissue is in a subject.

262. A method of increasing central nervous system transduction (e.g., increased crossing of the blood brain barrier) in a subject comprising administering the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

263. The method of embodiment 261 or 262, wherein the subject has, has been diagnosed with having, or is at risk of having a genetic disorder, e.g., a monogenic disorder or a polygenic disorder.

264. The method of any one of embodiments 261-263, wherein the subject has, has been diagnosed with having, or is at risk of having a neurological, e.g., a neurodegenerative disorder.

265. The method of any one of embodiments 261-264, wherein the subject has, has been diagnosed with having, or is at risk of having a neuro-oncological disorder.

266. The method of any one of embodiments 261-265, wherein the subject has, has been diagnosed with having, or is at risk of having a muscular disorder or a neuromuscular disorder.

267. A method of treating a subject having or diagnosed with having a genetic disorder, e.g., a monogenic disorder or a polygenic disorder, comprising administering to the subject the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

268. A method of treating a subject having or diagnosed with having a neurological disorder, e.g., a neurodegenerative disorder, comprising administering to the subject an effective amount of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

269. A method of treating a subject having or diagnosed with having a muscular disorder or a neuromuscular disorder, comprising administering to the subject an effective amount of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

270. A method of treating a subject having or diagnosed with having a neuro-oncological disorder, comprising administering to the subject an effective amount of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258.

271. The method of any one of embodiments 263-270, wherein the genetic disorder, neurological disorder, neurodegenerative disorder, muscular disorder, neuromuscular disorder, or neuro-oncological disorder is Huntington's Disease, Amyotrophic Lateral Sclerosis (ALS), Gaucher Disease, Dementia with Lewy Bodies, Parkinson's disease, Spinal Muscular Atrophy, Alzheimer's Disease, a leukodystrophy (e.g., Alexander disease, autosomal dominant leukodystrophy with autonomic diseases (ADLD), Canavan disease, cerebrotendinous xanthomatosis (CTX), metachromatic leukodystrophy (MLD), Pelizaeus-Merzbacher disease, or Refsum disease), or a cancer (e.g., a HER2/neu positive cancer or a glioblastoma).

272. The method of any one of embodiments 267-271, where treating comprises prevention of progression of the disease or disorder in the subject.

273. The method of embodiment 261-272, wherein the subject is a human.

274. The method of any one of embodiments 261-273, wherein the composition is administered to the subject intravenously, via intra-cisterna magna injection (ICM), intracerebrally, intrathecally, intracerebroventricularly, via intraparenchymal administration, intraarterially, or intramuscularly.

275. The method of any one of embodiments 261-274, wherein the composition is administered to the subject via focused ultrasound (FUS), e.g., coupled with the intravenous administration of microbubbles (FUS-MB), or MRI-guided FUS coupled with intravenous administration.

276. The method of any one of embodiments 261-275, wherein the composition is administered to the subject intravenously.

277. The method of any one of embodiments 261-276, wherein the composition is administered to the subject via intra-cisterna magna injection (ICM).

278. The method of any one of embodiments 261-277, wherein the composition is administered to the subject intraarterially.

279. The method of any one of embodiments 274-278, wherein administration of the composition results in a decreased presence, level, and/or activity of a gene, mRNA, protein, or combination thereof.

280. The method of any one of embodiments 274-278, wherein administration of the composition results in an increased presence, level, and/or activity of a gene, mRNA, protein, or a combination thereof.

281. The composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, for use in a method of delivering a payload to a cell or tissue.

282. The composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, for use in a method of treating a genetic disorder, a neurological disorder, a neurodegenerative disorder, a muscular disorder, a neuromuscular disorder, or a neuro-oncological disorder.

283. The composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, for use in the manufacture of a medicament.

284. The composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, for use in a method of increasing central nervous system transduction (e.g., increased crossing of the blood brain barrier) in a subject.

285. Use of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, in the manufacture of a medicament.

286. Use of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, in the manufacture of a medicament for treating a genetic disorder, a neurological disorder, a neurodegenerative disorder, a muscular disorder, a neuromuscular disorder, or a neuro-oncological disorder.

287. Use of the composition of any one of embodiments 1-127 or 144-254, the multispecific antibody molecule of any one of embodiments 128-143, or the pharmaceutical composition of embodiment 258, in the manufacture of a medicament for increasing central nervous system transduction (e.g., increased crossing of the blood brain barrier) in a subject.

[0015] The details of one or more embodiments of the disclosure are set forth in the accompanying description below. Other features, objects and advantages of the disclosure will be apparent from the description. In the description, the singular forms also include the plural unless the context clearly dictates otherwise. Certain terms are defined in the Definition section and throughout.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0017] **FIG. 1A** is a violin plot showing expression level of the payload on the Y-axis in various cell types as shown on the X-axis, which includes from left to right, microglia, astrocytes, endothelial cells subset 1, vascular smooth cells, pericytes, endothelial cells subset 2, committed oligodendrocytes, macrophages, vascular and leptomeningeal cells, committed oligodendrocytes subset 2, and mature oligodendrocytes. **FIG. 1A** relates to data shown in **Example 7** and **Table 22**. **FIG. 1B** is a violin plot showing expression level of ALPL on the Y-axis in various cell types as shown on the X-axis, which includes from left to right, microglia, astrocytes, endothelial cells subset 1, vascular smooth cells, pericytes, endothelial cells subset 2, committed oligodendrocytes, macrophages, vascular and leptomeningeal cells, committed oligodendrocytes subset 2, and mature oligodendrocytes.

[0018] **FIG. 2A** is a graph showing TTM-002 binding to ALPL at increasing concentrations of AAV by surface plasmon resonance (SPR) over time. **FIG. 2B** is a graph showing AAV9 binding to ALPL at increasing concentrations of AAV by SPR over time. **FIG. 2C** is a graph showing ALPL binding to TTM-002 at increasing concentrations of ALPL by surface plasmon resonance (SPR) over time. **FIG. 2D** is a graph showing ALPL binding to AAV9 at increasing concentrations of ALPL by SPR over time.

[0019] **FIG. 3A** is a graph showing ALPL binding to TTM-002 at increasing concentrations of ALPL by surface plasmon resonance (SPR) over time at a pH of 7.4, where the left half of the graph shows the association and the right half of the graph shows the dissociation. **FIG. 3B** is a graph showing ALPL binding to TTM-002 at increasing concentrations of ALPL by surface plasmon resonance (SPR) over time at a pH of 5.5, where the left half of the graph shows the association and the right half of the graph shows the dissociation.

[0020] **FIG. 4** is a graph showing the luciferase activity (RLU) as a measure of TTM-002 (right side of graph) or AAV9 (left side of graph) at 24-hours post-transduction and 48-hours post-transfection with siRNA 1, 2 or both siRNA 1 and 2 targeting ALPL or a non-ALPL control siRNA that did not knockdown ALPL.

[0021] **FIGs. 5A-5C** are a series of graphs demonstrating the effects of the small molecule inhibitor, TNAPi, of the ALPL receptor on TTM-002 transduction *in vitro* in HeLa cells expressing

ALPL. **FIG. 5A** is a graph showing luciferase activity as a measure of transduction of the TTM-002 capsid variant or the AAV9 capsid variant in the present of increasing concentrations of the TNAPi inhibitor. The concentrations tested include, from left to right on the X-axis, 0 nM (no inhibitor control), 24 nM, 48 nM, 95 nM, 190 nM, and 380 nM. **FIG. 5B** a graph showing luciferase activity as a measure of transduction of the TTM-002 capsid variant in the presence of increasing concentration of the TNAPi inhibitor or the DMSO vehicle control. The concentrations tested include, from left to right on the X-axis, 0 nM, 0.019 nM, 0.19 nM, 1.9 nM, 19 nM, and 190 nM. **FIG. 5C** is a graph showing the IC_{50} of the TNAPi inhibitor compared to the vehicle control for the TTM-002 capsid variant.

[0022] **FIGs. 6A-6C** are a series of graphs demonstrating the effects of the small molecule inhibitor, SBI-425, of the ALPL receptor on TTM-002 transduction *in vitro* in HeLa cells expressing ALPL.**FIG. 6A** is a graph showing luciferase activity as a measure of transduction of the TTM-002 capsid variant in the presence of increasing concentration of the SBI-425 inhibitor or the DMSO vehicle control. The concentrations tested include, from left to right on the X-axis, 0 nM, 0.00019 nM, 0.0019 nM, 0.019 nM, 0.19 nM, 1.9 nM, or 19.0 nM. **FIG. 6B** is a graph showing luciferase activity as a measure of transduction of the AAV9 capsid control in the presence of increasing concentration of the SBI-425 inhibitor or the DMSO vehicle control. The concentrations tested include, from left to right on the X-axis, 0 nM, 0.00019 nM, 0.0019 nM, 0.019 nM, 0.19 nM, 1.9 nM, or 19.0 nM. **FIG. 6C** is a graph showing the IC_{50} of the SNBI-425 inhibitor compared to the vehicle control for the TTM-002 capsid variant.

[0023] **FIG. 7A** is a series of graphs demonstrating ALPL binding to GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) (left graph) or GSGSKTINGHD**pS**PHKSGQNQ (SEQ ID NO: 4513) (right graph) at increasing concentrations of ALPL by SPR over time (seconds). **FIG. 7B** is a series of graphs demonstrating ALPL binding to GSGSNGHDSPHKSG (SEQ ID NO: 4500) (left graph) or GSGSNGHD**pS**PHKSG (SEQ ID NO: 4512) (right graph) at increasing concentrations of ALPL by SPR over time (seconds).

[0024] **FIG. 8A** is a series of graphs demonstrating the binding of GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) (left graph) or GSGSKTINGHD**pS**PHKSGQNQ (SEQ ID NO: 4513) (right graph) to ALPL at increasing concentrations of said peptides by SPR over time (seconds). **FIG. 8B** is a series of graphs demonstrating binding of GSGSNGHDSPHKSG (SEQ ID NO: 4500) (left graph) or GSGSNGHD**pS**PHKSG (SEQ ID NO: 4512) (right graph) to ALPL at increasing concentrations of said peptides by SPR over time (seconds).

[0025] **FIG. 9A** is a series of graphs showing binding of GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) (left graph) or GSGSKTINGHD**pS**PHKSGQNQ (SEQ ID NO: 4513) (right graph) to ALPL over time by Bio Layer Interferometry (BLI)/Octet. **FIG. 9B** is a series of graphs showing

binding of GSGSNGHDSPHKSG (SEQ ID NO: 4500) (left graph) or GSGSNGHD**pS**PHKSG (SEQ ID NO: 4512) (right graph) to ALPL over time by Bio Layer Interferometry (BLI)/Octet.

[0026] **FIG. 10A** is a graph showing the binding (OD450) to ALPL to increasing concentrations of GSGSNGHDSPHKSG (SEQ ID NO: 4500) or GSGSNGHD**pS**PHKSG (SEQ ID NO: 4512) (µg/mL) by ELISA. **FIG. 10B** is a graph showing the binding (OD450) to ALPL to increasing concentrations of GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) or GSGSKTINGHD**pS**PHKSGQNQ (SEQ ID NO: 4513) (µg/mL) by ELISA.

[0027] **FIG. 11A** is a graph depicting the antibody concentration in the top of the chamber prior to the transcytosis assay measured in µg/ml. The antibodies from left to right on the X-axis include: PT3 (non-ALPL binding control), MOPC (isotype control), Ab 9 (ALPL binding antibody), and Ab 22 (ALPL binding antibody). **FIG. 11B** is a graph showing the antibody concentration in the bottom chamber (pg/ml) for the antibodies indicated on the X-axis. **FIG. 11C** is a graph showing the percentage of the antibody detected in the bottom chamber relative to the load for the antibodies indicated on the X-axis. The left portion of the graph (labeled "MDCK ALPL Single Clone") depicts the percentage in the single clone MDCK ALPL expressing cells generated in Example 8 and the right portion of the graph (labeled "MDCK") shows the percentage in MDCK cells that do not express ALPL.

[0028] **FIG. 12** is a graph showing the luciferase activity (RLU) in cells pre-incubated with the antibody to ALPL as listed on the X-axis and described in **Table 40**, and subsequently transduced with an AAV particle comprising a TTM-002 capsid variant and encoding a GFP luciferase transgene. Low luciferase activity measured indicates that the antibody was able to compete for binding to ALPL with the TTM-002 capsid variant.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0029] Described herein, *inter alia*, are compositions comprising e.g., a fusion molecule or a conjugate molecule, comprising a ligand that binds to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL); and an active agent, e.g., a therapeutic agent or a diagnostic agent. In some embodiments, the ligand is fused or coupled, e.g., covalently or non-covalently to the active agent. In some embodiments, the GPI anchored protein is conserved in at least two to three species, e.g., at least three species (e.g., mice, NHPs (e.g., *Macaca fascicularis*), and/or humans). In some embodiments, the GPI anchored protein is present on the surface of a cell in the blood brain barrier. In some embodiments, the GPI anchored protein is ALPL, e.g., human or murine ALPL.

[0030] In some embodiments, a ligand to be used in a composition described herein is a ligand capable of binding ALPL. In some embodiments a ligand of the present disclosure is or comprises a peptide, a protein, an antibody molecule, a nucleic acid molecule (e.g., an aptamer), or a small molecule. In some embodiments, an active agent described herein is a therapeutic agent (e.g., a

protein (e.g., an enzyme), an antibody molecule, a nucleic acid molecule (e.g., an RNAi agent), or a small molecule). In some embodiments the active agent described herein is a diagnostic agent.

[0031] Without wishing to be bound by theory, it is believed in some embodiments, fusing or coupling, e.g., covalently (e.g., directly or via a linker) or non-covalently, a ligand that can bind ALPL to an active agent increases crossing of the blood brain barrier by the active agent relative to a active agent that is not fused or coupled to a ligand that can bind ALPL. Without wishing to be bound by theory, it is believed in some embodiments, that peptides comprising the amino acid sequences provided herein, e.g., in Tables 1, 2A, 2B, 2C, 13-19 (e.g., SEQ ID NOs: 2, 941, or 943), when fused or coupled, e.g., covalently (e.g., directly or via a linker) or non-covalently to an active agent, e.g., a therapeutic agent or a diagnostic agent, can enhance blood brain barrier crossing and biodistribution in the CNS of the active agent relative to the active agent alone.

Ligands

[0032] Disclosed herein are ligands that are capable of binding a protein present on a cell, e.g., a cell present in the blood brain barrier. In some embodiments, the ligand binds a GPI anchored protein. In some embodiments, the GPI anchored protein is conserved in at least two to three species, e.g., at least three species (e.g., mice, NHPs (e.g., *Macaca fascicularis*), and/or humans). In some embodiments, the GPI anchored protein is alkaline phosphatase issue-nonspecific isozyme (NM_000478.4, which is incorporated by reference herein) (ALPL).

[0033] ALPL is part of a family of membrane-bound glycoproteins that hydrolyze monophosphate esters at a high pH (see, e.g., Weiss et al., Isolation and characterization of a cDNA encoding a human liver/bone/kidney-type alkaline phosphatase. Proc. Nat. Acad. Sci., 83: 7182-7186 (1986), the contents of which are hereby incorporated by reference in their entirety). ALPL is highly conserved across humans, mice, and cynomolgus macaques (*Macaca fascicularis*) when compared by sequence alignment (e.g., as shown in **Table 24**). Additionally, in humans ALPL is expressed on endothelial cells and neurons, and at a low level on astrocytes. The highest level of ALPL expression in human is on endothelial cells. In mice, ALPL is more highly expressed on astrocytes, oligodendrocyte progenitor cells (OPCs), and to a lesser extent on endothelial cells. Without wishing to be bound by theory, it is believed in some embodiments that highly conserved nature of the ALPL receptor protein across species is predictive of cross-species compatibility of the AAV capsid variants described herein.

[0034] In some embodiments, the ligand binds an ALPL protein comprising an amino acid sequence or encoded by a nucleotide sequence provided in Table 32, or a sequence at least 70% (e.g., 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%) identical thereto. In some embodiments, the ligand binds a human ALPL protein, e.g., a human ALPL protein comprising the amino acid sequence of SEQ ID NO: 3, or an amino acid sequence at least 70% (e.g., 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%) identical thereto. In some embodiments, the ALPL is a murine ALPL, e.g., a

murine ALPL comprising the amino acid sequence of SEQ ID NO: 14, or an amino acid sequence at least 70% (e.g., 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%) identical thereto.

Table 32: Exemplary ALPL Sequences

Description	Sequence	SEQ ID NO:
Homo sapiens alkaline phosphatase, liver/bone/kidney (ALPL), transcript variant 1, mRNA, NM_000478.4	CCGGGCCTCACTCGGGCCCCGCGGCCGCCTTTATAAGGCGGCGGGGGTGGTGG CCCGGGCCGCGTTGCGCTCCCGCCACTCCGCGCCCGCTATCCTGGCTCCGTGC TCCCACGCGCTTGTGCCTGGACGGACCCTCGCCAGTGCTCTGCGCAGGATTGG AACATCAGTTAACATCTGACCACTGCCAGCCCACCCCCTCCCACCCACGTCGA TTGCATCTCTGGGCTCCAGGGATAAAGCAGGTCTTGGGGTGCACCATGATTTC ACCATTCTTAGTACTGGCCATTGGCACCTGCCTTACTAACTCCTTAGTGCCAG AGAAAGAGAAAGACCCCAAGTACTGGCGAGACCAAGCGCAAGAGACACTGAAA TATGCCCTGGAGCTTCAGAAGCTCAACACCAACGTGGCTAAGAATGTCATCAT GTTCCTGGGAGATGGGATGGGTGTCTCCACAGTGACGGCTGCCCGCATCCTCA AGGGTCAGCTCCACCACAACCCTGGGGAGGAGACCAGGCTGGAGATGGACAAG TTCCCCTTCGTGGCCCTCTCCAAGACGTACAACACCAATGCCCAGGTCCCTGA CAGCGCCGGCACCGCCACCGCCTACCTGTGTGGGGTGAAGGCCAATGAGGGCA CCGTGGGGGTAAGCGCAGCCACTGAGCGTTCCCGGTGCAACACCACCCAGGGG AACGAGGTCACCTCCATCCTGCGCTGGGCCAAGGACGCTGGGAAATCTGTGGG CATTGTGACCACCACGAGAGTGAACCATGCCACCCCCAGCGCCGCCTACGCCC ACTCGGCTGACCGGGACTGGTACTCAGACAACGAGATGCCCCCTGAGGCCTTG AGCCAGGGCTGTAAGGACATCGCCTACCAGCTCATGCATAACATCAGGGACAT TGACGTGATCATGGGGGGTGGCCGGAAATACATGTACCCCAAGAATAAAACTG ATGTGGAGTATGAGAGTGACGAGAAAGCCAGGGGCACGAGGCTGGACGGCCTG GACCTCGTTGACACCTGGAAGAGCTTCAAACCGAGATACAAGCACTCCCACTT CATCTGGAACCGCACGGAACTCCTGACCCTTGACCCCCACAATGTGGACTACC TATTGGGTCTCTTCGAGCCAGGGGACATGCAGTACGAGCTGAACAGGAACAAC GTGACGGACCCGTCACTCTCCGAGATGGTGGTGGTGGCCATCCAGATCCTGCG GAAGAACCCCAAAGGCTTCTTCTTGCTGGTGGAAGGAGGCAGAATTGACCACG GGCACCATGAAGGAAAAGCCAAGCAGGCCCTGCATGAGGCGGTGGAGATGGAC CGGGCCATCGGGCAGGCAGGCAGCTTGACCTCCTCGGAAGACACTCTGACCGT GGTCACTGCGGACCATTCCCACGTCTTCACATTTGGTGGATACACCCCCCGTG GCAACTCTATCTTTGGTCTGGCCCCCATGCTGAGTGACACAGACAAGAAGCCC TTCACTGCCATCCTGTATGGCAATGGGCCTGGCTACAAGGTGGTGGGCGGTGA ACGAGAGAATGTCTCCATGGTGGACTATGCTCACAACAACTACCAGGCGCAGT CTGCTGTGCCCCTGCGCCACGAGACCCACGGCGGGGAGGACGTGGCCGTCTTC TCCAAGGGCCCCATGGCGCACCTGCTGCACGGCGTCCACGAGCAGAACTACGT CCCCCACGTGATGGCGTATGCAGCCTGCATCGGGGCCAACCTCGGCCACTGTG CTCCTGCCAGCTCGGCAGGCAGCCTTGCTGCAGGCCCCCTGCTGCTCGCGCTG GCCCTCTACCCCCTGAGCGTCCTGTTCTGAGGGCCCAGGGCCCGGGCACCCAC AAGCCCGTGACAGATGCCAACTTCCCACACGGCAGCCCCCCCCTCAAGGGGCA GGGAGGTGGGGGCCTCCTCAGCCTCTGCAACTGCAAGAAAGGGGACCCAAGAA ACCAAAGTCTGCCGCCCACCTCGCTCCCCTCTGGAATCTTCCCCAAGGGCCAA ACCCACTTCTGGCCTCCAGCCTTTGCTCCCTCCCCGCTGCCCTTTGGCCAACA GGGTAGATTTCTCTTGGGCAGGCAGAGAGTACAGACTGCAGACATTCTCAAAG CCTCTTATTTTTCTAGCGAACGTATTTCTCCAGACCCAGAGGCCCTGAAGCCT CCGTGGAACATTCTGGATCTGACCCTCCCAGTCTCATCTCCTGACCCTCCCAC TCCCATCTCCTTACCTCTGGAACCCCCCAGGCCCTACAATGCTCATGTCCCTG TCCCCAGGCCCAGCCCTCCTTCAGGGGAGTTGAGGTCTTTCTCCTCAGGACAA GGCCTTGCTCACTCACTCACTCCAAGACCACCAGGGTCCCAGGAAGCCGGTGC CTGGGTGGCCATCCTACCCAGCGTGGCCCAGGCCGGGAAGAGCCACCTGGCAG GGCTCACACTCCTGGGCTCTGAACACACACGCCAGCTCCTCTCTGAAGCGACT CTCCTGTTTGGAACGGCAAAAAAAAATTTTTTTTTCTCTTTTTGGTGGTGGTT AAAAGGGAACACAAAACATTTAAATAAAACTTTCCAAATATTTCCGAGGACAA AAAAAAAAA	17
Homo sapiens alkaline phosphatase,	**MISPFLVLAIGTCLTNS**LVPEKEKDPKYWRDQAQETLKYALELQKLNTNVAKN VIMFLGDGMGVSTVTAARILKGQLHHNPGEETRLEMDKFPFVALSKTYNTNAQ VPDSAGTATAYLCGVKANEGTVGVSAATERSRCNTTQGNEVTSILRWAKDAGK	18

tissue-nonspecific isozyme isoform 1 (signal sequence underlined), NP_000469.3	SVGIVTTTRVNHATPSAAYAHSADRDWYSDNEMPPEALSQGCKDIAYQLMHNI RDIDVIMGGGRKYMYPKNKTDVEYESDEKARGTRLDGLDLVDTWKSFKPRYKH SHFIWNRTELLTLDPHNVDYLLGLFEPGDMQYELNRNNVTDPSLSEMVVVAIQ ILRKNPKGFFLLVEGGRIDHGHHEGKAKQALHEAVEMDRAIGQAGSLTSSEDT LTVVTADHSHVFTFGGYTPRGNSIFGLAPMLSDTDKKPFTAILYGNGPGYKVV GGERENVSMVDYAHNNYQAQSAVPLRHETHGGEDVAVFSKGPMAHLLHGVHEQ NYVPHVMAYAACIGANLGHCAPASSAGSLAAGPLLLALALYPLSVLF	
Homo sapiens alkaline phosphatase, placental-like 2 (ALPPL2), NM_031313.3	ATACTCCATACCTGGGATTTCCGCCTCGCCGCTCTCCGACTGCTTCCAGACAT GCAGGGGCCCTGGGTGCTGCTCCTGCTGGGCCTGAGGCTACAGCTCTCCCTGG GCATCATCCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCCAGGCAGCC GAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTGCACAGACAGCCGCCAAGAA CCTCATCATCTTCCTGGGTGACGGGATGGGGGTGTCTACGGTGACAGCTGCCA GGATCCTAAAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGACCTTCCTGGCC ATGGACCGCTTCCCGTACGTGGCTCTGTCCAAGACATACAGTGTAGACAAGCA TGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCA ACTTCCAGACCATTGGCTTGAGTGCAGCCGCCCGCTTTAACCAGTGCAACACG ACACGCGGCAACGAGGTCATCTCCGTGATGAATCGGGCCAAGAAAGCAGGAAA GTCAGTGGGAGTGGTAACCACCACACGGGTGCAGCATGCCTCGCCAGCCGGCG CCTACGCCCACACGGTGAACCGCAACTGGTACTCGGATGCCGACGTGCCTGCC TCGGCCCGCCAGGAGGGGTGCCAGGACATCGCCACGCAGCTCATCTCCAACAT GGACATTGATGTGATCCTAGGTGGAGGCCGAAAGTACATGTTTCCCATGGGGA CCCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGACCAGGCTGGAC GGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCACCAGGGTGCCCGGTACGT GTGGAACCGCACTGAGCTCCTGCAGGCTTCCCTGGACCCGTCTGTGACCCATC TCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCC ACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCTCCTGCTGAG CAGGAACCCCCGCGGCTTCTTCCTCTTCGTGGAGGGTGGTCGCATCGACCATG GTCATCATGAAAGCAGGGCTTACCGGGCACTGACTGAGACGATCATGTTCGAC GACGCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGAGCCT CGTCACTGCCGACCACTCCCACGTCTTCTCCTTCGGAGGCTACCCCCTGCGAG GGAGCTCCATCTTCGGGCTGGCCCCTGGCAAGGCCCGGGACAGGAAGGCCTAC ACGGTCCTCCTATACGGAAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCG GCCGGATGTTACGGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAG CAGTGCCCCTGGACGGAGAGACCCACGCAGGCGAGGACGTGGCGGTGTTCGCG CGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGCAGACCTTCATAGC GCACGTCATGGCCTTCGCCGCCTGCCTGGAGCCCTACACCGCCTGCGACCTGG CGCCCCGCGCCGGCACCACCGACGCCGCGCACCCGGGGCCGTCCGTGGTCCCC GCGTTGCTTCCTCTGCTGGCAGGGACCTTGCTGCTGCTGGGGACGGCCACTGC TCCCTGAGTGTCCCGTCCCTGGGGCTCCTGCTTCCCCATCCCGGAGTTCCCCT GCTCCCCACCTCCAGTCGTCCTGCCGGACCTCCACCTGGAGCTGTCACCCCCG GAGTCGCCACACAGACGTCCTGCCATGGAACCTTCCCCTCCCGGTGCACCCTG GGGACCGAGCCCTTGACACCACGCCCTTTGCTTTATCTTGCTCTTGAAATTTT GGCCCCAACTCCAGGGACTGGGGATTTGTGCCTGGCAGCTGCCTGCATTTCAG GAAAAGAGGAGGCTCAGACCATCCAGCCCCCGCCCATATCCTGAGGTGGATCA GGCAGGCTCTCTCCCCGGGGACATGAGGCACCCATACCTAGGACCCCCTGCGC CTTTTTTAGCTTCAGTCATGGCAGCACCTGAGGGACACAAGGACTTGGGTGCA TCAGGACGCCTTGGAGAAGCGTGGCTTCCTGCCACCCTGCAACCCACCCTCCC AGCCAAGGAGGCTGCTGTGGTGGGGATCCCCAGGGGGGCTTTGACACAGTCCT CTGCTGTCCCTCCACTGGGCTAATTCTACACCCCTGTGCCCCTCCTAGGGGCC CATGAGTCAGAGAGGCTTGCCCCAAGTCACAGCCACTCAGATGTTCGACGCCC CCTAAGGTCCATTCCAGCACCCACCTGAGTTCCGAGGAGCACCTGGGAAGCTC TGGGTGCAGGATAGCAGTCCAGAGTCCATGGCCCCGCCTAGGCCATCTGGGTG CTGGGCATGGATTTCTCAGCAAGGAAGACTCATTACCTTCCCTCCCTGGGCCT CCATTCTTCTGGGAAACACAAAGCAATAATAAAAGGAAGTGTTAGACAATGTA A	4
Homo sapiens alkaline phosphatase, placental-like 2 (ALPPL2), NP_112603.2	MQGPWVLLLLGLRLQLSLGIIPVEEENPDFWNRQAAEALGAAKKLQPAQTAAK NLIIFLGDGMGVSTVTAARILKGQKKDKLGPETFLAMDRFPYVALSKTYSVDK HVPDSGATATAYLCGVKGNFQTIGLSAAARFNQCNTTRGNEVISVMNRAKKAG KSVGVVTTTRVQHASPAGAYAHTVNRNWYSDADVPASARQEGCQDIATQLISN MDIDVILGGGRKYMFPMGTPDPEYPDDYSQGGTRLDGKNLVQEWLAKHQGARY VWNRTELLQASLDPSVTHLMGLFEPGDMKYEIHRDSTLDPSLMEMTEAALLLL	5

	SRNPRGFFLFVEGGRIDHGHHESRAYRALTETIMFDDAIERAGQLTSEEDTLS LVTADHSHVFSFGGYPLRGSSIFGLAPGKARDRKAYTVLLYGNGPGYVLKDGA RPDVTESESGSPEYRQQSAVPLDGETHAGEDVAVFARGPQAHLVHGVQEQTFI AHVMAFAACLEPYTACDLAPRAGTTDAAHPGPSVVPALLPLLAGTLLLLGTAT AP	
Homo sapiens alkaline phosphatase, placental (ALPP), NM_001632.5	ATACTCCATGCCCAGAATTCCTGCCTCGCCACTGTCCTGCTGCCCTCCAGACA TGCTGGGGCCCTGCATGCTGCTGCTGCTGCTGCTGCTGGGCCTGAGGCTACAG CTCTCCCTGGGCATCATCCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCG CGAGGCAGCCGAGGCCCTGGGTGCCGCCAAGAAGCTGCAGCCTGCACAGACAG CCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCTACGGTG ACAGCTGCCAGGATCCTAAAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGAT ACCCCTGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATACAATG TAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGG GTCAAGGGCAACTTCCAGACCATTGGCTTGAGTGCAGCCGCCCGCTTTAACCA GTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGATGAATCGGGCCAAGA AAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGCCTCG CCAGCCGGCACCTACGCCCACACGGTGAACCGCAACTGGTACTCGGACGCCGA CGTGCCTGCCTCCGCCCGCCAGGAGGGGTGCCAGGACATCGCTACGCAGCTCA TCTCCAACATGGACATTGACGTGATCCTAGGTGGAGGCCGAAAGTACATGTTT CGCATGGGAACCCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGAC CAGGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTG CCCGGTATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGTCT GTGACCCATCTCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCA CCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGC GCCTGCTGAGCAGGAACCCCCGCGGCTTCTTCCTCTTCGTGGAGGGTGGTCGC ATCGACCATGGTCATCATGAAAGCAGGGCTTACCGGGCACTGACTGAGACGAT CATGTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACA CGCTGAGCCTCGTCACTGCCGACCACTCCCACGTCTTCTCCTTCGGAGGCTAC CCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCTGGCAAGGCCCGGGACAG GAAGGCCTACACGGTCCTCCTATACGGAAACGGTCCAGGCTATGTGCTCAAGG ACGGCGCCCGGCCGGATGTTACCGAGAGCGAGAGCGGGAGCCCCGAGTATCGG CAGCAGTCAGCAGTGCCCCTGGACGAAGAGACCCACGCAGGCGAGGACGTGGC GGTGTTCGCGCGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGCAGA CCTTCATAGCGCACGTCATGGCCTTCGCCGCCTGCCTGGAGCCCTACACCGCC TGCGACCTGGCGCCCCCCGCCGGCACCACCGACGCCGCGCACCCGGGGCGGTC CGTGGTCCCCGCGTTGCTTCCTCTGCTGGCCGGGACCCTGCTGCTGCTGGAGA CGGCCACTGCTCCCTGAGTGTCCCGTCCCTGGGGCTCCTGCTTCCCCATCCCG GAGTTCTCCTGCTCCCCACCTCCTGTCGTCCTGCCTGGCCTCCAGCCCGAGTC GTCATCCCCGGAGTCCCTATACAGAGGTCCTGCCATGGAACCTTCCCCTCCCC GTGCGCTCTGGGGACTGAGCCCATGACACCAAACCTGCCCCTTGGCTGCTCTC GGACTCCCTACCCCAACCCCAGGGACTGCAGGTTGTGCCCTGTGGCTGCCTGC ACCCCAGGAAAGGAGGGGGCTCAGGCCATCCAGCCACCACCTACAGCCCAGTG GGTACCAGGCAGGCTCCCTTCCTGGGGAAAAGAAGCACCCAGACCCCGCGCCC CGCTGATCTTTGCTTCAGTCCTTGAATCACCTGTGGGACTTGAGGACTCGGGA TCTTCAGGACGCCTGGAGAAGGGTGGTTTCCTGCCACCCTGCTGGCCAAGGAG GCTCCTGGGGTGGGGATCACCAGGGGGATTTTGACACAGCCTTCGGCTGCCCC CCACTAAGCTAATTCCACACCCCTGTACCCCCCCCAGGGGGCCCTCTGCCTCAT GGCAAAGGCTTGCCCCAAATCTCAACTTCTCAGACGTTCCATACCCCCACATG CCAATTTCAGCACCCAACTGAGATCCGAGGAGCTCCTGGGAAGCCCTGGGTGC AGGACACTGGTCGAGAGCCAAAGGTCCCTCCCCAGACATCTGGACACTGGGCA TAGATTTCTCAAGAAGGAAGACTCCCCTGCCTCCCCAGGGCCTCTGCTCTCCT GGGAGACAAAGCAATAATAAAAGGAAGTGTTTGTAATCCCAGCACTTTGGGAG GCCGAGGTGGGCGGATCACGAGGTCAGGAGATGGAGACCATCCTGGCTAACAC GGTGAAACCCCTTATCTATGCGCCTGTAGTCCCAGCTACCCAGGAGGCTGAAG CAGGATAATCGCTTGAACCCGGGCGGCGGAGATTGCAGTGAGCCGAGGTCATG CCACTGCACTGCAGCCTGGGCGACAGAGCGAGATTCTGCCTCAAAAATAAACA AATAAATTTTAAAAATAAATAAATAATAAAAGGAAGTGTTAGACAATGTAA	6
Homo sapiens alkaline phosphatase,	MLGPCMLLLLLLLGLRLQLSLGIIPVEEENPDFWNREAAEALGAAKKLQPAQT AAKNLIIFLGDGMGVSTVTAARILKGQKKDKLGPEIPLAMDRFPYVALSKTYN VDKHVPDSGATATAYLCGVKGNFQTIGLSAAARFNQCNTTRGNEVISVMNRAK KAGKSVGVVTTTRVQHASPAGTYAHTVNRNWYSDADVPASARQEGCQDIATQL	7

placental type, protein, NP_001623.3	ISNMDIDVILGGGRKYMFRMGTPDPEYPDDYSQGGTRLDGKNLVQEWLAKRQG ARYVWNRTELMQASLDPSVTHLMGLFEPGDMKYEIHRDSTLDPSLMEMTEAAL RLLSRNPRGFFLFVEGGRIDHGHHESRAYRALTETIMFDDAIERAGQLTSEED TLSLVTADHSHVFSFGGYPLRGSSIFGLAPGKARDRKAYTVLLYGNGPGYVLK DGARPDVTESESGSPEYRQQSAVPLDEETHAGEDVAVFARGPQAHLVHGVQEQ TFIAHVMAFAACLEPYTACDLAPPAGTTDAAHPGRSVVPALLPLLAGTLLLLE TATAP	
Homo sapiens alkaline phosphatase, intestinal (ALPLI), NM_001631.5	CGGTTCCTGGTGTCCCCACTTCGCCTCCCTCCTGCTGCCCCCAAGACATGCAG GGGCCCTGGGTGCTGCTGCTGCTGGGCCTGAGGCTACAGCTCTCCCTGGGCGT CATCCCAGCTGAGGAGGAGAACCCGGCCTTCTGGAACCGCCAGGCAGCTGAGG CCCTGGATGCTGCCAAGAAGCTGCAGCCCATCCAGAAGGTCGCCAAGAACCTC ATCCTCTTCCTGGGCGATGGGTTGGGGGTGCCCACGGTGACAGCCACCAGGAT CCTAAAGGGGCAGAAGAATGGCAAACTGGGGCCTGAGACGCCCCTGGCCATGG ACCGCTTCCCATACCTGGCTCTGTCCAAGACATACAATGTGGACAGACAGGTG CCAGACAGCGCAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGCCAACTT CCAGACCATCGGCTTGAGTGCAGCCGCCCGCTTTAACCAGTGCAACACGACAC GCGGCAATGAGGTCATCTCCGTGATGAACCGGGCCAAGCAAGCAGGAAAGTCA GTAGGAGTGGTGACCACCACACGGGTGCAGCACGCCTCGCCAGCCGGCACCTA CGCACACACAGTGAACCGCAACTGGTACTCAGATGCTGACATGCCTGCCTCAG CCCGCCAGGAGGGGTGCCAGGACATCGCCACTCAGCTCATCTCCAACATGGAC ATTGACGTGATCCTTGGCGGAGGCCGCAAGTACATGTTTCCCATGGGGACCCC AGACCCTGAGTACCCAGCTGATGCCAGCCAGAATGGAATCAGGCTGGACGGGA AGAACCTGGTGCAGGAATGGCTGGCAAAGCACCAGGGTGCCTGGTATGTGTGG AACCGCACTGAGCTCATGCAGGCGTCCCTGGACCAGTCTGTGACCCATCTCAT GGGCCTCTTTGAGCCCGGAGACACGAAATATGAGATCCACCGAGACCCCACAC TGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGG AACCCCCGCGGCTTCTACCTCTTTGTGGAGGGCGGCCGCATCGACCATGGTCA TCATGAGGGTGTGGCTTACCAGGCACTCACTGAGGCGGTCATGTTCGACGACG CCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGACCCTCGTC ACCGCTGACCACTCCCATGTCTTCTCCTTTGGTGGCTACACCTTGCGAGGGAG CTCCATCTTCGGGTTGGCCCCCAGCAAGGCTCAGGACAGCAAAGCCTACACGT CCATCCTGTACGGCAATGGCCCGGGCTACGTGTTCAACTCAGGCGTGCGACCA GACGTGAATGAGAGCGAGAGCGGGAGCCCCGATTACCAGCAGCAGGCGGCGGT GCCCCTGTCGTCCGAGACCCACGGAGGCGAAGACGTGGCGGTGTTTGCGCGCG GCCCGCAGGCGCACCTGGTGCATGGTGTGCAGGAGCAGAGCTTCGTAGCGCAT GTCATGGCCTTCGCTGCCTGTCTGGAGCCCTACACGGCCTGCGACCTGGCGCC TCCCGCCTGCACCACCGACGCCGCGCACCCAGTTGCCGCGTCGCTGCCACTGC TGGCCGGGACCCTGCTGCTGCTGGGGGCGTCCGCTGCTCCCTGAGTGCCCCAC TCCGGAGTTATCCTGCTCCCCACCTCCGGGCGTCCTGCCCTGTTCCCCGTCCT GAGCCGCCACTTCCAGCGAACACACACAGGTGTCCTGCCGTTGGACCTTCACC TCCTAGAGATAAACCAGCCTCAGCTGGCGCAGCGGGGCCCTTCTTCCCTCCGC ATCCCCTTCAGGGAGCAGGAGCCCAGGGCGCCCTGGGAGCTGAGCCTGGGACT TCCAGGACCTCCCCTCAGGTTGTTCTCTGATTCTTCCTCCCAACCCCAGAGAC TGCAGATTTGTGCCATGCGGCTGCCTGCACCCCAGACAATAAAGGGACCAAAA CCACCCAACCCCCACCCTGCCTCTATCCTAAGGAAGACCAAGCAGGCCTGGAC CCAGAGACGTCCCCCATCGTGGGACACGACACACCCAGACCGCGTGCCCCACC GTCTTAGCTTCAATCCTGGCAGCACCTGGTAGACCCAAGGACTTGGGTGGATC AGGACACCTGAAGAAGAGAAGCTTCCGGCAACCCTGCAACCCACCCAAGGAGG CTACTGGATCGGGGATTCCCAGGGGGGCTTTGACACAGTCCTCTGCTGTCTCC CCACTAGGATCATTCCACACCCCTGCACCTGACCAAGGGACCAATGAGGCAGA GGCTTGCCCCAAGTCACAGCCACTCAGATGCTTCCTGCCCCCCAGTGCCCATT CCAGGTCACCAGATCCAAGGAGCGCTTGAGGAGCTCTGGGTACAGGGCAGCAA CCCAGAGCCCATGGGCCCTCCCGGGACATCTGGATGCTGGGCATAGATTTCTC AACAAGGAAGACTCCCCTGCCTCCTCAAGGTCTCCATTCTCCTAGGAGACAAA GCAATAATAAAAGGTGTTAGACAATGTAATGCCAGTACTACTTCCTAGGAGAA AAATCATGAGTGAGTGTGGGCACAGTATCTGGAGAGGTGGATAACGCAGGCCA GGAGGTACTGCTGAGGGGCAGATGATTGAGCAAGAGACTTGAACAGAGTGGGG GCTTGAGCAAGGCAGCACAGCAGTGCAAACGCCCTGGGGCAGTGTCAGCAGGT GCTCTGGGAGGCCAAGGGCTGGATCAGAGGGGTGGGGGTGGGTGGGCAGAGTG GGGAAAGCCTGAGGGGTCAGGAGAGTGGGGTGTGCATGGGGGACTGTGAAGTC TGGTTAGAGGGGTGTGGTTGGAGGTCTTTGAGGAGGGCTGTGACCTGCCCTGG	8

	TTGGGAAATAAGCACTCTGGCTGCTGCCAGGAGAAGGGTCTGGTCTTTTGGGC AGAGGGTGGGGGTGGTGGCAGGCTCAGGTGAAAGCTGGGGAAGGAGCTGACTC CAGGTGTTTCTGACCTCCCTCTGAAAGTATTCTGGAGCGCCCATCCCAATACA GCCATACTTAGTGAGTACACACCTGCTCCAAGAGAACATTGAAAAGAATAAAG GTGAAATCAACCACATTTTCCAGCAAATTTTGCAGTATTACAAATTTATTTGT ACATTTACAAAGGTGCAAAAAAGCATCTTGCTTTTGCAAGAAATAGTAACATC ATTCAATATGCTTTCTTATTTACTAAAACCTTGAAATAAAATTGTAAAACATC AGTTTGAA	
Homo sapiens alkaline phosphatase, intestinal (ALPLI) protein, NP_001622.2	MQGPWVLLLLGLRLQLSLGVIPAEEENPAFWNRQAAEALDAAKKLQPIQKVAK NLILFLGDGLGVPTVTATRILKGQKNGKLGPETPLAMDRFPYLALSKTYNVDR QVPDSAATATAYLCGVKANFQTIGLSAAARFNQCNTTRGNEVISVMNRAKQAG KSVGVVTTTRVQHASPAGTYAHTVNRNWYSDADMPASARQEGCQDIATQLISN MDIDVILGGGRKYMFPMGTPDPEYPADASQNGIRLDGKNLVQEWLAKHQGAWY VWNRTELMQASLDQSVTHLMGLFEPGDTKYEIHRDPTLDPSLMEMTEAALRLL SRNPRGFYLFVEGGRIDHGHHEGVAYQALTEAVMFDDAIERAGQLTSEEDTLT LVTADHSHVFSFGGYTLRGSSIFGLAPSKAQDSKAYTSILYGNGPGYVFNSGV RPDVNESESGSPDYQQQAAVPLSSETHGGEDVAVFARGPQAHLVHGVQEQSFV AHVMAFAACLEPYTACDLAPPACTTDAAHPVAASLPLLAGTLLLLGASAAP	9
Homo sapiens alkaline phosphatase, biomineralization associated (ALPL), transcript variant 2, mRNA NM_001127501.4	GCCGCGTTGCGCTCCCGCCACTCCGCGCCCGCTATCCTGGCTCCGTGCTCCCA CGCGCTTGTGCCTGGACGGACCCTCGCCAGTGCTCTGCGCAGAGAAAGAGAAA GACCCCAAGTACTGGCGAGACCAAGCGCAAGAGACACTGAAATATGCCCTGGA GCTTCAGAAGCTCAACACCAACGTGGCTAAGAATGTCATCATGTTCCTGGG AGATGGGATGGGTGTCTCCACAGTGACGGCTGCCCGCATCCTCAAGGGTCAGC TCCACCACAACCCTGGGGAGGAGACCAGGCTGGAGATGGACAAGTTCCCCTTC GTGGCCCTCTCCAAGACGTACAACACCAATGCCCAGGTCCCTGACAGTGCCGG CACCGCCACCGCCTACCTGTGTGGGGTGAAGGCCAATGAGGGCACCGTGGGGG TAAGCGCAGCCACTGAGCGTTCCCGGTGCAACACCACCCAGGGGAACGAGGTC ACCTCCATCCTGCGCTGGGCCAAGGACGCTGGGAAATCTGTGGGCATTGTGAC CACCACGAGAGTGAACCATGCCACCCCCAGCGCCGCCTACGCCCACTCGGCTG ACCGGGACTGGTACTCAGACAACGAGATGCCCCCTGAGGCCTTGAGCCAGGGC TGTAAGGACATCGCCTACCAGCTCATGCATAACATCAGGGACATTGACGTGAT CATGGGGGGTGGCCGGAAATACATGTACCCCAAGAATAAAACTGATGTGGAGT ATGAGAGTGACGAGAAAGCCAGGGGCACGAGGCTGGACGGCCTGGACCTCGTT GACACCTGGAAGAGCTTCAAACCGAGATACAAGCACTCCCACTTCATCTGGAA CCGCACGGAACTCCTGACCCTTGACCCCCACAATGTGGACTACCTATTGGGTC TCTTCGAGCCAGGGGACATGCAGTACGAGCTGAACAGGAACAACGTGACGGAC CCGTCACTCTCCGAGATGGTGGTGGTGGCCATCCAGATCCTGCGGAAGAACCC CAAAGGCTTCTTCTTGCTGGTGGAAGGAGGCAGAATTGACCACGGGCACCATG AAGGAAAAGCCAAGCAGGCCCTGCATGAGGCGGTGGAGATGGACCGGGCCATC GGGCAGGCAGGCAGCTTGACCTCCTCGGAAGACACTCTGACCGTGGTCACTGC GGACCATTCCCACGTCTTCACATTTGGTGGATACACCCCCCGTGGCAACTCTA TCTTTGGTCTGGCCCCCATGCTGAGTGACACAGACAAGAAGCCCTTCACTGCC ATCCTGTATGGCAATGGGCCTGGCTACAAGGTGGTGGGCGGTGAACGAGAGAA TGTCTCCATGGTGGACTATGCTCACAACAACTACCAGGCGCAGTCTGCTGTGC CCCTGCGCCACGAGACCCACGGCGGGGAGGACGTGGCCGTCTTCTCCAAGGGC CCCATGGCGCACCTGCTGCACGGCGTCCACGAGCAGAACTACGTCCCCCACGT GATGGCGTATGCAGCCTGCATCGGGGCCAACCTCGGCCACTGTGCTCCTGCCA GCTCGGCAGGCAGCCTTGCTGCAGGCCCCCTGCTGCTCGCGCTGGCCCTCTAC CCCCTGAGCGTCCTGTTCTGAGGGCCCAGGGCCCGGGCACCCACAAGCCCGTG ACAGATGCCAACTTCCCACACGGCAGCCCCCCCCTCAAGGGGCAGGGAGGTGG GGGCCTCCTCAGCCTCTGCAACTGCAAGAAAGGGGACCCAAGAAACCAAAGTC TGCCGCCCACCTCGCTCCCCTCTGGAATCTTCCCCAAGGGCCAAACCCACTTC TGGCCTCCAGCCTTTGCTCCCTCCCCGCTGCCCTTTGGCCAACAGGGTAGATT TCTCTTGGGCAGGCAGAGAGTACAGACTGCAGACATTCTCAAAGCCTCTTATT TTTCTAGCGAACGTATTTCTCCAGACCCAGAGGCCCTGAAGCCTCCGTGGAAC ATTCTGGATCTGACCCTCCCAGTCTCATCTCCTGACCCTCCCACTCCCATCTC CTTACCTCTGGAACCCCCCAGGCCCTACAATGCTCATGTCCCTGTCCCCAGGC CCAGCCCTCCTTCAGGGGAGTTGAGGTCTTTCTCCTCAGGACAAGGCCTTGCT CACTCACTCACTCCAAGACCACCAGGGTCCCAGGAAGCCGGTGCCTGGGTGGC CATCCTACCCAGCGTGGCCCAGGCCGGGAAGAGCCACCTGGCAGGGCTCACAC TCCTGGGCTCTGAACACACACGCCAGCTCCTCTCTGAAGCGACTCTCCTGTTT	10

	GGAACGGCAAAAAAAAATTTTTTTTTCTCTTTTTGGTGGTGGTTAAAAGGGAA CACAAAACATTTAAATAAAACTTTCCAAATATTTCCGAGGA	
Homo sapiens alkaline phosphatase, tissue-nonspecific isozyme isoform 2, protein, NP_001120973.2	MFLGDGMGVSTVTAARILKGQLHHNPGEETRLEMDKFPFVALSKTYNTNAQVP DSAGTATAYLCGVKANEGTVGVSAATERSRCNTTQGNEVTSILRWAKDAGKSV GIVTTTRVNHATPSAAYAHSADRDWYSDNEMPPEALSQGCKDIAYQLMHNIRD IDVIMGGGRKYMYPKNKTDVEYESDEKARGTRLDGLDLVDTWKSFKPRYKHSH FIWNRTELLTLDPHNVDYLLGLFEPGDMQYELNRNNVTDPSLSEMVVVAIQIL RKNPKGFFLLVEGGRIDHGHHEGKAKQALHEAVEMDRAIGQAGSLTSSEDTLT VVTADHSHVFTFGGYTPRGNSIFGLAPMLSDTDKKPFTAILYGNGPGYKVVGG ERENVSMVDYAHNNYQAQSAVPLRHETHGGEDVAVFSKGPMAHLLHGVHEQNY VPHVMAYAACIGANLGHCAPASSAGSLAAGPLLLALALYPLSVLF	12
Mus musculus alkaline phosphatase, liver/bone/kidney (Alpl), transcript variant 2, mRNA, NM_001287172.1	CTAGTGGGTTTGTGTGACAATCACATCTGAAGGCTCTCTTCACTCCAAGATGG CCCTCCTGTCGCCCACTCTGGACTTGGTGGTCACAGCAGTTGGTAGCTTCCTT CTGTTCGTGCTGGCCCTGGGCCTGCTCTGTTTCTTCACCTGTCGCCTGGCCAG GCCACTCAGGATCGGAACGTCAATTAACGTCAATTAACATCTGACGCTGCCCC CCCCCCCCTCTTCCCACCATCTGGGCTCCAGCGAGGGACGAATCTCAGGGTAC ACCATGATCTCACCATTTTTAGTACTGGCCATCGGCACCTGCCTTACCAACTC TTTTGTGCCAGAGAAAGAGAGAGACCCCAGTTACTGGCGACAGCAAGCCCAAG AGACCTTGAAAAATGCCCTGAAACTCCAAAAGCTCAACACCAATGTAGCCAAG AATGTCATCATGTTCCTGGGAGATGGTATGGGCGTCTCCACAGTAACCGCTGC CCGAATCCTTAAGGGCCAGCTACACCACAACACGGGCGAGGAGACCCGGCTGG AGATGGACAAATTCCCCTTTGTGGCCCTCTCCAAGACATATAACACCAACGCT CAGGTCCCTGACAGCGCGGGCACTGCCACTGCCTACTTGTGTGGCGTGAAGGC CAACGAGGGCACAGTGGGAGTGAGCGCAGCCACAGAGCGCACGCGATGCAACA CCACTCAGGGCAATGAGGTCACATCCATCCTGCGCTGGGCCAAGGATGCTGGG AAGTCCGTGGGCATTGTGACTACCACTCGGGTGAACCACGCCACACCCAGTGC AGCCTACGCACACTCGGCCGATCGGGACTGGTACTCGGATAACGAGATGCCAC CAGAGGCTCTGAGCCAGGGCTGCAAGGACATCGCATATCAGCTAATGCACAAT ATCAAGGATATCGACGTGATCATGGGTGGCGGCCGGAAATACATGTACCCGAA GAACAGAACTGATGTGGAATACGAACTGGATGAGAAGGCCAGGGGTACAAGGC TAGATGGCCTGGATCTCATCAGTATTTGGAAGAGCTTTAAACCCAGACACAAG CATTCCCACTATGTCTGGAACCGCACTGAACTGCTGGCCCTTGACCCCTCCAG GGTGGACTACCTCTTAGGTCTCTTTGAGCCCGGGGACATGCAGTATGAATTGA ATCGGAACAACCTGACTGACCCTTCGCTCTCCGAGATGGTGGAGGTGGCCCTC CGGATCCTGACCAAAAACCTCAAAGGCTTCTTCTTGCTGGTGGAAGGAGGCAG GATTGACCACGGACATCATGAGGGTAAGGCCAAGCAGGCTCTGCATGAAGCAG TGGAGATGGACCAGGCCATTGGCAAGGCAGGCGCCATGACATCCCAGAAAGAC ACCTTGACTGTGGTTACTGCTGATCATTCCCACGTTTTCACATTCGGTGGATA CACCCCCCGGGGCAACTCCATCTTTGGTCTGGCTCCCATGGTGAGCGACACGG ACAAGAAGCCCTTCACGGCCATCCTATATGGTAACGGGCCTGGCTACAAGGTG GTGGACGGTGAACGGGAAAATGTCTCCATGGTAGATTACGCTCACAACAACTA CCAGGCCCAGTCCGCTGTTCCCCTGCGCCATGAGACCCACGGTGGAGAAGACG TGGCGGTCTTTGCCAAGGGCCCGATGGCACACCTGCTTCACGGCGTCCATGAG CAGAACTACATTCCCCATGTGATGGCGTATGCCTCCTGCATTGGGGCCAACCT TGACCACTGTGCCTGGGCCGGCTCTGGGAGCGCACCCTCCCCAGGGGCCCTGC TGCTTCCACTGGCTGTGCTCTCCCTACGCACCCTGTTCTGAGGGTGCAGGTCC CACAAGCCCGCAATGGACAGCCAGCTCCCCTCCTTTTGTGGCCCACCACCGGG CAGCCCACACTCAAGGGAGAGGTCCAGGCAACTTCCAGCAGGAACAGAAGTTC GCTATCTGCCTTGCCTGTATCTGGAATCCTCCATGGGCCAGATTCCTGGCTCT GCCTTTATTCCCTAGTTATTGCCCTTTGGCCAGCAGGTTTCTCTCTTGGGCAG GCAAGACACAGACTGCACAGATTCCCAAAGCACCTTATTTTTCTACCAAATAT ATTCTCCAGACCCTGCAACCTCCATGGAACATTCCAGATCTGACCTTCTCTCC TCCATCCCTTCCCTTCCCTCTGGAACACTGGGCCCCATAGTCACGGCCAGTCC TCAAGCCCAACCCTCCCTGGGGGGAAGACCAGGTCTGCTCAGGATGAGACTCC CAGGAAGCCACCTCCGGGGTTGGCTGTCTACCCAGGGTTGCCAAGCTGGGAAG AACACTCCAGCCGGACAGGACACACACACACACTCCCCACCCAATTGCAGAGA CTCGCCAACCCTTCACTGAAGTGGCTCTCCTGTTTGGAATAGCGGGGTGGGGT GGGGGAGAAGAAAGAAAGAAAGAAAAAAAATTTTTAATTTCTCTTTTTGGTGT TGGTTAAAAGGGAACACAAGACATTTAAATAAAACATCCCAAATATTTCTGAG GCCAG	13

Mus musculus alkaline phosphatase, tissue-nonspecific isozyme, protein (signal peptide underlined), NP_001274101.1	**MISPFLVLAIGTCLTNS**FVPEKERDPSYWRQQAQETLKNALKLQKLNTNVAKN VIMFLGDGMGVSTVTAARILKGQLHHNTGEETRLEMDKFPFVALSKTYNTNAQ VPDSAGTATAYLCGVKANEGTVGVSAATERTRCNTTQGNEVTSILRWAKDAGK SVGIVTTTRVNHATPSAAYAHSADRDWYSDNEMPPEALSQGCKDIAYQLMHNI KDIDVIMGGGRKYMYPKNRTDVEYELDEKARGTRLDGLDLISIWKSFKPRHKH SHYVWNRTELLALDPSRVDYLLGLFEPGDMQYELNRNNLTDPSLSEMVEVALR ILTKNLKGFFLLVEGGRIDHGHHEGKAKQALHEAVEMDQAIGKAGAMTSQKDT LTVVTADHSHVFTFGGYTPRGNSIFGLAPMVSDTDKKPFTAILYGNGPGYKVV DGERENVSMVDYAHNNYQAQSAVPLRHETHGGEDVAVFAKGPMAHLLHGVHEQ NYIPHVMAYASCIGANLDHCAWAGSGSAPSPGALLLPLAVLSLRTLF	14
Macaca fascicularis alkaline phosphatase, biomineralization associated (ALPL), transcript variant X1, mRNA, XM_005544525.3	GCGTTGCGCTCCCGCCACTCCGCGCCCGCGATCCCGGCTCTGCGCTCCCACGC GCTTGTGCCTGGACGGACCCTCGTCAGTGCTCTGCGCAGGATTGGAACATCAG TTAACATCTGACCACTGCCAGCCCACCCCCTCCCACCCGCGTCGATCGCATCT CTGGGCTTCAGGGATAAAGCAGGTCTTGGGGTGCACCATGATTTCACCATTCT TAGTACTGGCCATTGGCACCTGCCTTACCAACTCCTTAGTGCCAGAGAAAGAG AAAGACCCCAAGTACTGGCGAGACCAAGCGCAAGAGACACTGAAATATGCCCT GGAGCTTCAGAAGCTCAACACCAATGTGGCTAAGAATGTCATCATGTTCCTGG GAGATGGGATGGGCGTCTCCACAGTGACGGCCACCCGCATCCTCAAGGGTCAG CTCCACCACAACCCTGGGGAGGAGACCAGGCTGGAGATGGACAAGTTCCCCTT CGTGGCCCTCTCCAAGACGTACAACACCAATGCCCAGGTCCCTGACAGTGCCG GCACCGCCACCGCCTACCTGTGTGGGGTGAAGGCCAACGAGGGCACCGTGGGG GTAAGCGCAGCCACCGAGCGTTCCCGGTGCAACACCACCCAGGGGAACGAGGT CACCTCCATCCTGCGCTGGGCCAAGGACGCTGGGAAATCTGTGGGCATTGTAA CCACCACAAGAGTGAACCATGCCACCCCCAGCGCCGCCTATGCCCACTCAGCT GACCGGGACTGGTACTCAGACAACGAGATGCCCCCTGAGGCCTTGAGCCAGGG CTGCAAGGACATCGCCTACCAGCTTGTGCATAACATCAGGGACATTGACGTGA TCATGGGGGGTGGCCGGAAATACATGTACCCCAAGAATAAAACTGATGTGGAG TATGAGATTGACGAGAAAGCCAGGGGCACGAGGCTGGACGGCCTGGACCTCGT TAACATCTGGAAGAGCTTCAAACCGAGACACAAGCACTCCCACTTCATCTGGA ACCGCACGGAACTCCTGACCCTTGACCCCCACAATGTGGACTACCTATTGGGT CTCTTTGAGCCGGGGGACATGGAGTACGAGCTGAACAGGAACAACGTGACGGA CCCGTCACTCTCCGAGATGGTGGTGGTGGCCATCCAGATCCTGCGGAAGAACC CCAAAGGCTTCTTCTTGCTGGTGGAAGGAGGCAGGATCGACCACGGGCACCAT GAAGGCAAAGCCAAGCAGGCCCTGCACGAGGCGGTAGAGATGGACCGGGCCAT CGGGCAGGCAGGCAGCATGACCTCCTTGGAAGACACTCTGACCGTGGTCACCG CGGACCATTCCCACGTCTTCACCTTTGGTGGATACACCCCCCGTGGCAACTCT ATCTTTGGTCTGGCCCCCATGCTGAGTGACACAGACAAGAAGCCCTTCACTGC CATCCTGTATGGCAATGGGCCTGGCTACAAGGTGGTGGGCGGTGAACGAGAGA ATGTCTCCATGGTGGACTATGCTCACAACAACTACCAGGCGCAGTCTGCTGTG CCCCTGCGCCACGAGACCCACGGCGGGGAGGATGTGGCCGTCTTCTCCAAGGG CCCCATGGCACACCTGCTGCACGGCGTCCATGAGCAGAACTACATCCCCCACG TGATGGCGTACGCAGCCTGCATCGGGGCCAACCTCGACCACTGTGCCCCTGCC AGCTCGGCAGGCAGCCTTGCTGCAGGCCCCCTGCTGCTCCCCCTGGCCCTCTT CCCCCTGAGCATCCTGTTCTGAGGGCCCAGGGCCCGGGCACCCACGAGCCCGT GACACGCCAACTTCCCACTCCCCAGTGCTGCCCACCGCCCGGCAGCCCACCCC GCAAGGGGCAGGGAGGTGGGGGCCTCCTCAGCCTCTGCAACTGCGAGAAAGGG GACCCAGGAAACCAAAGTCTGCCGCCCACCTCGCTCCCCTCTGGAATCTTCCC CGAGGGCCAAACCCACTTCTGGCCTCCAGCCTTTGCTCCCTCCCCGCTGCCCT TTGGCCAACAGGGTAGATTTCTCTTGGGCAAGCAGAGAGTACAGACTGCAGAA ATTCTCAAAGCCTCTTATTTTTCTAGCAAACATATTTCTCCAGACCCAGAGGC CCTGAAGCCTCCATGGAACATTCCGGATCTGACCCTCCCACTCTCATCTCCTT CCCTCTAGAACCCCCCAGGCCCTACCATGCTCATGTCCCTGTCCTCAGGCCCA GCCCTTCTTCAGGGGAGATGAGGTCTTTCTCCTCAGGACAAGGCCTCGCTCAC TCACTCCAAGGCCACCGGGGTCCCAGGAAGCTGGTGCCTGGGTGGCCATCCTA CCCGGCGTGGCCCAGGCCAGGAAGAGCCACCTGGCAGGGCTCACACTCCTGGG CTCTGAACACGCATGCCAGCTCCTCTCTGAAGCGATTCTCCCATTTGGAACGG CAAAAAAAAATTTTTTTCTCTTTTTGGTGGTGGTTAAAAGGGAACACAAAACA TTTAAATAAAACTTTCCAAATATTTCTGAGGACA	15
Macaca fascicularis, alkaline	**MISPFLVLAIGTCLTNS**LVPEKEKDPKYWRDQAQETLKYALELQKLNTNVAKN VIMFLGDGMGVSTVTATRILKGQLHHNPGEETRLEMDKFPFVALSKTYNTNAQ VPDSAGTATAYLCGVKANEGTVGVSAATERSRCNTTQGNEVTSILRWAKDAGK	16

phosphatase, tissue-nonspecific isozyme (signal peptide underlined), XP_005544582.1	SVGIVTTTRVNHATPSAAYAHSADRDWYSDNEMPPEALSQGCKDIAYQLVHNI RDIDVIMGGGRKYMYPKNKTDVEYEIDEKARGTRLDGLDLVNIWKSFKPRHKH SHFIWNRTELLTLDPHNVDYLLGLFEPGDMEYELNRNNVTDPSLSEMVVVAIQ ILRKNPKGFFLLVEGGRIDHGHHEGKAKQALHEAVEMDRAIGQAGSMTSLEDT LTVVTADHSHVFTFGGYTPRGNSIFGLAPMLSDTDKKPFTAILYGNGPGYKVV GGERENVSMVDYAHNNYQAQSAVPLRHETHGGEDVAVFSKGPMAHLLHGVHEQ NYIPHVMAYAACIGANLDHCAPASSAGSLAAGPLLLPLALFPLSILF	

[0035] In some embodiments, the GPI anchor protein described herein is CD59, LY6E, CA4, GPC5, NTM, HYAL2, LSAMP, BST2, EMP2, ALPL, CPM, NCAM1, EFNA1, PIBF1, SEC24B, PRNP, TFPI, OPCML, CD109, DPM3, CNTN4, PIGN, HBP1, CNTN2, CD55, NEGR1, EFNA5, RECK, NRN1, CNTN1, GPAA1, PGAP1, PIGF, PIGK, MDGA2, DPM1, SVIP, NTNG1, CNTN5, GPC6, PIGG, TMEM8A, THY1, GPIHBP1, PIGT, PIGL, ZFAND2B, PLAUR, DPM2, or GPC1.

[0036] In some embodiments, the ligand is or comprises a peptide, a protein, an antibody molecule, a nucleic acid molecule (e.g., an aptamer), or a small molecule.

[0037] In some embodiments, the ligand is not a component of a viral particle, e.g., an AAV viral particle. In some embodiments, the ligand is not a component of a capsid protein, e.g., an AAV capsid protein.

[0038] In some embodiments, the ligand is covalently attached, e.g., directly or indirectly via a linker, to an active agent described herein (e.g., a therapeutic agent or a diagnostic agent). In some embodiments, the ligand is conjugated, e.g., directly or indirectly via a linker, to an active agent described herein (e.g., a therapeutic agent or a diagnostic agent). In some embodiments, the ligand is fused to the active agent, e.g., as part of a fusion peptide or protein.

[0039] In some embodiments, the ligand is conjugated directly to an active agent described herein. In some embodiments, direct conjugation includes but is not limited to formation of a covalent bond between a reactive group on the ligand and a corresponding group or acceptor on the active agent; modification (e.g., genetic modification) of the ligand or active agent to be conjugated to a reactive group (e.g., a sulfhydryl group or a carboxyl group) that forms a covalent attachment to the other molecule to be conjugated under appropriate conditions. For example, a desired active group may be introduced to the ligand, active agent, or both and a disulfide bond may be formed.

[0040] In some embodiments, the ligand is coupled or fused, e.g., conjugated, to the active agent non-covalently, e.g., by a hydrophobic bond, an electrostatic interaction, and/or an ionic bond.

[0041] In some embodiments, the ligand is conjugated to the ligand by a linker. In some embodiments, the linker is a cleavable linker (e.g., an acid-labile linker, peptidase-sensitive linker, photolabile linker, dimethyl linker or disulfide-containing linker). In some embodiments, the linker is a non-cleavable linker. In some embodiments, the linker is an enzyme sensitive linker or a pH sensitive linker. In some embodiments, the pH sensitive linker comprises a hydrazine/hydrazone linker or a disulfide linker. In some embodiments, the enzyme sensitive linker comprises a peptide based linker, e.g., a peptide linker sensitive to a protease (e.g., a lysosomal protease); or a beta-

glucuronide linker. In some embodiments, the non-cleavable linker is a linker comprising a thioether group or a maleimidocaproyl group. In some embodiments, the linker is a chemical linker. In some embodiments, the linker is a peptide linker, e.g., a flexible polypeptide. In some embodiments, the linker is a glycine serine linker. In some embodiments, the linker is a cross-linker, e.g., a cross-linker selected from BMPS, EMCS, GMBS, HBVS, LC-SM CC, MBS, MPBH, SBAP, SIA, SIAB, SMCC, SMPB, SMPH, sulfo-EMCS, sulfo-GMBS, sulfoKMUS, sulfo-MBS, sulfo-SIAB, sulfo-SMCC, and sulfo-SMPB, or SVSB (succinimidyl(4-vinylsulfone)benzoate).

[0042] In some embodiments, a ligand may be conjugated to an active agent described herein using a bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidom ethyl) cyclohexane-l-carboxylate (SMCC), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate H), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and bis-active fluorine compounds (such as l,5-difluoro-2, 4-dinitrobenzene).

[0043] In some embodiments, the ligand and the active agent are fused or coupled post-translationally, e.g., using click chemistry. In some embodiments, the ligand and the active agent are fused or couple via chemically induced dimerization.

[0044] In some embodiments, a ligand may be conjugated to an active agent described herein using a method described in Shadish JA and DeForest CA, Site-Selective Protein Modification: From Functionalized Proteins to Functional Biomaterials. Matter 2020 2:50-70; Fu et al. Antibody drug conjugate: the "biological missile" for targeted cancer therapy. Signal Transduction and Targeted Therapy 2022 7:93; and Drago et al. Unlocking the potential of antibody-drug conjugates for cancer therapy. Nat Rev Clin Oncol 2021 18:327-344; Eyford et al. A Nanomule Peptide Carrier Delivers siRNA Across the Intact Blood Brain Barrier to Attenuate Ischemic Stroke. Front Mol Biosci 2021 8:611367; A microfluidic method for synthesis of transferrin-lipid nanoparticle loaded with siRNA LOR-1284 for therapy of acute myeloid leukemia. Nanoscale 2014 6(16):9742-9751; or US20220125823A1; which are all hereby incorporated by reference in their entirety.

[0045] In some embodiments, the ligand is present N-terminal relative to the active agent. In some embodiments the ligand is present C-terminal relative to the active agent. In some embodiments, the ligand is fused or coupled at or near the C-terminus of the active agent, wherein the active agent is a therapeutic protein, enzyme, or antibody molecule. In some embodiments, the ligand is fused or coupled within 20, 30, 40, 50, 60, 70, 80, 90, 100, or more amino acids from the C-terminus of the therapeutic protein, enzyme, or antibody molecule.

[0046] In some embodiments, binding to ALPL results in one or both of increased cell signaling and/or transcytosis. In some embodiments, binding to ALPL results in increased crossing of the blood brain barrier, e.g., as compared to a reference sequence of SEQ ID NO: 138.

Peptides

[0047] Disclosed herein are ligands comprising peptides or proteins, for binding a protein on cell, e.g., a cell present in the blood brain barrier. In some embodiments, the protein is a GPI anchored protein. In some embodiments, the protein is ALPL, e.g., human or murine ALPL. In some, embodiments, the peptide is an isolated, e.g., recombinant, peptide. In some embodiments, the nucleic acid encoding the peptide, is an isolated, e.g., recombinant nucleic acid.

[0048] The present disclosure also provides peptides and associated AAV particles comprising an AAV capsid variant and a peptide for enhanced or improved transduction of a target cell or tissue (e.g., a cell or tissue of the CNS). In some embodiments, the peptide may increase distribution of an AAV particle to a cell, region, or tissue of the CNS. The cell of the CNS may be, but is not limited to, neurons (e.g., excitatory, inhibitory, motor, sensory, autonomic, sympathetic, parasympathetic, Purkinje, Betz, etc.), glial cells (e.g., microglia, astrocytes, oligodendrocytes) and/or supporting cells of the brain such as immune cells (e.g., T cells). The tissue of the CNS may be, but is not limited to, the cortex (e.g., frontal, parietal, occipital, temporal), thalamus, hypothalamus, striatum, putamen, caudate nucleus, hippocampus, entorhinal cortex, basal ganglia, or deep cerebellar nuclei. In some embodiments, the peptide may increase distribution of an AAV particle to the CNS (e.g., the cortex) after intravenous administration.

[0049] In some embodiments, a peptide of a ligand described herein may vary in length. In some embodiments, the peptide is about 3 to about 20 amino acids in length. As non-limiting examples, the peptide may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 3-5, 3-8, 3-10, 3-12, 3-15, 3-18, 3-20, 5-10, 5-15, 5-20, 10-12, 10-15, 10-20, 12-20, or 15-20 amino acids in length. In some embodiments, a peptide comprises about 6 to 12 amino acids in length, e.g., about 9 amino acids in length. In some embodiments, a peptide comprises about 5 to 10 amino acids in length, e.g., about 7 amino acids in length. In some embodiments, a peptide comprises about 7 to 11 amino acids in length, e.g., about 8 amino acids in length. In some embodiments, a peptide comprises about 4 to 9 amino acids in length, e.g., about 6 amino acids in length.

[0050] In some embodiments, a ligand described herein comprises a protein or a peptide, which comprises a sequence as set forth in Table 1 (e.g., comprising the amino acid sequence of any one of SEQ ID NOs: 200-940, 1800-2241, 2242-2886, or 2887-3076). In some embodiments, the peptide may comprise a sequence as set forth in Table 2A, 2B, or 2C. In some embodiments, the peptide may comprise a sequence set forth in Table 13 or 14. In some embodiments, the peptide may comprise a sequence as set forth in Table 15. In some embodiments, the peptide may comprise a sequence as set forth in Table 16. In some embodiments, the peptide may comprise a sequence as set forth in Table

17. In some embodiments, the peptide may comprise a sequence as set forth in Table 18. In some embodiments, the peptide may comprise a sequence as set forth in Table 19. In some embodiments, the peptide is isolated, e.g., recombinant.

Table 1. Exemplary Peptide Sequences

Peptide Sequence	SEQ ID NO:	Peptide Sequence	SEQ ID NO:	Peptide Sequence	SEQ ID NO:	Peptide Sequence	SEQ ID NO:
GSGSPHSKAQN QQT	200	GSLHHDNHGQN QQT	385	GSVFGVPSG QNQQT	570	GSIAMTSHG QNQQT	755
GHDSPHKSGQN QQT	201	GIMARDSSGQN QQT	386	GSGLPDRNL QNQQT	571	GSPGVSPSG QNQQT	756
GSGSPHARMQN QQT	202	GVVHITNSGQN QQT	387	GSGTHNSAIQ NQQT	572	GSGQNQQTG SSSRV	757
GSGSPHVKSQN QQT	203	GSGQNQHSAPF NQT	388	GSGMIIASMQ NQQT	573	GSGQHLPLL GNQQT	758
GQDSPHKSGQN QQT	204	GSGQTSGLKQN QQT	389	GGITWTDSG QNQQT	574	GSDHSHRGG QNQQT	759
GSGSPHASRQN QQT	205	GSGQNQQTSLS NTA	390	GSGQNQQAS GRQQT	575	GSGIVTKLG QNQQT	760
GSGSPHASRQN KQT	206	GSGQNQAVHNK SQT	391	GSGQNQQPH LKSLT	576	GSGQDVTKT GNQQT	761
GSGSPHVKIQNQ QT	207	GVHTHLPSGQN QQT	392	GPPQHMTSG QNQQT	577	GSGQNQQSH GRIGT	762
GSGSPHSKAKN QQT	208	GHLTMHNSGQN HQT	393	GSGQNQQAS LPSRT	578	GSGQNQQIN HRSPT	763
GSGSPHKKNQN QQT	209	GSGSSSRPYQNQ QT	394	GSGQIVSTQT NQQT	579	GSGDDSRVG QNQQT	764
GSGSPHVRMQN QQT	210	GILLATPSGQNQ QT	395	GSGKGHSAG QNQQT	580	GSGQSTLKRI NQQT	765
GSGSPHASRQK QQT	211	GSGQNAGSFPN QQT	396	GSGQNTRLQ LGQQT	581	GSGSQHSKA QNQQT	766
GHSSPHRSGQN QQT	212	GSRDGHTVGQN QQT	397	GSVGSRPVG QNQQT	582	GSGQNQQHA SSNNT	767
GMRTYHLSGQN QQT	213	GSLLISTSGQNQ QT	398	GSSHTLALG QNQQT	583	GSRTYQVSG QNQQT	768
GSGSPHTRGQN QQT	214	GSGAMPSHGQN QQT	399	GMYEYSQSG QNQQT	584	GSGQNQGLL SSPQT	769
GSGIIPVSSQNQ QT	215	GALVSPISGQNQ QT	400	GNGQNQQHS ILHGT	585	GSGGGLQHN QNQQT	770
GSEYGHKSGQN QQT	216	GSLSSHGVGQN QQT	401	GSGYNQPHL QNQQT	586	GSGQNQQTT AATRM	771
GRGQNVSSVHR QQT	217	GSGQNQQASLA MRT	402	GPLVNASSG QNQQT	587	GSGQNQRAS ILVQT	772
GSSHRFYGGQN QQT	218	GPGLGSHSGQN QQT	403	GSGQNQQVL TTART	588	GSGQNLGLL GAQQT	773
GYFVAAWSGQN QQT	219	GHDSQHKSGQN QQT	404	GSGQNQHSV HNDQT	589	GSLDLGRSG QNQQT	774
GSVLHSHAAQN QQT	220	GSGLTLSATQN QQT	405	GAGLIMHSG QNQQT	590	GNSQVKVSG QNQQT	775
GSGDLVVSTQN QQT	221	GSGQVVAHVGN QQT	406	GMGRHSASG QNQQT	591	GSSGSHQYG QNQQT	776
GSYGMAASGQN QQT	222	GSGLRTMTTQN QQT	407	GSHSQSGHG QNQQT	592	GSGQNQQQR DGTLT	777
GLNHFGASGQN QQT	223	GSGQVGRLLQN QQT	408	GSSTTIVSGQ NHQT	593	GRGQHVSVA NNQQT	778

GSTGSHSAGQNQHT	224	GSGQLSHQSVNQQT	409	GRHLVTASGQNQQT	594	GDSSSRISGQNQQT	779
GLAGHTVSGQNQQT	225	GSGDRYQTLQNQQT	410	GSGQNQQHANLNQT	595	GSGQNQQHSLSSQT	780
GIILGASSGQNQQT	226	GSGQNQQLKSSAQT	411	GSGSTHSKAQNQQT	596	GSLMDVHRGQNQQT	781
GSGVSTYNIQNQQT	227	GSGQNQYSIPVAQT	412	GSGQNKQMLSGNTT	597	GSIQYQSSGQNQQT	782
GSLVSVQTGQNQQT	228	GSGERLHLTQNQQT	413	GSGQVHNPTQNQQT	598	GLGSKNPSGQNQQT	783
GQSSPHRSGQNQQT	229	GSGHNQQVRTAPNT	414	GSGQNQQIPHVHQT	599	GSGQLVLTLQNQQT	784
GREYGHKSGQNQQT	230	GGLSHVMSGQNQQT	415	GSLHAGLSGQNQQT	600	GSGQNQQTSQPLPG	785
GHTLTLSSGQNQQT	231	GSGQSHRDVLNQQT	416	GPAQHGTSGQNQQT	601	GSGQNQQNLGKLNT	786
GSITLIPSGQNQQT	232	GSGQNLAGRMDQQT	417	GEKAVTSSGQNQQT	602	GTTAHQPSGQNQQT	787
GSNGFTALGQNQQT	233	GSGQNQQTNRGNPM	418	GSGQNQQTMANGQR	603	GSGQNRAQIGTQQT	788
GSGHSSHSVQNQQT	234	GSGQSYQRDHNQQT	419	GSGSPHSKDQNQQT	604	GSGQYVHVSSNQQT	789
GSGIPQRSGKNQQT	235	GSLLSAGMGQNQHT	420	GSFSMGYGGQNQQT	605	GSGQNQQTAHAFNI	790
GSGDTLHMLQNQQT	236	GSGQNQQTAIYRNI	421	GSGTHLVSLQNQQT	606	GSGQNQRTMVATQT	791
GERHTVLSGQNQQT	237	GSGQNQQTSGTTNC	422	GSGQMQPHVQNQQT	607	GSGQNPIRGAMQQT	792
GSGMPQSHIQNQQT	238	GMTSHSVSGQNQQT	423	GSGQNQQVAGLNNT	608	GSGYVITGSQNQQT	793
GSGQLSGIGGNQQT	239	GSSQSTGYQPNQQT	424	GSSQNQQHDMRLRT	609	GRGPKQSNIQNQQT	794
GSGQNRKPASFAQT	240	GSLKPTTLGQNQQT	425	GPASLPISGQNQQT	610	GSGQNQQTMLGKPC	795
GSGSVSQLGQNQQT	241	GRMFSLGSGQNQQT	426	GSGQNQQPPLATRT	611	GSGQNQQVGSTVRT	796
GSDFLGTHGQNQQT	242	GSGQNQQTALGVKC	427	GSSRVPVSGQNQQT	612	GNVTTQKSGQNQQT	797
GQIVQNPSGQNQQT	243	GAMVSHSSGQNQQT	428	GSGQNQQTNLGHTT	613	GSGNPVSHLQNQQT	798
GSGTQIPSQQNQQT	244	GSGQNQQRNSDSVT	429	GSGQNQQLVSRVQT	614	GSLSHMESGQNQQT	799
GSGQNQQSAREGLT	245	GSGQSMTLHLNQQT	430	GPNSYPVSGQKQQT	615	GRAPTNLSGQNQQT	800
GSGLGMSTGQNQQT	246	GSGQVHQAEVNQQT	431	GHAHYQASGQNQQT	616	GSGQNQQTVMTARA	801
GSGLPVLSGQNQQT	247	GSGQNQSQNHLQQT	432	GSGQALLSTGNQQT	617	GSGMPASRLQNQQT	802
GSGHSIRTDQNQQT	248	GSLLTTASGQNQQT	433	GSGQLPRQMTNQQT	618	GVVRNHQSGQNQQT	803
GSGQSVQTVVNQQT	249	GSGLIRTAAQNQQT	434	GSGFPKSTEQNQQT	619	GSGQNQHSVQVRQT	804
GSGQNRAQSRFQQT	250	GSGQNQQTVSRQST	435	GSRETSLSGQNQQT	620	GSGQNTGHLTMQQT	805
GGGDLGRSSQNQQT	251	GSGQYANHGINQQT	436	GSGQNQQGTGVSHT	621	GSGQNQQYAGKILT	806
GGGTKMDSGQNQQT	252	GSRSTGPSGQNQQT	437	GSRTVPVYGQNQQT	622	GSGNPHVRNQNQQT	807

GSGSPHPSRQNQQT	253	GRGVQQKLQQNQQT	438	GSNAQSAHGQNQQT	623	GSGQNGGSSNRQQT	808
GSGQFTNAGMNQQT	254	GSGQNQQVHLSTGT	439	GAFHLAASGQNQQT	624	GSGQRLSQGVNHQT	809
GGRNGHTVGQNQQT	255	GSGQNQQLSAKSST	440	GSGQYRSSSDNQQT	625	GSGQNAHAKEGQQT	810
GSGFGPQTGQNQQT	256	GSGYKAARPQNQQT	441	GSGQVYISTPNQQT	626	GSSPAPNSGQNQQT	811
GRTDSHTSGQNQQT	257	GSAGISPSGQNQQT	442	GSGVSTQLLQNQQT	627	GLAHKTSSGQNQQT	812
GYEVLGSSGQNQQT	258	GSGQNRAHAFLQQT	443	GSGQLGLSVTNQQT	628	GSGQNQQTPGAHKT	813
GSVHLSVTGQNQQT	259	GSGLSGITMQNQQT	444	GSGSNMRLSQNQQT	629	GSGQNQQSLSGSFT	814
GFMSYKGSGQNQQT	260	GPGSAHSSGQNQQT	445	GSGQNLHSGLPQQT	630	GSGQNQQSTGTSRT	815
GNIAGSVSGQNQQT	261	GSSHTQALGQNQQT	446	GSSHTLALGQNKQT	631	GSGQNQQTVQSNLV	816
GSGSHRDVSQNQQT	262	GSGVHGVSSQNQQT	447	GSGQNQHSLPAHRT	632	GSGQNQQLGSRQCT	817
GGLGSMSSGQNQQT	263	GSSGRDMGGQNQQT	448	GSGQNQGTVYPNQT	633	GSGQNQYLRLELQT	818
GSGHLPQSAQNQQT	264	GERAFPTSGQNQQT	449	GSGQNQQPSLRQST	634	GSGQNQQTSPRLQT	819
GGVLVGGSGQNQQT	265	GGRIVSLSGQNQQT	450	GSGQNARLKDNQQT	635	GSGQNQQTTSSNMT	820
GTHPYTSSGQNQQT	266	GSGQNSYSHTSQQT	451	GHAGSTGSGQNQQT	636	GTASTYNSGQNQQT	821
GSGQNQQLKENRST	267	GLGYPGSSGQNQQT	452	GSGQALSSSGNQQT	637	GSGQNQQTMPQHKI	822
GSGQNQQTSPHNHT	268	GSGPQSHTGQNQQT	453	GSGASESHRQNQQT	638	GSGQSHLHTGNQQT	823
GSGTLYPQSQNQQT	269	GSGQNQQLSRDAST	454	GVGVITSSGQNQQT	639	GVKGVGHSGQNQQT	824
GSGQNQQSNWITKT	270	GSGQILHSVPNQQT	455	GSLYGQSLGQNQQT	640	GSGKVTKQSQNQQT	825
GSGYTSLFLQNQQT	271	GSGFHTDSRQNQQT	456	GSGQMSDVHGNQQT	641	GSGQNQQTALEKSL	826
GSGVMTHVLQNQQT	272	GSGQSHSLATNQQT	457	GSGQNQQHSSKATT	642	GSGYKDTYGQNQQT	827
GSVSDVRAGQNQQT	273	GSGQNQQTLSKPWT	458	GSGQNQQTSVSQQT	643	GSGQNQQSGTFLST	828
GSGQSHMATLNQQT	274	GSGHAAISQQNQQT	459	GSGQKMWKLDNQQT	644	GSGQNTGQHMMQQT	829
GSGLSVHLAQNQQT	275	GSGQNQQQIGGNST	460	GSGQNVSMQVNQQT	645	GSGKNQQRPGLDQT	830
GSGLSHATQQNQQT	276	GGGPMAGSGQNQQT	461	GSGQNQRATLSNQT	646	GSGQSREISLNQQT	831
GSGLSVQSGQNQQT	277	GMRMEYQSGQNQQT	462	GSGQASSKSANQQT	647	GTPTSPSSGQNQQT	832
GSGHMTYREKNQQT	278	GSGQNQQGTLLHQT	463	GSGKNQTPIPKGQT	648	GKPAGGLSGQNQQT	833
GSKGVPTPGQNQQT	279	GSGQNQRSSGGVQT	464	GSGQNQQTRQEGST	649	GSGQNHRSADMQQT	834
GSGLLPLSSQNQQT	280	GSGQNQRGALATQT	465	GASSLATSGQNQQT	650	GSGQNQQTLPSLSL	835
GNGLYAVSGQNQQT	281	GSGTVHAATQNQQT	466	GSGQRGSLTENQQT	651	GSPYMGATGQNQQT	836

GFNGSPSSGQNQQT	282	GSRMTQQFGQNQQT	467	GSEQTRQRGQNQQT	652	GSGHAKAVGQNQQT	837
GSGQIRHSDQNQQT	283	GSSSPGASGQNQQT	468	GSGQNQQTLTASKE	653	GHMKGVTSGQNQQT	838
GGQVAPSSGQNQQT	284	GHPSPHVSGQNQQT	469	GSGTSGKTGKNQQT	654	GSGQNQKILTLDQT	839
GSFSMHTHGQNQQT	285	GSGSHHASRQNQQT	470	GQLVTFTSGQNQQT	655	GSGQNQQTKVGHSA	840
GSGQNQQVIQGSNT	286	GAVGHSYSGQNQQT	471	GSGQNQQSANKILT	656	GIARTTISGQNQQT	841
GRVLHSHAGQNQQT	287	GSRSQYDIGQNQQT	472	GSGQNQQHHSSHTT	657	GSGQNQQTSVGFRT	842
GSGQNQQTSLQDQT	288	GSGQGPQERGNQQT	473	GSGQNQKGMQPNQT	658	GSGQNQQTMIANIR	843
GSGLGRAPVQNQQT	289	GSIAHVGTGQNQQT	474	GSGQLVSGLYNQQT	659	GDMTRSSSGQNQQT	844
GNGFSSASGQNQQT	290	GSGQNQQKQNHGNT	475	GSSVGVPSGQNQQT	660	GSGHMSDLRQNQQT	845
GSGQMASRESNQQT	291	GSGQNQQALGSQRT	476	GSGQNQQWDSRRQT	661	GRGAVMASGQNQQT	846
GPGLPNHSGQNQQT	292	GSGAITHMPQNQQT	477	GSEQTRQSGQNQQT	662	GSGQNQQLSGKSVT	847
GNIQWQGSGQNQQT	293	GSGQRNPLLLNQQT	478	GSGIGSHIPQNQQT	663	GSHTLVVSGQNQQT	848
GMSAHMSSGQNQQT	294	GSSGIPVSHQNQQT	479	GSGQNQRLHGVDQT	664	GSGPWSAGLQNQQT	849
GHSFVNRSGQNQQT	295	GVHSLTPSGQNQQT	480	GEVSRVLSGQNQQT	665	GSGQHSPHALNQQT	850
GRAVMDHSGQNQQT	296	GVIVLHGSGQNQQT	481	GSGQNQQKVSPLLT	666	GSGQNQQPNSGSMT	851
GALTVMQSGQNQQT	297	GGTRVVDSGQNQQT	482	GSGLALERSQNQQT	667	GSGLAHLGGQNQQT	852
GSGQRSPVLPNQQT	298	GSGGVTYQSQNQQT	483	GPDRIGSSGQNQQT	668	GSSVRYEPKQNQQT	853
GSGQNGHLSLKQQT	299	GSGQNQAGHGPGQT	484	GSGQNQDHQNKQQT	669	GSGQNQQARPLELT	854
GSLPRGTSDQNQQT	300	GSGQLVTSGPNQQT	485	GSGQNQQTALYNNT	670	GSGQPRSTGINQQT	855
GVAGSLVSGQNQQT	301	GSGIAAQRTQNQQT	486	GSGAVHLTAQNQQT	671	GSGQNQANWVKVQT	856
GRGGIPQSGQNQQT	302	GSTPAGVGGQNQQT	487	GSLVSTQSGQNQQT	672	GSGHLFQSGQNQQT	857
GSGQYASSIPNQQT	303	GSGQNQQTSTGVHS	488	GSGVSARMVQNQQT	673	GSGQNRGISISQQT	858
GTDFGRQSSQNQQT	304	GSGQIRQLVDNQQT	489	GSGQTRMPLANQQT	674	GSGTHYDNRQNQQT	859
GIFMQTPSGQNQQT	305	GSLIGMQSGQNQQT	490	GSGISSRNMQNQQT	675	GSGQNQQTSTTPLP	860
GSGQNQQTRLVDLT	306	GSGQIKGKMDNQQT	491	GSGEKVHSGQNQQT	676	GSGQVHASQVNQKT	861
GTREMPLSGQNQQT	307	GSGSDMSSWQNQQT	492	GSGQNQQKLSSMST	677	GSSGHRESGQNQQT	862
GSRLVHVHGQNQQT	308	GRGQNQQHTGLATT	493	GSGQNQQTGQHMRV	678	GLSAEKSSGQNQQT	863
GSGRLVPNGPNQQT	309	GSGQNQQTLYSSNT	494	GSGMIHTTAQNQQT	679	GSGQEHRSLANQQT	864
GSGYLRESPQNQQT	310	GSGQTQVLKSNQQT	495	GSGQNWPALKGQQT	680	GSGQTVVRIANQQT	865

GARIQNASGQK QQT	311	GSRTLSNVGQN QQT	496	GASHMSISG QNQQT	681	GSGQNVSSV HRQQT	866
GLSNPMPSGQN QQT	312	GSGVQHSLPQN QQT	497	GSDQNQQLG YSKQT	682	GSGASRMSI QNQQT	867
GSTVQDTRGQN QQT	313	GNYLHQASGQN QQT	498	GIPSIRESGQ NQQT	683	GVAFIGSSGQ NQQT	868
GPFGMPSSGQN QQT	314	GSGGTSVHQQN QQT	499	GSGIPSVKFQ NQQT	684	GSGQNQQTV PTRQT	869
GSGQNHGVLSN QQT	315	GMDHSRPSGQN QQT	500	GSGQNQQTS VSQNV	685	GSGQAAKSS QNQQT	870
GSGYSMSQAQN QQT	316	GSGQNQQSMGT FTT	501	GSGQNQQIG ESRMT	686	GSGQNQQVA IRTST	871
GSGMLTHTLQN QQT	317	GSGQNQQTPLR PPT	502	GSGSSSMSFQ NQQT	687	GSVHMQNA GQNQQT	872
GRGSPHASRQN QQT	318	GSGQNQHHSVS QQT	503	GSGQKQERA VSKQT	688	GSGMRQAG VQNQQT	873
GLSWPSTSGQN QQT	319	GSGQLRSLSTNQ QT	504	GCTTRLNSG QNQQT	689	GSGQNQQVG GKTVT	874
GNSMERTSGQN QQT	320	GSGSPRQLSQN QQT	505	GSGQNQQIIS TKIT	690	GVHDMRVS GQNQQT	875
GSGMSPSTLQN QQT	321	GSGQNQQTTAS SHT	506	GSGQNQQKS LNGNT	691	GSGQHVSVA NNQQT	876
GSGHGQVLSQN QQT	322	GRGQVVSTHQN QQT	507	GSGIPAPRLQ NQQT	692	GSAAMSVRG QNQQT	877
GRGQIYSTGGN QQT	323	GSAQVSMVGQN QQT	508	GSGQIRESM GNQQT	693	GVSRGGPSG QNQQT	878
GVVAAHNSGQN QQT	324	GSSTLVTIGKNQ QT	509	GSGQNSGVH FNQQT	694	GSGQMVHTI GNQQT	879
GDSSLRHSGQN QQT	325	GFAHQASSGQN QQT	510	GSGQNIGHSL PQQT	695	GRGGSMAET QNQQT	880
GSLVSQGAGQN QQT	326	GSGQPVLSISNQ QT	511	GSGERSISVQ NQQT	696	GSGHTNPTR QNQQT	881
GSLLQAHSGQN QQT	327	GSGQSHRSELN QQT	512	GSGLKPNVL QNQQT	697	GSGEAARYE QNQQT	882
GSGHIYVGIQNQ QT	328	GSSVGSPIGQNQ QT	513	GSGQVAYAQ GNQQT	698	GSGQNERHL VLQQT	883
GHHTTVQSGQN QQT	329	GSGMPIRNVQN QQT	514	GSGQSSYGS GNQQT	699	GSGQNQQSK QQVLT	884
GSRQSKRNELN QQT	330	GSSTRVDSGQN QQT	515	GSGQNQAMT HGDQT	700	GSGQARAHR GNQQT	885
GSGQNQQHVSS PRT	331	GSGQNQQTAMR STT	516	GSGQNQALV SMGQT	701	GSGQNQQPL DTSRT	886
GSSKELLWGQN QQT	332	GSGQNQQHSSS HLT	517	GSGQNPSFM RGQQT	702	GSGQNQQLA NMVTT	887
GSLSTPSSGQNQ QT	333	GSRNGHAVGQN QQT	518	GSGQNQQSH LRTNT	703	GSGQMKDL HRNQQT	888
GSIGYAGQGQN QQT	334	GLGAYQSSGQN QQT	519	GYTRLETSG QNQQT	704	GSGQNQHLS SFVQT	889
GSGQNQRVSNS QQT	335	GPGLSGHSGQN QQT	520	GSGQSYDMR GNQQT	705	GSGQNQQPS SRVTT	890
GSGYASHVQQN QQT	336	GSTGIVSSGQNQ QT	521	GSRTTQDIGQ NQQT	706	GSGQNQQLA ITLGT	891
GSGEYSRSGQN QQT	337	GSRTTQVIGQNQ QT	522	GSGHPYKAA QNQQT	707	GSGQNQQTV GNPAT	892
GSVSTHSSGQN QQT	338	GSGLLHRAQQN QQT	523	GRLSNAHGG QNQQT	708	GSGQNQGRA HPMQT	893
GSGQNQHSLGN YQT	339	GSGQNAQQAAA QQT	524	GSGQNQRAV LNDQT	709	GSGQLIASV VNQQT	894

GSGGLDTRGQNQQT	340	GSGQNQQSALRTQT	525	GGSHTYGGGQNQQT	710	GSSVRSLVGQNQQT	895
GNILHATSGQNQQT	341	GSGFLSDTRQNQQT	526	GSSVNSMIGQNQQT	711	GGAGSAHSGQNQQT	896
GSGQSYTMTQNQQT	342	GSGLLYHDQQNQQT	527	GNSSMMGSGQNQQT	712	GSDQNQQTMSSTRT	897
GSGQNQHSAPNSQT	343	GSGQNQHYSLHKQT	528	GNRDRPSSGQNQQT	713	GSGQNQQMAGAFRT	898
GSGQNQQTMDHNRT	344	GSGHSPLPQQNQQT	529	GSGNMHASRQNQQT	714	GSLGNLQRGQNQQT	899
GSNGGVGTGQNQQT	345	GNGHSMRPNQNQQT	530	GFIFPKVSGQNQQT	715	GSGPSISHGQNQQT	900
GAGSIIPSGQNQQT	346	GSGLKWSTLQNQQT	531	GSGQNQQLKNSTST	716	GSGQNQQSSFNVQT	901
GSGQTHGGQHNQQT	347	GSGQMGRQAVNQQT	532	GSGQNQQSQYMPRT	717	GSGQNQQTGQATHN	902
GSNLSFQSGQNQQT	348	GSGQNQQTSGVLTL	533	GSGQRMADIGNQQT	718	VSGSPHSKAQNQQT	903
GATLQVHSGQNQQT	349	GSGQNQQALHNPHT	534	GSGQNQSHYPSQQT	719	CSGSPHSKAQNQQT	904
GSGFNQRSEQNQQT	350	GSGQNQQVIPNSKT	535	GSDGKMHRGQNQQT	720	GSGSPHRKAQNQQT	905
GSGSLRDFDQNQQT	351	GSPLQDRVGQNQQT	536	GSGSVGFIGQNQQT	721	GRGSPHSKAQNQQT	906
GSGDSITGKQNQQT	352	GSGQNQYSSTNPQT	537	GLHGMTLSGQNQQT	722	GSGSPHSKAQNKQT	907
GSGQDRNIVQNQQT	353	GAMTVTISGQNQQT	538	GSDQSKRGDSNQQT	723	GSGSPHSKAQTQQT	908
GSGLSHSHQQNQQT	354	GSGQNQQLQTLIRT	539	GSLFLATGGQNQQT	724	GSGSTHASRQNQQT	909
GSGQNQQTGMSSVK	355	GSGLRQTSQQNQQT	540	GSGQNQQPSAFSKT	725	GSGSPHKYGQNQQT	910
GSVTHGISGQNQQT	356	GSGQNQQTGLRQQT	541	GSGQLPQSGLNQQT	726	GSGSPHKFGQNQQT	911
GVVAHQPSGQNQQT	357	GSGQTRQMKDNQQT	542	GSGSKQNALQNQQT	727	VSGSPHKFGQNQQT	912
GSGPILGQLQNQQT	358	GSGQNHGLQSGQQT	543	GSGQRRELSQNQQT	728	GSGSPHSKAQNHQT	913
GSGHVPNSGLNQQT	359	GSGQSHRQPENQQT	544	GSGQREPKASNQQT	729	GSGSPHSKAQHQQT	914
GDAGVRSSGQNQQT	360	GSGQDRHIVQNQQT	545	GSGQNQQHPSTQQT	730	GSGSPHKTYQNQQT	915
GSGSQLMSLQNQQT	361	GSGQNQQLPHSNLT	546	GSQSTLGLGQNQQT	731	VSGSPHASRQNQQT	916
GSGLDYSQRQNQQT	362	GSGQLSVPYDNQQT	547	GSGQNQQMPGLSST	732	GSGSPHKFGKNQQT	917
GSGQSSGRLINKQT	363	GSGRNPQTQPLQQT	548	GSGQNQQTVGGKNL	733	GSGSPHASRQNQHT	918
GSSVSPSSGQNQQT	364	GSGQPYSTGLNQQT	549	GSSREFHSGQNQQT	734	GSHSPHKSGQNQQT	919
GSGQVVGLSGNQQT	365	GSGQNQQTHGGLRD	550	GSGQNQQTVPSNLV	735	GSGQNQQRRMSPST	920
GSNMGVPLGQNQQT	366	GAYGMVSSGQNQQT	551	GSGQNAYSSQAQQT	736	GSGSPHSKPQNQQT	921
GSFYPSSTGQNQQT	367	GSGIQSSYSQNQQT	552	GSGQNKDHSTRRQT	737	GSGSPHKFGQKQQT	922
GSGQNQQTRLTDLT	368	GPRLSDQSGQNQQT	553	GQLGSVGSGQDQQT	738	VSGSPHGARQNQQT	923

GPTNGRSSGQNQQT	369	GSGQNQQTHPSPCT	554	GSGQHAAPGHNQQT	739	GSGSPHSKAQKQQT	924
GSGLLHGKLQNQQT	370	GSGQSFQMHTNQQT	555	GSGQNQQTSQSPPT	740	GSHSPHKRGQNQQT	925
GANMGHVSGQNQQT	371	GSGQNQQTGNPKHT	556	GSGNYRDHEQNQQT	741	GSGQNRQRLKGLET	926
GSGQNQQSGRGDLT	372	GFSSAVHSGQNQQT	557	GSGQHSNQHVNQQT	742	GSGSPHKLGQNQQT	927
GSHGHYASGQKQQT	373	GSGQNQQTSMSNAT	558	GSGQTARNGINQQT	743	GSGSPHKTSKNQQT	928
GSGDLRISPQNQQT	374	GSGQDMKQHHNQQT	559	GSGQNQQHYGSQGT	744	GSGSPHKIGQNQQT	929
GSGMPVILGQNQQT	375	GLRLSTPSGQNQQT	560	GSGSPQASRQNQQT	745	GSGQDSPHVRNQQT	930
GRGVITSSGQNHQT	376	GSGQNQQTSVYMNT	561	GSGFSHSMGKNQQT	746	GSGSPHKTSQNQQT	931
GSGHSVSGPQNQQT	377	GSGQNQYSQSSMQT	562	GSGQSHSLETNQQT	747	GSGSPHASRKNQQT	932
GSRNGHTVGRNQQT	378	GSGQNQQSMADHTT	563	GTEQTRQSGQNQQT	748	GSHSPHKSGQKQQT	933
GAGVHMVSGQNQQT	379	GWERSFVSGQNQQT	564	GSGRHLASVQNQQT	749	GSGSPHKTSQKQQT	934
GSGQNHRPSVLQQT	380	GLLAGKSSGQNQQT	565	GLGSKNHSGQNQQT	750	GSGSPHVRGQNKQT	935
GSGSPRDSIQNQQT	381	GKSFVPQSGQNQQT	566	GSGQNQQTSHFPSA	751	GSGSPHKTTQNQQT	936
GSGQGIHSSVNQQT	382	GSGQMQSAGSNQQT	567	GSGQLSGTPQNQQT	752	GSGPVRALRQNQQT	937
GSGQQLSITPNQQT	383	GSDQNQRLTSSMQT	568	GSGQNQQAPHKKET	753	GSGSPHVRGQKQQT	938
GGYHSQTSGQNQQT	384	GESRAVLSGQNQQT	569	GSGQNQQTLRGSLE	754	CSGSPHKTSQNQQT	939
CSHSPHKSGQNQQT	940	DAGSPHSKAQNQQ	1909	GSGSPHASRQNQQ	2019	GNDSPHKSVQNQQ	2129
GHDSPHKSGQNQQ	1800	GSGSPHSKGQNQQ	1910	GSGSPHASRQNKQ	2020	GHDSPHKSAQNYQ	2130
GSGSPHSKAQNQQ	1801	DGGSPHSKAQNQQ	1911	GSGSPHVKIQNQQ	2021	GSASPHSKALNQQ	2131
GSGSPHSKAQNRH	1802	ASGSPHSKAHNQQ	1912	GSGSPHSKAKNQQ	2022	GHESPHKSAQNRQ	2132
GSGSPHSKVQNQQ	1803	GSGSPHSKAQNTY	1913	GSGSPHKKNQNQQ	2023	GQDSPHKIGQNQQ	2133
MSGSPHSKAQNQQ	1804	GSGSPHSKSQNQH	1914	GSGSPHVRMQNQQ	2024	GHDSPHKSGQNHL	2134
GRGSPHSKAQNQQ	1805	GGGSPHSKAQDKQ	1915	GSGSPHASRQKQQ	2025	GHDSPHKSGQYQH	2135
RNGSPHSKAQNQQ	1806	GSGSPHSKAQNHL	1916	GHSSPHRSGQNQQ	2026	GNDSPHKSVQNHQ	2136
GSGSPHSKARDQQ	1807	GSGSPHSKAQIGM	1917	GSGSPHTRGQNQQ	2027	GHDSPHKSGQNQW	2137
GSGSPHSKAPNLQ	1808	GSGSPHSKALNKQ	1918	CSGSPHSKAQNQQ	2028	GHDSPHKSVQNQH	2138
TSGSPHSKAQNQQ	1809	GGGSPHSKAQNPQ	1919	GSGSPHRKAQNQQ	2029	GHDSPHKSGQNQH	2139
GSGSPHSKAHVRQ	1810	GTGSPHSKAPNQL	1920	GSGSPHSKAQNKQ	2030	GHDSPHKSGQTRQ	2140
GSGSPHSKAPNQH	1811	GSGSPHSKAQLQQ	1921	GSGSPHSKAQTQQ	2031	GHDSPHKSGQNLH	2141

ISGSPHSKAQNQQ	1812	GGGSPHSKAQYQQ	1922	CSGSPHKTSQNQQ	2032	GHDSPHKSAQNQE	2142
GPGSPHSKAHNQQ	1813	GGGSPHSKAQHQQ	1923	CSHSPHKSGQNQQ	2033	GHDSPHKSGQHLQ	2143
GSGSPHSKTQSQQ	1814	GSGSPHSKAQRMS	1924	GQSSPHRSGQNQQ	2034	GHDSPHKSRLNQP	2144
ESGSPHSKAQNQQ	1815	GSGSPHSKAQGIL	1925	GRGSPHASRQNQQ	2035	GQDSPHKSGQNQD	2145
GSGSPHSKAQPAK	1816	GSGSPHSKAQDRQ	1926	GSGSPHASRKNQQ	2036	GHDSPHKSGRNQQ	2146
SSGSPHSKAQNQQ	1817	GSGSPHSKARDWQ	1927	GSGSPHASRQNQH	2037	GHDSPHKSGQNLL	2147
GNGSPHSKAQNQQ	1818	GSGSPHSKAQNTH	1928	GSGSPHKFGKNQQ	2038	GHDSPHKSGQLVI	2148
GSGSPHSKSQTQQ	1819	GSGSPHSKAQERS	1929	GSGSPHKFGQKQQ	2039	GHDSPHKSRQSQQ	2149
ASGSPHSKAQNQQ	1820	GSGSPHSKAQNYQ	1930	GSGSPHKFGQNQQ	2040	GHDSPHKSGRTQE	2150
GSGSPHSKAQNLA	1821	GSGSPHSKAQRTC	1931	GSGSPHKIGQNQQ	2041	GHDSPHKSVQTHQ	2151
GSGSPHSKSQNQL	1822	GSGSPHSKAQIGH	1932	GSGSPHKLGQNQQ	2042	GHDSPHKSGQNQP	2152
NSGSPHSKAQNQQ	1823	GSGSPHSKAQGAI	1933	GSGSPHKTSKNQQ	2043	GHDSPHKSGQTQQ	2153
GSGSPHSKAQGQQ	1824	GSGSPHSKAQVPP	1934	GSGSPHKTSQKQQ	2044	GPDSPHKIGQNQQ	2154
VSGSPHSKAQNQQ	1825	GSGSPHSKAQVQQ	1935	GSGSPHKTSQNQQ	2045	GHDSPHKSVQNQQ	2155
GSGSPHSKALNRQ	1826	GSGSPHSKALMRQ	1936	GSGSPHKTTQNQQ	2046	GHDSPHKSRQDQH	2156
LSGSPHSKAQNQQ	1827	GSGSPHSKAQYSV	1937	GSGSPHKTYQNQQ	2047	GPDSPHKSGQKQQ	2157
GSGSPHSKAHNQQ	1828	GSGSPHSKVPNLQ	1938	GSGSPHKYGQNQQ	2048	GHDSPHKSRQSQH	2158
GSGSPHSKTQNQQ	1829	GSGSPHSKAQAIT	1939	GSGSPHSKAQHQQ	2049	GHDSPHKSVQNQL	2159
GGGSPHSKAQTQQ	1830	GSGSPHSKAQKTL	1940	GSGSPHSKDQNQQ	2050	GYDSPHKSGQYQH	2160
GSGSPHSKAQNPP	1831	GSGSPHSKAQNQW	1941	GSGSPHSKPQNQQ	2051	GHDSPHKSRQNQQ	2161
GSGSPHSKAQNLQ	1832	GSGSPHSKAQLHH	1942	GSGSPHVRGQKQQ	2052	GHDSPHKSWVRQQ	2162
GGGSPHSKAQNQQ	1833	GSGSPHSKAQNII	1943	GSGSPHVRGQNKQ	2053	GHESPHKSGQNQH	2163
GSGSPHSKAQYQQ	1834	MEGSPHSKAQNQQ	1944	GSHSPHKRGQNQQ	2054	GHDSPHKIGHNQQ	2164
GGGSPHSKAQNKQ	1835	GSGSPHSKAQGHH	1945	GSHSPHKSGQKQQ	2055	GHDSPHKSNAWQQ	2165
GSGSPHSKAQDQE	1836	GSGSPHSKAQSKV	1946	GSHSPHKSGQNQQ	2056	GHDSPHKSGQSVP	2166
KSGSPHSKAQNQQ	1837	GSGSPHSKAQLPS	1947	VSGSPHASRQNQQ	2057	GHESPHKSGQNIQ	2167
GGGSPHSKAQNQL	1838	GSGSPHSKAIGKQ	1948	VSGSPHGARQNQQ	2058	GHDSPHKSVQNHL	2168
GSGSPHSKAQNHQ	1839	GGGSPHSKSQNQQ	1949	VSGSPHKFGQNQQ	2059	GHDSPHKIGLDQQ	2169
GSGSPHSKAQDQQ	1840	GSGSPHSKAQAIH	1950	GSGSPHSKAQYYV	2060	ASGSPHSKAQHQQ	2170

GGGSPHSKSQNQL	1841	GSGSPHSKAQHGL	1951	GSGSPHSKLRRQQ	2061	GHDSPHKRGPDQQ	2171
GNGSPHSKAQNKQ	1842	GSGSPHSKAQFMC	1952	GSGSPHSKAGCGQ	2062	GMGSPHSKTQNQQ	2172
GSGSPHSKGHWQQ	1843	VSGSPHSKAQGQQ	1953	GSGSPHSRAQNQQ	2063	GHDSPHKSGESQQ	2173
GSGSPHSKAPNQQ	1844	GGGSPHSKAQNQM	1954	GSGSPHSKRLRQQ	2064	GHDSPHKHGQNHQ	2174
GSGSPHSKAQNQL	1845	GSGSPHSKAQHLQ	1955	GSGSPHSLRRNQQ	2065	GTGSPHSKAQNQL	2175
GSGSPHSKRPEQQ	1846	ENGSPHSKAQNQQ	1956	GSGSPHSRGRNQQ	2066	GHDSPHKSVQNKQ	2176
GSGSPHSKAQRTM	1847	GSGSPHSKTQNHQ	1957	GSGSPHSSRRNQQ	2067	GQVSPHKSGQNQQ	2177
GNGSPHSKAQNQH	1848	GSGSPHSKAQPAR	1958	GSGSPHSKAFRLQ	2068	GHDSPHKSGQRQL	2178
HSGSPHSKAQNQQ	1849	GSGSPHSKAQSLQ	1959	GSCSPHRKAQNQQ	2069	GHDSPHKIGQNQL	2179
GGGSPHSKALNQQ	1850	GSGSPHSKSQSQL	1960	GSGSPHFLRQNQQ	2070	GHDSPHKSGQIIV	2180
GSGSPHSKALHQH	1851	GSASPHSKAHSQQ	1961	GSGSPHSLRFNQQ	2071	GYDSPHKSGQKQS	2181
GTGSPHSKAQNHQ	1852	GSGSPHSKAQMPS	1962	GSGSPHSKWLLQQ	2072	GNGSPHSKAQNQE	2182
GSGSPHSKAQHRI	1853	GSGSPHSKAQGSL	1963	GSGSPHSKRRLQQ	2073	GDDSPHKSVQNQQ	2183
GSGSPHSKAQYIH	1854	GSGSPHSKSQNQQ	1964	GSGSPHSKAQRKL	2074	GHDSPHKSVQSHQ	2184
GGGSPHSKAHNQQ	1855	GNGSPHSKSQNQQ	1965	GSGSPHSKALRRQ	2075	GHDSPHKSGQFVV	2185
GSGSPHSKAQKFE	1856	GSGSPHSKAQVPA	1966	GSGSPHSKAQRLR	2076	GHDSPHKSRQNLQ	2186
ESGSPHSKAQNHQ	1857	GNGSPHSKAQNLQ	1967	GSGSPHSKAQRRL	2077	GHNSPHKSGQNQE	2187
GSGSPHSKAQFPS	1858	GSGSPHSKAQDKQ	1968	GSGSPHSKARRQQ	2078	GHDSPHKSGQSQP	2188
PSGSPHSKAQNQQ	1859	GSGSPHSKAHYQQ	1969	GSGSPHSKARRLQ	2079	GHESPHKSGQNEQ	2189
GNGSPHSKAQNPL	1860	GSGSPHSKAQVPS	1970	GSGSPHSKSRRQQ	2080	GHDSPHKSGQNQL	2190
GGGSPHSKAQSQQ	1861	GGGSPHSKAQNHQ	1971	GLLSPHWKAQNQQ	2081	GHDSPHKSAQNLL	2191
GSGSPHSKAQAIK	1862	GSGSPHSKARGEQ	1972	GSGSPHSKARLRQ	2082	ASGSPHSKAINQQ	2192
GSGSPHSKGQNRQ	1863	GGGSPHSKAQYQH	1973	GSGSPHSKASKRQ	2083	GNGSPHKRGQNQQ	2193
GSGSPHSKAQSQQ	1864	GSGSPHSKAPGQQ	1974	GSGSPHVRRQNQQ	2084	GHDSPHKSGQSLQ	2194
GSVSPHGKAQNQL	1865	KNGSPHSKAQNQQ	1975	GSGSPHSKAQLYR	2085	GHDSPHKSAQNHQ	2195
ASGSPHSKAQNQL	1866	GSGSPHSKRLEQQ	1976	GSGSPHSKAQLTV	2086	GHDSPHKSGRNRQ	2196
RSGSPHSKAQNQQ	1867	GSGSPHSKAQNQS	1977	GHDSPHKRGQHRQ	2087	GHDSPHKYGQNEQ	2197
GSGSPHSKAQYQH	1868	GSGSPHSKAQKVI	1978	GHDSPHKSGQKQQ	2088	GNGSPHSKAPNLQ	2198
GSGSPHTKAQNPQ	1869	GSGSPHSKAQNND	1979	GHDSPHKSGLTQQ	2089	GHDSPHKSQQNQQ	2199

GSGSPHSKGQNPP	1870	GSGSPHSKAQSVH	1980	GDDSPHKSGRNQQ	2090	GHDSPHKSVQSKQ	2200
GSGSPHSKAQHQL	1871	GSGSPHSKAQPLG	1981	GHDSPHKSGLNQQ	2091	GNDSPHKIGHNQQ	2201
GSGSPHSKAQSPP	1872	KEGSPHSKAQNQQ	1982	GHESPHKSAQNHQ	2092	GGGSPHSKAQDQQ	2202
GSGSPHSKAQAKL	1873	GSGSPHSKAHNQE	1983	GHDSPHKSAQNQW	2093	GQDSPHKSGQNPL	2203
GSGSPHSKTKSQQ	1874	GSGSPHSKAQIQQ	1984	GHDSPHKSGQNTH	2094	ASGSPHSKAQNHQ	2204
GSGSPHSKAQDRP	1875	GSGSPHSKAQVRN	1985	GHDSPHKSGRRRQ	2095	GHDSPHKSGRDQK	2205
GIGSPHSKAQNLG	1876	GSGSPHSKAPSNQ	1986	GHDSPHKSAQNQQ	2096	GHDSPHKSVHNQQ	2206
GSGSPHSKAQAFH	1877	GSGSPHSKAQVGH	1987	GHDSPHKSGQYQQ	2097	GHDSPHKSGQWKR	2207
GSGSPHSKAQKQQ	1878	GSGSPHSKAQRDI	1988	GNYSPHKIGQNQQ	2098	GSGSPHSKAENRQ	2208
GSGSPHSKAQNAQ	1879	GSGSPHSKAQMPN	1989	GHDSPHKSRQNDQ	2099	GHDSPHKSGQSQQ	2209
WSGSPHSKAQNQQ	1880	AIGSPHSKAQNQQ	1990	GHDSPHKSGQIRQ	2100	GHDSPHKSRQAQQ	2210
GSGSPHSKAHNQL	1881	GSGSPHSKARGLQ	1991	GHDSPHKIGQNQH	2101	GHDSPHKSVQNHQ	2211
GNGSPHSKAQNHQ	1882	GSGSPHSKLQKQQ	1992	GYDSPHKSGQKQQ	2102	GHDSPHKSKQNQQ	2212
GGGSPHSKAQNLQ	1883	GSGSPHSKAPSLQ	1993	GHDSPHKSGQSVQ	2103	GHDSPHKSAQNQL	2213
GSGSPHSKAQKLN	1884	GSGSPHSKAQRDQ	1994	GHESPHKSGRSQQ	2104	GHDSPHKSGQTQP	2214
GGGSPHSKSQNQH	1885	GSGSPHSKNRDQQ	1995	GHDSPHKSGQNKL	2105	GHDSPHKLWINQQ	2215
GSGSPHSKSQNVQ	1886	GSGSPHSKAQAKG	1996	GHDSPHKTGQNQQ	2106	GPDSPHKSGQNQQ	2216
GSGSPHSKAQAQQ	1887	GSGSPHSKAQSAH	1997	GRGSPHKRGQNQQ	2107	GHDSPHKSVQKQL	2217
DSGSPHSKAQNQQ	1888	GNGSPHSKSQNQH	1998	GSGSPHTKAQNPP	2108	GHPSPHWKGQNQQ	2218
ASGSPHSKAPNQQ	1889	GSGSPHSKSQNHQ	1999	GQDSPHKSGQHQQ	2109	GHDSPHKSGRNQL	2219
GSGSPHSKAQTPP	1890	RSGSPHSKAQDQQ	2000	GHDSPHKSGQIQH	2110	GSGSPHSKVQDQQ	2220
IDGSPHSKAQNQQ	1891	GSGSPHSKAQSTM	2001	GHDSPHKSGPRQQ	2111	GHDSPHKMGRNQQ	2221
GSGSPHNKAQNHQ	1892	GSGSPHSKAQREM	2002	GHDSPHKSGHTQQ	2112	GHDSPHKSGISIQ	2222
GSGSPHSKAQPPA	1893	GGGSPHSKSQNRQ	2003	GHDSPHKSGQRQH	2113	GHDSPHKSVQNLQ	2223
GSGSPHSKAQERP	1894	GSGSPHSKAQYRA	2004	GSGSPHTKAQNQQ	2114	GHDSPHKMAHNQQ	2224
GSGSPHSKAQDLQ	1895	GGGSPHSKAQRQQ	2005	GHDSPHKSAQSQQ	2115	GHDSPHKHGQNQQ	2225
GGGSPHSKAQNPP	1896	GSGSPHSKNQWQQ	2006	GHESPHKSGQNQQ	2116	GHDSPHKSVQSQQ	2226
GSGSPHSKAQAMH	1897	GSGSPHSKAQRMN	2007	GHDSPHKSLQNQQ	2117	GHDSPHKSGQTVC	2227
GSGSPHSKALNQQ	1898	GSGSPHAKAQNHQ	2008	GHGSPHSKAQNPQ	2118	GQDSPHKSGQYQQ	2228

GSGSPHSKAQHPS	1899	GSGSPHSKAGDSQ	2009	GHDSPHKSGRNQE	2119	GHDSPHKSGQQIM	2229
GLGSPHSKSQNQQ	1900	GSGSPHSKLKSQQ	2010	GHDSPHKSGQTQL	2120	GHDSPHKSRQNEQ	2230
GTGSPHSKAQNQQ	1901	GSGSPHSKAQKIS	2011	GHDSPHKSEKNQQ	2121	GHDSPHKSGLNHQ	2231
GSGSPHSKAPGLQ	1902	GSGSPHSKAPSMQ	2012	GRDSPHKSGQDQQ	2122	GYDSPHKSGQNQQ	2232
GSGSPHSKAQGIR	1903	GSGSPHSKASPRQ	2013	GHDSPHKTGHNQQ	2123	GHDSPHKSGQNLQ	2233
GSGSPHSKAQAPA	1904	GSGSPHSKRMEQQ	2014	GYDSPHKSGQTQQ	2124	GHDSPHKSRQDQQ	2234
GSGSPHSKSQSQQ	1905	GSGSPHSKAQYQN	2015	GHESPHKSGQTQQ	2125	GDDSPHKSGQKQL	2235
GSGSPHSKAQIPP	1906	GSGSPHARMQNQQ	2016	GHDSPHKSGQSKQ	2126	GSGSPHSKAQNQA	2236
GSGSPHSKAQTQL	1907	GSGSPHVKSQNQQ	2017	GHDSPHKTGQNQP	2127	GDDSPHKSGHNQQ	2237
GSGSPHSKAQAPS	1908	GQDSPHKSGQNQQ	2018	GHDSPHKSGQSPQ	2128	GHDSPHKSGQMIH	2238
GHDSPHKSGRNHQ	2239	GHDSPHKSVQNRQ	2240	GHDSPHKSGQKMN	2241	TINGHDSPHKSRLNQP	2728
TINGSGSPHSKAQNQQ	2242	TDRGSGSPHSKAQNQQ	2404	TINGSGSPHSKAQSTM	2566	TVDGHDSPHKSGQKQQ	2729
TINGHDSPHKSGQNQQ	2243	TINGSGSPHSKAQIPP	2405	TVNASGSPHSKAQNQL	2567	TINGQDSPHKSGQNQD	2730
TIIGSGSPHSKAQNRH	2244	TVKGSGSPHSKAQDQQ	2406	TINGSGSPHSKAQREM	2568	TIEGHDSPHKSGRNQQ	2731
TFPGSGSPHSKVQNQQ	2245	NADGSGSPHSKAQNQQ	2407	TVHGSGSPHSKAQSQQ	2569	TTNGHDSPHKSGQNLL	2732
TEKMSGSPHSKAQNQQ	2246	TDKVSGSPHSKAQNQQ	2408	TINGGGSPHSKSQNRQ	2570	TINGHDSPHKSGQLVI	2733
EINGRGSPHSKAQNQQ	2247	TITGSGSPHSKAQTQL	2409	TINGSGSPHSKAQYRA	2571	TVNGHDSPHKSRQSQQ	2734
TVNRNGSPHSKAQNQQ	2248	TINGSGSPHSKAQAPS	2410	TINGGGSPHSKAQRQQ	2572	TINGHDSPHKSGRTQE	2735
TVNGSGSPHSKARDQQ	2249	NCVGSGSPHSKAQNQQ	2411	TEPMSGSPHSKAQNQQ	2573	TINGHDSPHKSVQTHQ	2736
TFNGSGSPHSKAPNLQ	2250	TIRDAGSPHSKAQNQQ	2412	TINGSGSPHSKNQWQQ	2574	TSNGHDSPHKSGQNQP	2737
TEKTSGSPHSKAQNQQ	2251	TVKDSGSPHSKAQNQQ	2413	ETAGSGSPHSKAQNQQ	2575	VINGHDSPHKSGQTQQ	2738
TINGSGSPHSKAHVRQ	2252	NALGSGSPHSKAQNQQ	2414	TINGSGSPHSKAQRMN	2576	TINGPDSPHKIGQNQQ	2739
TVNGSGSPHSKAPNQH	2253	VINGSGSPHSKGQNQQ	2415	NNLGSGSPHSKAQNQQ	2577	AVNGHDSPHKSVQNQQ	2740
TEKISGSPHSKAQNQQ	2254	TVNGGGSPHSKAQNQQ	2416	TINGSGSPHAKAQNHQ	2578	TINGHDSPHKSRQDQH	2741
TINGPGSPHSKAHNQQ	2255	TIQDGGSPHSKAQNQQ	2417	TIIKNGSPHSKAQNQQ	2579	AINGPDSPHKSGQKQQ	2742
TVNGSGSPHSKTQSQQ	2256	TISGGGSPHSKAQNQQ	2418	TINGSGSPHSKAGDSQ	2580	TINGHDSPHKSRQSQH	2743
SINESGSPHSKAQNQQ	2257	TSNASGSPHSKAHNQQ	2419	TINGSGSPHSKLKSQQ	2581	TIYGHDSPHKSVQNQL	2744
TERTSGSPHSKAQNQQ	2258	TINGSGSPHSKAQNTY	2420	TINGSGSPHSKAQKIS	2582	TVNGHDSPHKSGQNLL	2745
TINGSGSPHSKAQPAK	2259	TINGSGSPHSKSQNQH	2421	TEYNSGSPHSKAQNQQ	2583	TENKSGSPHSKAQNQQ	2746

TEKSSGSPHSKAQNQQ	2260	TINGGGSPHSKAQDKQ	2422	TINGSGSPHSKAPSMQ	2584	TTNGQDSPHKSGQNQQ	2747
TSYGNGSPHSKAQNQQ	2261	TEFVSGSPHSKAQNQQ	2423	AINGSGSPHSKAQNQQ	2585	TDKGSGSPHSKAQNQQ	2748
TEKGSGSPHSKAQNQQ	2262	TVNGSGSPHSKAQNHL	2424	TINGSGSPHSKASPRQ	2586	TIDGHDSPHKSGRNQQ	2749
TINGSGSPHSKSQTQQ	2263	TREISGSPHSKAQNQQ	2425	TINGSGSPHSKRMEQQ	2587	TINGYDSPHKSGQYQH	2750
TERISGSPHSKAQNQQ	2264	TINGSGSPHSKAQIGM	2426	TINGSGSPHSKAQYQN	2588	TDNGHDSPHKSRQNQQ	2751
TERASGSPHSKAQNQQ	2265	TIDGSGSPHSKALNKQ	2427	TINGSGSPHSKAQYYV	2589	TINGHDSPHKSWVRQQ	2752
ELHGSGSPHSKAQNQQ	2266	TIIGGGSPHSKAQNPQ	2428	TINGSGSPHSKLRRQQ	2590	TINGHESPHKSGQNQH	2753
AINGSGSPHSKAQNLA	2267	QGEGSGSPHSKAQNQQ	2429	TINGSGSPHSKAGCGQ	2591	TVNGHDSPHKIGHNQQ	2754
TVNGSGSPHSKSQNQL	2268	TINGTGSPHSKAPNQL	2430	SMNGSGSPHSRAQNQQ	2592	TCNGHDSPHKSGRNQQ	2755
TERNSGSPHSKAQNQQ	2269	TVNGSGSPHSKAQLQQ	2431	TINGSGSPHSKRLRQQ	2593	TINGNGSPHSKAQNHQ	2756
SVNGNGSPHSKAQNQQ	2270	TFNGGGSPHSKAQYQQ	2432	TINGSGSPHSLRRNQQ	2594	NVVGHDSPHKSGQNQQ	2757
TFNGSGSPHSKAQGQQ	2271	SINGSGSPHSKTQSQQ	2433	TINGSGSPHSRGRNQQ	2595	TINGHDSPHKSNAWQQ	2758
TERVSGSPHSKAQNQQ	2272	TVNGGGSPHSKAQHQQ	2434	TINGSGSPHSSRRNQQ	2596	TDAGHDSPHKSGQNQQ	2759
TINGSGSPHSKALNRQ	2273	SEKGSGSPHSKAQNQQ	2435	TINGSGSPHSKAFRLQ	2597	TEVGHDSPHKSGQNQQ	2760
TERLSGSPHSKAQNQQ	2274	NVNGSGSPHSKAQNQQ	2436	TINGSCSPHRKAQNQQ	2598	SELGHDSPHKSGQNQQ	2761
TDNGSGSPHSKAHNQQ	2275	GGEGSGSPHSKAQNQQ	2437	TINGSGSPHFLRQNQQ	2599	TINGHDSPHKSGQSVP	2762
TFHGSGSPHSKTQNQQ	2276	TINGSGSPHSKAQRMS	2438	TINGSGSPHSLRFNQQ	2600	TINGHESPHKSGQNIQ	2763
TINGGGSPHSKAQTQQ	2277	TINGSGSPHSKAQGIL	2439	TINGSGSPHSKWLLQQ	2601	TINGHDSPHKSVQNHL	2764
TSNGSGSPHSKAQNPP	2278	EFVGSGSPHSKAQNQQ	2440	TINGSGSPHSKRRLQQ	2602	TINGHDSPHKIGLDQQ	2765
TINGSGSPHSKAQNLQ	2279	TIIGSGSPHSKAQDRQ	2441	TINGSGSPHSKAQRKL	2603	TSNASGSPHSKAQHQQ	2766
TVHGNGSPHSKAQNQQ	2280	SDKGSGSPHSKAQNQQ	2442	TINGSGSPHSKALRRQ	2604	TINGHDSPHKRGPDQQ	2767
TINGGGSPHSKAQNQQ	2281	TEQVSGSPHSKAQNQQ	2443	TINGSGSPHSKAQRLR	2605	TINGMGSPHSKTQNQQ	2768
TENMSGSPHSKAQNQQ	2282	TEHVSGSPHSKAQNQQ	2444	YLSGSGSPHSKAQNQQ	2606	TIKGHDSPHKSGESQQ	2769
TENVSGSPHSKAQNQQ	2283	TINGSGSPHSKARDWQ	2445	TINGSGSPHSKAQRRL	2607	TINGHDSPHKHGQNHQ	2770
TSSGSGSPHSKAQYQQ	2284	TENASGSPHSKAQNQQ	2446	TINGSGSPHSKARRQQ	2608	TVNGTGSPHSKAQNQL	2771
TIDGGGSPHSKAQNKQ	2285	EVQGSGSPHSKAQNQQ	2447	TINGSGSPHSKARRLQ	2609	TIIGHDSPHKSGQYQH	2772
TEKVSGSPHSKAQNQQ	2286	TINGSGSPHSKAQNTH	2448	TINGSGSPHSKSRRQQ	2610	TSNGHDSPHKSVQNKQ	2773
AINGSGSPHSKAQDQE	2287	TINGSGSPHSKAPNLQ	2449	TINGLLSPHWKAQNQQ	2611	IVNGQVSPHKSGQNQQ	2774
TCNKSGSPHSKAQNQQ	2288	TINGSGSPHSKAQERS	2450	TINGSGSPHSKARLRQ	2612	TVNGHDSPHKSGQRQL	2775

TINGGGSPHSKAQNQL	2289	TSNGSGSPHSKAQNYQ	2451	TINGSGSPHSKASKRQ	2613	TVNGHDSPHKIGQNQL	2776
NINGGGSPHSKAQNQQ	2290	TEYISGSPHSKAQNQQ	2452	TINGSGSPHVRRQNQQ	2614	TINGHDSPHKSGQIIV	2777
TEHLSGSPHSKAQNQQ	2291	TINGSGSPHSKAQRTC	2453	TINGSGSPHSKAQLYR	2615	IGNGHESPHKSGQNQQ	2778
AEMGSGSPHSKAQNQQ	2292	TINGSGSPHSKAQIGH	2454	GLSGSGSPHSKAQNQQ	2616	EVMGHDSPHKSGQNQQ	2779
ATNGSGSPHSKAQNHQ	2293	NCWGSGSPHSKAQNQQ	2455	TINGSGSPHSKAQLTV	2617	TINGYDSPHKSGQKQS	2780
AIKGSGSPHSKAQDQQ	2294	TINGSGSPHSKAQGAI	2456	TINGHDSPHKRGQHRQ	2618	TIHGNGSPHSKAQNQE	2781
TINGGGSPHSKSQNQL	2295	TDVNSGSPHSKAQNQQ	2457	MPEGHDSPHKSGQNQQ	2619	YQVGHDSPHKSGQNQQ	2782
TVNGNGSPHSKAQNKQ	2296	SDIGSGSPHSKAQNQQ	2458	MEGGHDSPHKSGQNQQ	2620	TIKGDDSPHKSVQNQQ	2783
TINGSGSPHSKGHWQQ	2297	TINGSGSPHSKAQVPP	2459	MEYGHDSPHKSGQNQQ	2621	TINGHDSPHKSVQSHQ	2784
TDKTSGSPHSKAQNQQ	2298	TINGSGSPHSKAQVQQ	2460	AEWGHDSPHKSGQNQQ	2622	TINGHDSPHKSGQFVV	2785
TFKGSGSPHSKAPNQQ	2299	TINGSGSPHSKALMRQ	2461	CEWGHDSPHKSGQNQQ	2623	TVNGHDSPHKSRQNLQ	2786
TVNGSGSPHSKAQNQL	2300	TINGSGSPHSKAQYSV	2462	ANNGQDSPHKSGQNQQ	2624	ATNGHNSPHKSGQNQE	2787
TINGSGSPHSKRPEQQ	2301	NSIGSGSPHSKAQNQQ	2463	IPEGHDSPHKSGQNQQ	2625	AINGHDSPHKSAQNQQ	2788
TINGSGSPHSKAQRTM	2302	TINGSGSPHSKVPNLQ	2464	ADMGHDSPHKSGQNQQ	2626	TEHGHDSPHKSGQNQQ	2789
TEKASGSPHSKAQNQQ	2303	AINGSGSPHSKAQSQQ	2465	IEYGHDSPHKSGQNQQ	2627	TIYGHDSPHKSGQSQP	2790
SDQGSGSPHSKAQNQQ	2304	TINGSGSPHSKAQAIT	2466	ADYGHDSPHKSGQNQQ	2628	TISGHESPHKSGQNEQ	2791
TEITSGSPHSKAQNQQ	2305	TINGSGSPHSKAQKTL	2467	IETGHDSPHKSGQNQQ	2629	AIIGHDSPHKSAQNQQ	2792
TDKSSGSPHSKAQNQQ	2306	TVNGSGSPHSKAQNQW	2468	MEWGHDSPHKSGQNQQ	2630	AIDGHDSPHKSGQNQL	2793
TIDGSGSPHSKAQNQQ	2307	TINGSGSPHSKAQLHH	2469	CEYGHDSPHKSGQNQQ	2631	TIMGHDSPHKSVQNQQ	2794
TVNGNGSPHSKAQNQH	2308	TEQTSGSPHSKAQNQQ	2470	RINGHDSPHKSGQKQQ	2632	EVGGHDSPHKSGQNQQ	2795
NTNGSGSPHSKAQNQQ	2309	TINGSGSPHSKAQNII	2471	MEIGHDSPHKSGQNQQ	2633	TINGHDSPHKSAQNLL	2796
TETHSGSPHSKAQNQQ	2310	NSLGSGSPHSKAQNQQ	2472	LEYGHDSPHKSGQNQQ	2634	TINASGSPHSKAINQQ	2797
TINGGGSPHSKALNQQ	2311	TIPMEGSPHSKAQNQQ	2473	ADWGHDSPHKSGQNQQ	2635	AINGNGSPHKRGQNQQ	2798
TINGSGSPHSKALHQH	2312	TINGSGSPHSKAQGHH	2474	IEIGHDSPHKSGQNQQ	2636	SEMGHDSPHKSGQNQQ	2799
TINGTGSPHSKAQNHQ	2313	TDRTSGSPHSKAQNQQ	2475	DIMGHDSPHKSGQNQQ	2637	AQQGHDSPHKSGQNQQ	2800
TINGSGSPHSKAQHRI	2314	TINGSGSPHSKAQSKV	2476	FEQGHDSPHKSGQNQQ	2638	AINGHDSPHKSGQSLQ	2801
TINGSGSPHSKAQYIH	2315	EVVGSGSPHSKAQNQQ	2477	MEFGHDSPHKSGQNQQ	2639	TINGSGSPHSKAPNQQ	2802
TENISGSPHSKAQNQQ	2316	TINGSGSPHSKAQLPS	2478	CDQGHDSPHKSGQNQQ	2640	CGEGHDSPHKSGQNQQ	2803
TIIGGGSPHSKAHNQQ	2317	TINGSGSPHSKAIGKQ	2479	LPEGHDSPHKSGQNQQ	2641	TVNGHDSPHKSAQNHQ	2804

TINGSGSPHSKA QKFE	2318	TEPTSGSPHSKA QNQQ	2480	IENGHDSPHK SGQNQQ	2642	TVNGHDSPH KSGQTQL	2805
TSNESGSPHSKA QNHQ	2319	TVNGGGSPHSK SQNQQ	2481	MESGHDSPH KSGQNQQ	2643	TNNGHDSPH KSGRNRQ	2806
TINGSGSPHSKA QFPS	2320	TINGSGSPHSKA QAIH	2482	AEIGHDSPHK SGQNQQ	2644	TINGHDSPH KYGQNEQ	2807
TERPSGSPHSKA QNQQ	2321	TINGSGSPHSKA QHGL	2483	VEYGHDSPH KSGQNQQ	2645	TINGNGSPHS KAPNLQ	2808
TINGNGSPHSKA QNPL	2322	SELGSGSPHSKA QNQQ	2484	IINGHDSPHK SGLTQQ	2646	SINGHDSPHK SQQNQQ	2809
SIKGNGSPHSKA QNQQ	2323	TINGSGSPHSKA QFMC	2485	TSNGDDSPH KSGRNQQ	2647	TIGGHDSPH KSGQNQQ	2810
TERMSGSPHSK AQNQQ	2324	TINVSGSPHSKA QGQQ	2486	IEVGHDSPHK SGQNQQ	2648	TINGHDSPH KSVQSKQ	2811
TERSSGSPHSKA QNQQ	2325	TINGGGSPHSKA QNQM	2487	MEMGHDSPH KSGQNQQ	2649	ELVGHDSPH KSGQNQQ	2812
TELHSGSPHSKA QNQQ	2326	TVNGSGSPHSK AQHLQ	2488	AEVGHDSPH KSGQNQQ	2650	ELMGHDSPH KSGQNQQ	2813
TELTSGSPHSKA QNQQ	2327	TIRENGSPHSKA QNQQ	2489	MDAGHDSPH KSGQNQQ	2651	TINGNDSPH KIGHNQQ	2814
TINGSGSPHSKA HNQQ	2328	TINGSGSPHSKT QNHQ	2490	VEWGHDSPH KSGQNQQ	2652	TIKGGGSPHS KAQDQQ	2815
TINGGGSPHSKA QSQQ	2329	TINGSGSPHSKA QPAR	2491	AEQGHDSPH KSGQNQQ	2653	TVNGHDSPH KSGQTQQ	2816
TINGSGSPHSKA QAIK	2330	TVNGSGSPHSK AQSLQ	2492	LEWGHDSPH KSGQNQQ	2654	TINGQDSPH KSGQNPL	2817
TENTSGSPHSKA QNQQ	2331	TINGSGSPHSKS QSQL	2493	MELGHDSPH KSGQNQQ	2655	TVNASGSPH SKAQNHQ	2818
TIDGSGSPHSKG QNRQ	2332	TINGSASPHSKA HSQQ	2494	METGHDSPH KSGQNQQ	2656	TINGHDSPH KSGRDQK	2819
NINGSGSPHSKA QSQQ	2333	TWQNSGSPHSK AQNQQ	2495	MEAGHDSPH KSGQNQQ	2657	TINGHDSPH KSVHNQQ	2820
TINGSVSPHGKA QNQL	2334	TINGSGSPHSKA QDRQ	2496	IESGHDSPHK SGQNQQ	2658	TINGHDSPH KSGQWKR	2821
TSNASGSPHSKA QNQL	2335	TINGSGSPHSKA QMPS	2497	MEVGHDSPH KSGQNQQ	2659	TIDGSGSPHS KAENRQ	2822
TEARSGSPHSKA QNQQ	2336	TNNGGGSPHSK AQNLQ	2498	CEIGHDSPHK SGQNQQ	2660	NEIGHDSPH KSGQNQQ	2823
TEKNSGSPHSKA QNQQ	2337	TINGSGSPHSKA QGSL	2499	ATNGHDSPH KSGLNQQ	2661	AINGHDSPH KSGQSQQ	2824
TANGSGSPHSK AQYQQ	2338	TEVTSGSPHSKA QNQQ	2500	MDGGHDSPH KSGQNQQ	2662	IINGHDSPHK SRQAQQ	2825
TVNGSGSPHSK AQYQH	2339	SINGGGSPHSKA QYQQ	2501	QEVGHDSPH KSGQNQQ	2663	TPNGHDSPH KSGQNQQ	2826
TINGSGSPHTKA QNPQ	2340	TVIGSGSPHSKS QNQQ	2502	ADQGHDSPH KSGQNQQ	2664	ITNGHDSPH KSGQTQQ	2827
TINGSGSPHSKG QNPP	2341	AVNVSGSPHSK AQNQQ	2503	TINGHESPHK SAQNHQ	2665	TINGHDSPH KSVQNHQ	2828
TIIGSGSPHSKAQ HQL	2342	TVNGNGSPHSK SQNQQ	2504	TINGHDSPHK SAQNQW	2666	TINGHDSPH KSKQNQQ	2829
TINGSGSPHSKA QSPP	2343	TDRNSGSPHSK AQNQQ	2505	NMNGHDSPH KSGQNTH	2667	TINGHDSPH KSAQNQL	2830
TIYGSGSPHSKA QNQQ	2344	TINGSGSPHSKA QVPA	2506	IEMGHDSPH KSGQNQQ	2668	TVNGHDSPH KSGQTQP	2831
TINGSGSPHSKA QAKL	2345	GVLGSGSPHSK AQNQQ	2507	TINGHDSPHK SGRRRQ	2669	TDQGHDSPH KSGQNQQ	2832
TDKNSGSPHSK AQNQQ	2346	TLNGNGSPHSK AQNLQ	2508	ISNGHDSPHK SAQNQQ	2670	TINGHDSPH KLWINQQ	2833

TINGSGSPHSKTKSQQ	2347	AINGSGSPHSKAQDKQ	2509	TGNGHDSPHKSGQYQQ	2671	GINGPDSPHKSGQNQQ	2834
TINGSGSPHSKAQDRP	2348	TSNGSGSPHSKAHYQQ	2510	TINGNYSPHKIGQNQQ	2672	SEIGHDSPHKSGQNQQ	2835
TINGIGSPHSKAQNLG	2349	TINGSGSPHSKAQVPS	2511	TINGHDSPHKSRQNDQ	2673	TINGHDSPHKSVQKQL	2836
TINGSGSPHSKAQSQQ	2350	TELRSGSPHSKAQNQQ	2512	QQQGHDSPHKSGQNQQ	2674	TINGHPSPHWKGQNQQ	2837
TENLSGSPHSKAQNQQ	2351	NINGSGSPHSKAQNHQ	2513	HDWGHDSPHKSGQNQQ	2675	TVNGHDSPHKSGRNQL	2838
TINGSGSPHSKAQAFH	2352	TVNGGGSPHSKAQNHQ	2514	IEGGHDSPHKSGQNQQ	2676	TIKGSGSPHSKVQDQQ	2839
TINGSGSPHSKAQKQQ	2353	TINGSGSPHSKARGEQ	2515	TFNRSGSPHSKAQNQQ	2677	SEKGHDSPHKSGQNQQ	2840
TFSGSGSPHSKAQNLQ	2354	TINGGGSPHSKAQYQH	2516	AINGHDSPHKSGQIRQ	2678	WSAGHDSPHKSGQNQQ	2841
AINGSGSPHSKAQNAQ	2355	TEDLSGSPHSKAQNQQ	2517	TINGHDSPHKIGQNQH	2679	ELAGHDSPHKSGQNQQ	2842
TESWSGSPHSKAQNQQ	2356	TINGSGSPHSKAPGQQ	2518	AINGYDSPHKSGQKQQ	2680	TINGHDSPHKMGRNQQ	2843
TTNGSGSPHSKAHNQL	2357	TIPKNGSPHSKAQNQQ	2519	TESGHDSPHKSGQNQQ	2681	TINGHDSPHKSGISIQ	2844
TVNGNGSPHSKAQNHQ	2358	TINGSGSPHSKAQSLQ	2520	TINGHDSPHKSGQSVQ	2682	TSNGHDSPHKSVQNLQ	2845
TEDKSGSPHSKAQNQQ	2359	TINGSGSPHSKRLEQQ	2521	TINGHESPHKSGRSQQ	2683	QTQGHDSPHKSGQNQQ	2846
TESASGSPHSKAQNQQ	2360	TERGSGSPHSKAQNQQ	2522	TINGHDSPHKSGQNKL	2684	TINGHDSPHKMAHNQQ	2847
TNNGSGSPHSKAQNQQ	2361	TVNGSGSPHSKAPNQQ	2523	TINGHDSPHKTGQNQQ	2685	AINGSGSPHSKAQTQQ	2848
TSNGGGSPHSKAQNLQ	2362	TSNGSGSPHSKAQNQS	2524	TINGRGSPHKRGQNQQ	2686	TINGHDSPHKHGQNQQ	2849
TDKMSGSPHSKAQNQQ	2363	TINGSGSPHSKAQKVI	2525	TINGSGSPHTKAQNPP	2687	GADGHDSPHKSGQNQQ	2850
EVHGSGSPHSKAQNQQ	2364	TEGISGSPHSKAQNQQ	2526	TINGQDSPHKSGQHQQ	2688	VGEGHDSPHKSGQNQQ	2851
TINGSGSPHSKAQKLN	2365	TINGSGSPHSKAQNND	2527	SINGHDSPHKSGQIQH	2689	ANEGHDSPHKSGQNQQ	2852
TINGGGSPHSKSQNQH	2366	TINGSGSPHSKAQSVH	2528	AINGHDSPHKSGPRQQ	2690	TEAKSGSPHSKAQNQQ	2853
TVNGGGSPHSKAQSQQ	2367	TINGSGSPHSKAQPLG	2529	TVNGHDSPHKSGHTQQ	2691	TINGHDSPHKSVQSQQ	2854
TTNGSGSPHSKAQYQH	2368	TINKEGSPHSKAQNQQ	2530	SINGHDSPHKSGQRQH	2692	TIPGSGSPHSKAQNLQ	2855
TISGSGSPHSKAQYQH	2369	TCNASGSPHSKAQNQQ	2531	SLNGSGSPHTKAQNQQ	2693	TINGHDSPHKSGQTVC	2856
TESTSGSPHSKAQNQQ	2370	AINGSGSPHSKAHNQE	2532	AINGHDSPHKSAQSQQ	2694	ELRGHDSPHKSGQNQQ	2857
TINGSGSPHSKSQNVQ	2371	TEGLSGSPHSKAQNQQ	2533	SIYGHESPHKSGQNQQ	2695	CQIGHDSPHKSGQNQQ	2858
SINGSGSPHSKAQAQQ	2372	TRDASGSPHSKAQNQQ	2534	TVNGHDSPHKSLQNQQ	2696	GVMGHDSPHKSGQNQQ	2859
TVNGSGSPHSKAQNLQ	2373	TSNGSGSPHSKAQNLQ	2535	TINGHGSPHSKAQNPQ	2697	ACDGHDSPHKSGQNQQ	2860
TVRDSGSPHSKAQNQQ	2374	TGNGSGSPHSKAQIQQ	2536	TSNGYDSPHKSGQKQQ	2698	TINGQDSPHKSGQYQQ	2861
TFNASGSPHSKAPNQQ	2375	TVNGGGSPHSKAQNLQ	2537	TVNGHDSPHKSGRNQE	2699	TINGHDSPHKSGQQIM	2862

TDRMSGSPHSK AQNQQ	2376	TDRSSGSPHSKA QNQQ	2538	TTNGHDSPH KSGQTQL	2700	TINGHDSPH KSRQNEQ	2863
TINGSGSPHSKA QTPP	2377	TINGSGSPHSKA QVRN	2539	AINGHDSPH KSEKNQQ	2701	ASNGHDSPH KSGLNHQ	2864
TIKGSGSPHSKA QNQQ	2378	TINGSGSPHSKA PSNQ	2540	IINGRDSPHK SGQDQQ	2702	TVNGHDSPH KSGQSQP	2865
NHIGSGSPHSKA QNQQ	2379	TINGSGSPHSKA QVGH	2541	TISGHDSPHK TGHNQQ	2703	NELGHDSPH KSGQNQQ	2866
TINGSGSPHSKA QYQH	2380	NAIGSGSPHSKA QNQQ	2542	SINGYDSPHK SGQTQQ	2704	AAEGHDSPH KSGQNQQ	2867
TIPIDGSPHSKAQ NQQ	2381	AENGSGSPHSK AQNQQ	2543	TINGHESPHK SGQTQQ	2705	GQNGHDSPH KSGQNQQ	2868
TINGSGSPHSKA QGQQ	2382	TINGSGSPHSKA QRDI	2544	TINGHDSPHK SGQSKQ	2706	NEFGHDSPH KSGQNQQ	2869
TFNGSGSPHNK AQNHQ	2383	TINGSGSPHSKA QMPN	2545	AIIGHESPHK SGQNQQ	2707	TSIGYDSPHK SGQNQQ	2870
ESDGSGSPHSKA QNQQ	2384	TVNGSGSPHSKS QNQQ	2546	TINGHDSPHK TGQNQP	2708	TDNGHDSPH KSGQNLQ	2871
TINGSGSPHSKA QPPA	2385	TIPAIGSPHSKAQ NQQ	2547	AINGHDSPH KSGQSPQ	2709	TITGHDSPHK SRQDQQ	2872
TINGSGSPHSKA QERP	2386	TINGSGSPHSKA RGLQ	2548	TIKGNDSPHK SVQNQQ	2710	AEHGHDSPH KSGQNQQ	2873
TIKGSGSPHSKA QDLQ	2387	TELGSGSPHSKA QNQQ	2549	TEFGHDSPH KSGQNQQ	2711	TINGDDSPH KSGQKQL	2874
TDLKSGSPHSKA QNQQ	2388	AETGSGSPHSKA QNQQ	2550	TINGHDSPHK SAQNYQ	2712	EILGHDSPHK SGQNQQ	2875
TINGGGSPHSKA QNPP	2389	TINGSGSPHSKL QKQQ	2551	TFNGSASPHS KALNQQ	2713	TIHGSGSPHS KAQNQA	2876
TINGSGSPHSKA QAMH	2390	TINGSGSPHSKA PSLQ	2552	TINGHESPHK SAQNRQ	2714	AINGDDSPH KSGHNQQ	2877
TVPNSGSPHSKA QNQQ	2391	TINGSGSPHSKA QRDQ	2553	TTNGHDSPH KSGQNQQ	2715	TSNGHNSPH KSGQNQE	2878
TVIGSGSPHSKA LNQQ	2392	TDVGSGSPHSK AQNQQ	2554	TIKGQDSPHK IGQNQQ	2716	TINGHDSPH KSGQMIH	2879
TINGSGSPHSKA QHPS	2393	TINGSGSPHSKN RDQQ	2555	TVNGHDSPH KSGQNHL	2717	NAIGHDSPH KSGQNQQ	2880
TINGLGSPHSKS QNQQ	2394	SINGSGSPHSKA PNLQ	2556	SINGHDSPHK SGQYQH	2718	VINGHDSPH KSGRNHQ	2881
TINGTGSPHSKA QNQQ	2395	TINGSGSPHSKA QAKG	2557	TINGNDSPHK SVQNHQ	2719	TITGHDSPHK SVQNRQ	2882
TINGSGSPHSKA PGLQ	2396	TVNGSGSPHSK AQDKQ	2558	TITGHDSPHK SGQNQW	2720	TINGHDSPH KSGQKMN	2883
TINGSGSPHSKA QGIR	2397	TINGGGSPHSKA QNPQ	2559	TNNGHDSPH KSVQNQH	2721	TIHGHDSPH KSGQSQQ	2884
TESHSGSPHSKA QNQQ	2398	TINGSGSPHSKA QSAH	2560	TIDGHDSPHK SGQNQH	2722	TEIGHDSPHK SGQNQQ	2885
TINGSGSPHSKA QAPA	2399	TINGNGSPHSKS QNQH	2561	TVNGHDSPH KSGQTRQ	2723	TINGHDSPH KSGQYQH	2886
TINGSGSPHSKS QSQQ	2400	TVPTSGSPHSKA QNQQ	2562	TVNGHDSPH KSGQNLH	2724	NCLGSGSPH SKAQNQQ	2403
AEHGSGSPHSK AQNQQ	2401	TIDGSGSPHSKS QNHQ	2563	AISGHDSPHK SGLNQQ	2725	AINRSGSPHS KAQDQQ	2565
TEDRSGSPHSKA QNQQ	2402	TDVKSGSPHSK AQNQQ	2564	AINGHDSPH KSAQNQE	2726	TITGHDSPHK SGQHLQ	2727
GSGSPHSKAQN RHT	2887	GSGSPHSKAQS QQT	2936	GSGSPHSKA QRMST	2985	GSGSPHSKA QNNDQ	3034
GSGSPHSKVQN QQT	2888	ASGSPHSKAQN QLT	2937	GSGSPHSKA QGILT	2986	GSGSPHSKA QSVHT	3035

MSGSPHSKAQN QQT	2889	RSGSPHSKAQN QQT	2938	GSGSPHSKA QDRQT	2987	GSGSPHSKA QPLGT	3036
GSGSPHSKARD QQT	2890	GSGSPHSKAQY QHT	2939	GSGSPHSKA RDWQT	2988	GSGSPHSKA HNQET	3037
GSGSPHSKAPNL QT	2891	GSGSPHTKAQN PQS	2940	GSGSPHSKA QNTHD	2989	GSGSPHSKA QNLQI	3038
TSGSPHSKAQN QQT	2892	GSGSPHSKGQN PPT	2941	GSGSPHSKAP NLQI	2990	GSGSPHSKA QIQQT	3039
GSGSPHSKAHV RQT	2893	GSGSPHSKAQH QLT	2942	GSGSPHSKA QERST	2991	GSGSPHSKA QVRNT	3040
GSGSPHSKAPNQ HT	2894	GSGSPHSKAQSP PT	2943	GSGSPHSKA QNYQT	2992	GSGSPHSKA PSNQT	3041
ISGSPHSKAQNQ QT	2895	GSGSPHSKAQA KLT	2944	GSGSPHSKA QRTCT	2993	GSGSPHSKA QVGHT	3042
GSGSPHSKTQSQ QT	2896	GSGSPHSKTKSQ QT	2945	GSGSPHSKA QIGHT	2994	GSGSPHSKA QRDIT	3043
GSGSPHSKAQN QST	3032	GSGSPHSKAQD RPT	2946	GSGSPHSKA QGAIT	2995	GSGSPHSKA QMPNT	3044
ESGSPHSKAQN QQI	2898	GSGSPHSKAQS QQL	2947	GSGSPHSKA QVPPT	2996	GSGSPHSKA RGLQT	3045
GSGSPHSKAQPA KT	2899	GSGSPHSKAQA FHT	2948	GSGSPHSKA QVQQI	2997	GSGSPHSKL QKQQT	3046
SSGSPHSKAQNQ QT	2900	GSGSPHSKAQK QQD	2949	GSGSPHSKAL MRQT	2998	GSGSPHSKA PSLQT	3047
GSGSPHSKSQTQ QN	2901	GSGSPHSKAQN AQT	2950	GSGSPHSKA QYSVT	2999	GSGSPHSKA QRDQT	3048
ASGSPHSKAQN QQT	2902	WSGSPHSKAQN QQT	2951	GSGSPHSKVP NLQT	3000	GSGSPHSKN RDQQT	3049
GSGSPHSKAQN LAT	2903	GSGSPHSKAHN QLT	2952	GSGSPHSKA QSQQI	3001	GSGSPHSKA QAKGT	3050
GSGSPHSKSQNQ LT	2904	GSGSPHSKAQN QQY	2953	GSGSPHSKA QAITT	3002	GSGSPHSKA QSAHT	3051
NSGSPHSKAQN QQT	2905	GSGSPHSKAQK LNT	2954	GSGSPHSKA QKTLT	3003	GSGSPHSKS QNHQT	3052
GSGSPHSKAQG QQT	2906	GSGSPHSKSQN VQT	2955	GSGSPHSKA QNQWT	3004	RSGSPHSKA QDQQT	3053
VSGSPHSKAQN QQT	2907	GSGSPHSKAQA QQT	2956	GSGSPHSKA QLHHT	3005	GSGSPHSKA QSTMT	3054
GSGSPHSKALNR QS	2908	GSGSPHSKAQN LQA	2957	GSGSPHSKA QNIII	3006	GSGSPHSKA QREMT	3055
LSGSPHSKAQN QQT	2909	DSGSPHSKAQN QQT	2958	GSGSPHSKA QGHHT	3007	GSGSPHSKA QYRAT	3056
GSGSPHSKAHN QQT	2910	ASGSPHSKAPN QQT	2959	GSGSPHSKA QSKVT	3008	GSGSPHSKN QWQQT	3057
GSGSPHSKTQN QQT	2911	GSGSPHSKAQTP PT	2960	GSGSPHSKA QLPST	3009	GSGSPHSKA QRMNT	3058
GSGSPHSKAQNP PT	2912	GSGSPHSKAQY QHA	2961	GSGSPHSKAI GKQT	3010	GSGSPHAKA QNHQT	3059
GSGSPHSKAQN LQT	2913	GSGSPHSKAQG QQA	2962	GSGSPHSKA QAIHT	3011	GSGSPHSKA GDSQT	3060
GSGSPHSKAQY QQT	2914	GSGSPHNKAQN HQT	2963	GSGSPHSKA QHGLT	3012	GSGSPHSKL KSQQT	3061
GSGSPHSKAQD QET	2915	GSGSPHSKAQPP AT	2964	GSGSPHSKA QFMCT	3013	GSGSPHSKA QKIST	3062
KSGSPHSKAQN QQT	2916	GSGSPHSKAQE RPT	2965	VSGSPHSKA QGQQT	3014	GSGSPHSKA PSMQT	3063
GSGSPHSKAQN HQT	2917	GSGSPHSKAQD LQT	2966	GSGSPHSKA QHLQT	3015	GSGSPHSKA SPRQT	3064

GSGSPHSKAQDQQT	2918	GSGSPHSKAQAMHT	2967	GSGSPHSKTQNHQN	3016	GSGSPHSKRMEQQT	3065
GSGSPHSKGHWQQT	2919	GSGSPHSKALNQQT	2968	GSGSPHSKAQPART	3017	GSGSPHSKAQYQNT	3066
GSGSPHSKAPNQQT	2920	GSGSPHSKAQHPST	2969	GSGSPHSKAQSLQT	3018	RSGSPHSKAQNQQI	3067
GSGSPHSKAQNQLI	2921	GSGSPHSKAPGLQT	2970	GSGSPHSKSQSQLT	3019	GSGSPHTKAQNPPT	3068
GSGSPHSKRPEQQT	2922	GSGSPHSKAQGIRT	2971	GSGSPHSKAQDRQS	3020	GSGSPHTKAQNQQT	3069
GSGSPHSKAQRTMT	2923	GSGSPHSKAQAPAT	2972	GSGSPHSKAQMPST	3021	ASGSPHSKAQHQQT	3070
GSGSPHSKAQNQQH	2924	GSGSPHSKSQSQQI	2973	GSGSPHSKAQGSLT	3022	ASGSPHSKAINQQT	3071
HSGSPHSKAQNQQT	2925	GSGSPHSKAQIPPT	2974	GSGSPHSKSQNQQT	3023	GSGSPHSKAPNQQH	3072
GSGSPHSKALHQHT	2926	GSGSPHSKAQTQLT	2975	GSGSPHSKAQVPAT	3024	ASGSPHSKAQNHQT	3073
GSGSPHSKAQHRIT	2927	GSGSPHSKAQAPST	2976	GSGSPHSKAQDKQT	3025	GSGSPHSKAENRQT	3074
GSGSPHSKAQYIHT	2928	GSGSPHSKGQNQQT	2977	GSGSPHSKAHYQQT	3026	GSGSPHSKVQDQQT	3075
GSGSPHSKAQKFET	2929	ASGSPHSKAHNQQT	2978	GSGSPHSKAQVPST	3027	GSGSPHSKAQTQQA	3076
ESGSPHSKAQNHQT	2930	GSGSPHSKAQNTYA	2979	GSGSPHSKARGEQT	3028	GSGSPHSKAQNQAT	3077
GSGSPHSKAQFPST	2931	GSGSPHSKSQNQHI	2980	GSGSPHSKAPGQQT	3029	GSGSPHSKGQNRQT	2935
PSGSPHSKAQNQQT	2932	GSGSPHSKAQNHLT	2981	GSGSPHSKAQSLQI	3030	GSGSPHSKAQLQQT	2984
GSGSPHSKAHNQQR	2933	GSGSPHSKAQIGMT	2982	GSGSPHSKRLEQQT	3031	GSGSPHSKAQKVIT	3033
GSGSPHSKAQAIKT	2934	GSGSPHSKALNKQT	2983	GHDSPHKHGQNHQT	3160	GHDSPHKSGQWKRT	3202
GHDSPHKSGQIRQT	3078	GHDSPHKSGQNQHA	3119	LHDSPHKSGQNQQT	3161	GHDSPHKSGQSQQI	3203
GHDSPHKIGQNQHA	3079	GHDSPHKSGQTRQT	3120	GHDSPHKSVQNKQT	3162	GHDSPHKSRQAQQT	3204
GYDSPHKSGQKQQT	3080	GHDSPHKSGQNLHT	3121	GQVSPHKSGQNQQT	3163	GHDSPHKSVQNHQI	3205
GHDSPHKSGQNQQT	3081	GHDSPHKSGLNQQT	3122	GHDSPHKSGQRQLT	3164	GHDSPHKSKQNQQA	3206
GHDSPHKSGQSVQT	3082	GHDSPHKSAQNQET	3123	GHDSPHKIGQNQLT	3165	GHDSPHKSAQNQLN	3207
GHESPHKSGRSQQT	3083	GHDSPHKSGQHLQT	3124	GHDSPHKSGQIIVT	3166	GHDSPHKSGQTQPT	3208
GHDSPHKSGQNKLE	3084	GHDSPHKSRLNQPT	3125	IHDSPHKSGQNQQT	3167	FHDSPHKSGQNQQT	3209
GHDSPHKTGQNQQK	3085	GHDSPHKSGQKQQT	3126	GYDSPHKSGQKQST	3168	GHDSPHKLWINQQT	3210
GRGSPHKRGQNQQT	3086	GQDSPHKSGQNQDT	3127	MHDSPHKSGQNQQT	3169	GPDSPHKSGQNQQT	3211
GHDNPHKSGQNQQT	3087	GHDSPHKSGRNQQT	3128	GDDSPHKSVQNQQT	3170	GHDSPHKSVQKQLT	3212
GQDSPHKSGQHQQA	3088	GHDSPHKSGQNLLT	3129	GHDSPHKSVQSHQT	3171	GHDSPHKSGRNQLA	3213
GHDSPHKSGQIQHT	3089	GHDSPHKSGQLVIT	3130	GHDSPHKSGQFVVT	3172	VHDSPHKSGQNQQS	3214

GHDSPHKSGPRQQT	3090	GHDSPHKSRQSQQT	3131	GHDSPHKSRQNLQT	3173	GHDSPHKMGRNQQS	3215
GHDSPHKSGHTQQT	3091	GHDSPHKSGRTQET	3132	GHNSPHKSGQNQET	3174	GHDSPHKSGISIQT	3216
GHDSPHKSGQRQHT	3092	GHDSPHKSVQTHQT	3133	GHDSPHKSAQNQQI	3175	VHDSPHKSGQNQQT	3217
GHDSPHKSAQSQQT	3093	GHDSPHKSGQNQPA	3134	NHDSPHKSGQNQQT	3176	GHDSPHKSVQNLQT	3218
GHESPHKSGQNQQS	3094	KHDSPHKSGQNQQT	3135	GHDSPHKSGQSQPT	3177	GHDSPHKMAHNQQT	3219
GHDSPHKSLQNQQT	3095	GHDSPHKSGQTQQT	3136	GHESPHKSGQNEQT	3178	GHDSPHKHGQNQQN	3220
GHDSPHKSGRNQET	3096	GPDSPHKIGQNQQS	3137	GHDSPHKSAQNQQT	3179	GHDSPHKSVQSQQS	3221
GHDSPHKSGQTQLT	3097	GHDSPHKSVQNQQT	3138	GHDSPHKSGQNQLT	3180	GHDSPHKSGQTVCT	3222
GHDSPHKSEKNQQT	3098	GHDSPHKSRQDQHT	3139	GHDSPHKSAQNLLT	3181	GQDSPHKSGQYQQI	3223
GRDSPHKSGQDQQT	3099	GPDSPHKSGQKQQT	3140	THDSPHKSGQNQQT	3182	GHDSPHKSGQQIMT	3224
GHDSPHKTGHNQQT	3100	GHDSPHKSRQSQHT	3141	GNGSPHKRGQNQQT	3183	GHDSPHKSRQNEQS	3225
AHDSPHKSGQNQLT	3101	GHDSPHKSVQNQLT	3142	GHDSPHKSGQSLQT	3184	GHDSPHKSGLNHQT	3226
GYDSPHKSGQTQQT	3102	AHDSPHKSGQNQQT	3143	GHDSPHKSAQNHQT	3185	GYDSPHKSGQNQQT	3227
GHESPHKSGQTQQI	3103	GQDSPHKSGQNQQS	3144	GHDSPHKSGRNRQT	3186	GHDSPHKSGQNLQT	3228
GHDSPHKSGQSKQA	3104	GHDSPHKSGRNQQI	3145	EHDSPHKSGQNQQT	3187	GHDSPHKSRQDQQT	3229
GHESPHKSGQNQQT	3105	GYDSPHKSGQYQHT	3146	GHDSPHKYGQNEQT	3188	GDDSPHKSGQKQLT	3230
GHDSPHKTGQNQPP	3106	GHDSPHKSRQNQQT	3147	RHDSPHKSGHNQQT	3189	GDDSPHKSGHNQQT	3231
GHDSPHKSGQSPQT	3107	GHDSPHKSWVRQQT	3148	GHDSPHKSQQNQQT	3190	GHDSPHKSGQMIHT	3232
GNDSPHKSVQNQQT	3108	GHESPHKSGQNQHS	3149	GHDSPHKSGQNQQI	3191	GHDSPHKSGRNHQS	3233
GHDSPHKSAQNYQT	3109	GHDSPHKIGHNQQT	3150	QHDSPHKSGQNQQT	3192	GHDSPHKSVQNRQT	3234
AHDSPHKIGQNHQT	3110	WHDSPHKSGQNQQT	3151	RHDSPHKIVQNQQT	3193	GHDSPHKSGQKMNT	3235
GHESPHKSAQNRQT	3111	RHDSPHKSGQNQQT	3152	YHDSPHKSGQNQQT	3194	GHDSPHKSGQSQQN	3236
GHDSPHKSGQNQQG	3112	GHDSPHKSNAWQQT	3153	GHDSPHKSVQSKQT	3195	GHDSPHKSGQYQHA	3237
GQDSPHKIGQNQQT	3113	GHDSPHKSGQSVPT	3154	GNDSPHKIGHNQQT	3196	GHDSPHKSGQNQWT	3117
GHDSPHKSGQNHLT	3114	GHESPHKSGQNIQP	3155	HHDSPHKSGQNQQT	3197	GHDSPHKRGPDQQS	3158
GHDSPHKSGQYQHT	3115	GHDSPHKSVQNHLN	3156	GHDSPHKSGQTQQI	3198	GHDSPHKSGRDQKT	3200
GNDSPHKSVQNHQT	3116	GHDSPHKIGLDQQT	3157	GQDSPHKSGQNPLT	3199	GHDSPHKSVHNQQN	3201
GHDSPHKSVQNQHT	3118	GHDSPHKSGESQQT	3159	KTISKRGSPHSKAQNQQT	4098	KGLGGSGSPHSKAQNQQT	4099
KTINGHDSPHSKAQNLQT	4100	KTINGSGSPHSKTCIQQT	4196	KEIYGSGSPHSKAQNQQT	4292	KTINGSGSPHKRGQKQQT	4388

KTINGHDSPHSK AQNQQI	4101	KTINGSGSPHSK WLTQQT	4197	KELSGSGSPH SKAQNQQT	4293	KTINGSGSPH KRGQNQET	4389
KTINGSGSPHFT RQNQQT	4102	KTINGSGSPHSK WVVQQT	4198	KETIGSGSPH SKAQNQQT	4294	KTINGSGSPH KRGQNQLT	4390
KTINGSGSPHSL PWNQQT	4103	KTINGSGSPHSK YRLQQT	4199	KEVLGSGSP HSKAQNQQT	4295	KTINGSGSPH KRGRNQQT	4391
KTINGHDSPHSK AQNHQT	4104	KTINGSGSPHSK YSKQQT	4200	KFALGHDSP HKSGQKQQT	4296	KTINGSGSPH KSGGNQQT	4392
KTMNGHDSPHS KAQNQQT	4105	KTINGSGSPHSK YSRQQT	4201	KIINGHDSPH KSGQNLVL	4297	KTINGSGSPH KSGHNQET	4393
KPYKGSGSPHSK AQNQQT	4106	KTINGSGSPHSL KRNQQT	4202	KIINGHDSPH KSGQRNYT	4298	KTINGSGSPH KSGHNQLT	4394
KRLWGSGSPHS KAQNQQT	4107	KTINGSGSPHSL WFNQQT	4203	KIINGHDSPH SKAQNQQT	4299	KTINGSGSPH KSGHNQQN	4395
KRMRGSGSPHS KAQNQQT	4108	KTINGSGSPHSL WPNQQT	4204	KLNPGHDSP HKSGQTQQT	4300	KTINGSGSPH KSGLNQLT	4396
KRTYGSGSPHSK AQNQQT	4109	KTINGSGSPHSL WTNQQT	4205	KLNRGHDSP HKSGQNQQS	4301	KTINGSGSPH KSGPNQQT	4397
KTINCLRSPHSK AQNQQT	4110	KTINGSGSPHSM RRNQQT	4206	KLSSGHDSPH KSGQNQQN	4302	KTINGSGSPH KSGQGQQT	4398
KTINFSRSPHSK AQNQQT	4111	KTINGSGSPHSP CLNQQT	4207	KNINGHDSP HSKAQNQQT	4303	KTINGSGSPH KSGQHLQT	4399
KTINGLRSPHFK AQNQQT	4112	KTINGSGSPHSQ WQNQQT	4208	KNNDGSGSP HSKAQNQQT	4304	KTINGSGSPH KSGQHQQT	4400
KTINGNRSPHNK AQNQQT	4113	KTINGSGSPHSR CANQQT	4209	KNVMGSGSP HSKAQNQQT	4305	KTINGSGSPH KSGQKHQT	4401
KTINGPRSPHYK AQNQQT	4114	KTINGSGSPHSRI RNQQT	4210	KPINGHDSPH KSGQNKLS	4306	KTINGSGSPH KSGQKQQS	4402
KTINGQASPHW KAQNQQT	4115	KTINGSGSPHSR KSNQQT	4211	KPINGHDSPH KSGQNLSS	4307	KTINGSGSPH KSGQNEQT	4403
KTINGRCSPHSK AQNQQT	4116	KTINGSGSPHSR LWNQQT	4212	KPINGHDSPH SKAQNQQT	4308	KTINGSGSPH KSGQNHQT	4404
KTINGRHSPHSK AQNQQT	4117	KTINGSGSPHSR RFNQQT	4213	KRINGHDSPH SKAQNQQT	4309	KTINGSGSPH KSGQNKQT	4405
KTINGRKSPHRK AQNQQT	4118	KTINGSGSPHSR RPNQQT	4214	KSCSGHDSP HKSGQNQQS	4310	KTINGSGSPH KSGQNKTS	4406
KTINGRLSPHW KAQNQQT	4119	KTINGSGSPHSR SCNQQT	4215	KSINGHDSPH KSGQNLAS	4311	KTINGSGSPH KSGQNQEA	4407
KTINGRLSPHYK AQNQQT	4120	KTINGSGSPHSR SKNQQT	4216	KSINGHDSPH KSGQNLFL	4312	KTINGSGSPH KSGQNQET	4408
KTINGRPSPHMK AQNQQT	4121	KTINGSGSPHSR TKNQQT	4217	KSINGHDSPH KSGQNLLM	4313	KTINGSGSPH KSGQNQKT	4409
KTINGRSSPHWK AQNQQT	4122	KTINGSGSPHSR WLNQQT	4218	KSINGHDSPH KSGQNLLQ	4314	KTINGSGSPH KSGQNQQI	4410
KTINGRWSPHSK AQNQQT	4123	KTINGSGSPHSS VCNQQT	4219	KSINGHDSPH KSGQNSLG	4315	KTINGSGSPH KSGQNQQR	4411
KTINGSGSPHAP CQNQQT	4124	KTINGSGSPHSS WRNQQT	4220	KSINGHDSPH KSGQNTLQ	4316	KTINGSGSPH KSGQNQRT	4412
KTINGSGSPHA WAQNQQT	4125	KTINGSGSPHSV CQNQQT	4221	KSINGHDSPH KSSSNQQT	4317	KTINGSGSPH KSGQNQYT	4413
KTINGSGSPHCM RQNQQT	4126	KTINGSGSPHSV LCNQQT	4222	KSINGHDSPH KYKLNQQT	4318	KTINGSGSPH KSGQRQQT	4414
KTINGSGSPHFC SQNQQT	4127	KTINGSGSPHSV RRNQQT	4223	KSINGSGSPH KSGQKQQT	4319	KTINGSGSPH KSGQSQQT	4415
KTINGSGSPHFL FQNQQT	4128	KTINGSGSPHSV SCNQQT	4224	KSINGSGSPH KSGQNQQT	4320	KTINGSGSPH KSGQYQRT	4416
KTINGSGSPHFW AQNQQT	4129	KTINGSGSPHSW ALNQQT	4225	KSINGSGSPH SKAQGLST	4321	KTINGSGSPH KSGRNQQA	4417

KTINGSGSPHLC AQNQQT	4130	KTINGSGSPHSW ITNQQT	4226	KSINGSGSPH SKAQLLGT	4322	KTINGSGSPH KSRHNQQT	4418
KTINGSGSPHLR YQNQQT	4131	KTINGSGSPHSW PMNQQT	4227	KSINGSGSPH SKTSWQQT	4323	KTINGSGSPH KSRQYQQT	4419
KTINGSGSPHLY YQNQQT	4132	KTINGSGSPHSW RSNQQT	4228	KSMNGHDSP HSKAQNQQT	4324	KTINGSGSPH RKAQAPGT	4420
KTINGSGSPHPL CQNQQT	4133	KTINGSGSPHSY FLNQQT	4229	KSTLGSGSPH SKAQNQHT	4325	KTINGSGSPH SKAAMKQT	4421
KTINGSGSPHRI RQNQQT	4134	KTINGSGSPHSY TYNQQT	4230	KSTLGSGSPH SKAQNQQN	4326	KTINGSGSPH SKAGRQQT	4422
KTINGSGSPHRL FQNQQT	4135	KTINGSGSPHSY WQNQQT	4231	KSTVGSGSPH SKAQTQQT	4327	KTINGSGSPH SKAGRTQT	4423
KTINGSGSPHSC GQNQQT	4136	KTINGSGSPHTL CQNQQT	4232	KTCKESGSPH SKAQNQQT	4328	KTINGSGSPH SKAKSNQT	4424
KTINGSGSPHSC LRNQQT	4137	KTINGSGSPHWL RQNQQT	4233	KTCKGSGSP HSKAQNQQT	4329	KTINGSGSPH SKALKTQT	4425
KTINGSGSPHSC LSNQQT	4138	KTINGSGSPHWP SQNQQT	4234	KTCKSSGSPH SKAQNQQT	4330	KTINGSGSPH SKAPRTQT	4426
KTINGSGSPHSC RLNQQT	4139	KTINGSGSPHYL RQNQQT	4235	KTDMGSGSP HSKAQNQQT	4331	KTINGSGSPH SKAQAART	4427
KTINGSGSPHSC SLNQQT	4140	KTINGSGSPHYT RQNQQT	4236	KTDNGIGSPH SKAQNQQT	4332	KTINGSGSPH SKAQAILT	4428
KTINGSGSPHSK ACTLQT	4141	KTINGSLSPHLW AQNQQT	4237	KTEGGSGSP HSKAQNQQT	4333	KTINGSGSPH SKAQCRGT	4429
KTINGSGSPHSK AFRAQT	4142	KTINGSPSPHCQ AQNQQT	4238	KTEHHSGSP HSKAQNQQT	4334	KTINGSGSPH SKAQGLRT	4430
KTINGSGSPHSK AIRKQT	4143	KTINGSRSPHLC AQNQQT	4239	KTEKDSGSP HSKAQNQQT	4335	KTINGSGSPH SKAQKGVL	4431
KTINGSGSPHSK AQASRT	4144	KTINGSRSPHWR AQNQQT	4240	KTELGHDSP HKRGQNQQT	4336	KTINGSGSPH SKAQKSNT	4432
KTINGSGSPHSK AQFELT	4145	KTINGSVSPHWL AQNQQT	4241	KTESVSGSPH SKAQNQQT	4337	KTINGSGSPH SKAQNNKF	4433
KTINGSGSPHSK AQIVIT	4146	KTINGTFSPHRK AQNQQT	4242	KTETNSGSPH SKAQNQQT	4338	KTINGSGSPH SKAQNRRT	4434
KTINGSGSPHSK AQLART	4147	KTINGWTSPHR KAQNQQT	4243	KTETYSGSPH SKAQNQQT	4339	KTINGSGSPH SKAQPKQT	4435
KTINGSGSPHSK AQLQRT	4148	KTINRGISPHSK AQNQQT	4244	KTEWLSGSP HSKAQNQQT	4340	KTINGSGSPH SKAQRAPT	4436
KTINGSGSPHSK AQNARR	4149	KTINTVRSPHSK AQNQQT	4245	KTFNGSGSPH KSGQNQQT	4341	KTINGSGSPH SKAQREHT	4437
KTINGSGSPHSK AQNCPR	4150	KTKLRSGSPHSK AQNQQT	4246	KTGLRHDSP HKSGQKQQT	4342	KTINGSGSPH SKAQRFGT	4438
KTINGSGSPHSK AQNMRR	4151	KTRLRSGSPHSK AQNQQT	4247	KTGLRHDSP HKSGQNQQS	4343	KTINGSGSPH SKAQRPCT	4439
KTINGSGSPHSK AQNRRV	4152	KWLLGSGSPHS KAQNQQT	4248	KTGVTHDSP HKSGQKQQT	4344	KTINGSGSPH SKAQRQAT	4440
KTINGSGSPHSK AQPSRT	4153	KWSQGSGSPHS KAQNQQT	4249	KTIDGHESPH SKAQNQQT	4345	KTINGSGSPH SKAQRQPT	4441
KTINGSGSPHSK AQQVKT	4154	KWYLGSGSPHS KAQNQQT	4250	KTIEGHDSPH KSGQTQQT	4346	KTINGSGSPH SKAQTKLT	4442
KTINGSGSPHSK AQQVRT	4155	KYHSGSGSPHS KAQNQQT	4251	KTIHGHDSPH SKAQNQQT	4347	KTINGSGSPH SKAQTTHT	4443
KTINGSGSPHSK AQRLKT	4156	KYLPGSGSPHSK AQNQQT	4252	KTIHGHESPH SKAQNQQT	4348	KTINGSGSPH SKAQVQRT	4444
KTINGSGSPHSK AQRRAT	4157	KAINGGGSPHSK TQNQQT	4253	KTIIGHDSPH KSGQNRSS	4349	KTINGSGSPH SKAQVVRT	4445
KTINGSGSPHSK AQRRGT	4158	KAINGHDSPHK RSPNQQT	4254	KTIIGHDSPH KSGQRLGT	4350	KTINGSGSPH SKAQWPNT	4446

KTINGSGSPHSK AQRRRT	4159	KAINGHDSPHKS FSPQQT	4255	KTIIGSGSPH KSGQNQQT	4351	KTINGSGSPH SKAQYPST	4447
KTINGSGSPHSK AQRTRT	4160	KAINGHDSPHKS GENQQP	4256	KTIKGHDSPH KSGQNMLF	4352	KTINGSGSPH SKARALQT	4448
KTINGSGSPHSK AQRVHT	4161	KAINGHDSPHKS GQLART	4257	KTILGSGSPH SKAQNLQT	4353	KTINGSGSPH SKARDQHT	4449
KTINGSGSPHSK AQTYRT	4162	KAINGHDSPHKS GQNAFL	4258	KTINGCSSPH WKAQNQQT	4354	KTINGSGSPH SKARFQQT	4450
KTINGSGSPHSK AQVRKT	4163	KAINGHDSPHKS GQNAYT	4259	KTINGGGSTH SKAQNQQT	4355	KTINGSGSPH SKARRTQT	4451
KTINGSGSPHSK ARGRQT	4164	KAINGHDSPHKS GQNFAS	4260	KTINGHDSPH KAGQSQQT	4356	KTINGSGSPH SKARSLQT	4452
KTINGSGSPHSK ARLCQT	4165	KAINGHDSPHKS GQNLAS	4261	KTINGHDSPH KRGQNVPS	4357	KTINGSGSPH SKARVIQT	4453
KTINGSGSPHSK ARLKQT	4166	KAINGHDSPHKS GQNLGS	4262	KTINGHDSPH KRGRSYQT	4358	KTINGSGSPH SKAWYLQT	4454
KTINGSGSPHSK ARNSQT	4167	KAINGHDSPHKS GQNLKF	4263	KTINGHDSPH KTGQNPPT	4359	KTINGSGSPH SKGGGQQT	4455
KTINGSGSPHSK ARWVQT	4168	KAINGHDSPHKS GQNLLK	4264	KTINGHDSPH SKAENQQT	4360	KTINGSGSPH SKGSRQQT	4456
KTINGSGSPHSK AVRWQT	4169	KAINGHDSPHKS GQNLSR	4265	KTINGHDSPH SKALSLQT	4361	KTINGSGSPH SKLQRQQT	4457
KTINGSGSPHSK AYTRQT	4170	KAINGHDSPHKS GQNLSS	4266	KTINGHDSPH SKAQGQQT	4362	KTINGSGSPH SKMLRQQT	4458
KTINGSGSPHSK CQSQQT	4171	KAINGHDSPHKS GQNSLG	4267	KTINGHDSPH SKAQHQQT	4363	KTINGSGSPH SKSSIKQT	4459
KTINGSGSPHSK FLRQQT	4172	KAINGHDSPHKS GQNTLQ	4268	KTINGHDSPH SKAQIQQT	4364	KTINGSGSPH SKVRFQQT	4460
KTINGSGSPHSK FRFQQT	4173	KAINGHDSPHKS GQNTSL	4269	KTINGHDSPH SKAQKQQT	4365	KTINGSGSPH SVVWNQQT	4461
KTINGSGSPHSK FRLQQT	4174	KAINGHDSPHKS GQRLGT	4270	KTINGHDSPH SKAQNLSS	4366	KTINGSTSPH KLAQNQQP	4462
KTINGSGSPHSK FRRQQT	4175	KAINGHDSPHKS GQRNYT	4271	KTINGHDSPH SKAQNPQT	4367	KTINRHDSPH KSGQRPST	4463
KTINGSGSPHSK GMKQQT	4176	KAINGHDSPHKS GQRPST	4272	KTINGHDSPH SKAQNQET	4368	KTINRIMSPH SKAQNQQT	4464
KTINGSGSPHSK KLRQQT	4177	KAINGHDSPHKS GQRPVT	4273	KTINGHDSPH SKAQNQHT	4369	KTINTARSPH SKAQNQQT	4465
KTINGSGSPHSK KRPQQT	4178	KAINGHDSPHKS GQVPST	4274	KTINGHDSPH SKAQNQLT	4370	KTISGHDSPH SKAQNQQT	4466
KTINGSGSPHSK KSRQQT	4179	KAINGHDSPHKS LSNQQT	4275	KTINGHDSPH SKAQNQPT	4371	KTISGSGSPH KSGQNQQT	4467
KTINGSGSPHSK LYRQQT	4180	KAINGHDSPHKS VLSQQT	4276	KTINGHDSPH SKAQNQQA	4372	KTITGHDSPH KSGQRLGT	4468
KTINGSGSPHSK LYWQQT	4181	KAINGHDSPHKT LQNQQT	4277	KTINGHDSPH SKAQNTGS	4373	KTITGSGSPH KSGQNQQT	4469
KTINGSGSPHSK PRMQQT	4182	KAINGHNSPHSK AQNQQT	4278	KTINGHDSPH SKAQSQQT	4374	KTIYGHDSP HKSGQRLGT	4470
KTINGSGSPHSK RFPQQT	4183	KAINGLDSPHSK AQNQQT	4279	KTINGHDSPH SKAQTQQT	4375	KTLNGHDSP HKSGQNLFL	4471
KTINGSGSPHSK RFRQQT	4184	KAINGSGSPHKS GQNQQT	4280	KTINGHDSPH SKAQYQQT	4376	KTLNGHDSP HKSGQNLSS	4472
KTINGSGSPHSK RPYQQT	4185	KAINGSGSPHSK AQGQQT	4281	KTINGHDSPH SKARNQQT	4377	KTLSFHDSPH KSGQNQQS	4473
KTINGSGSPHSK RRMQQT	4186	KAINGSGSPHSK AQLSGT	4282	KTINGHDSPH SKLPGQQT	4378	KTSNGSGSP HSKAQNTMT	4474
KTINGSGSPHSK RSKQQT	4187	KAINGSGSPHSK AQNGSL	4283	KTINGHDSPH SKSPNQQT	4379	KTTNGHDSP HSKAQNQQT	4475

KTINGSGSPHSK RSRQQT	4188	KAINGSGSPHSK AQNSLL	4284	KTINGHESPH KSGQNAFL	4380	KTVNGGGSP HSKAQNQQT	4476
KTINGSGSPHSK RTMQQT	4189	KAINGSGSPHSK AVGLQT	4285	KTINGIGSPH SKAPNEQT	4381	KTVNGHDSP HKSGQNVSL	4477
KTINGSGSPHSK RTRQQT	4190	KAINGSGSPHSK SLLQQT	4286	KTINGQDSPH KSGQNLHM	4382	KTVNGHDSP HKSGQRPST	4478
KTINGSGSPHSK RVRQQT	4191	KAINGSGSPHSK SLPQQT	4287	KTINGRGSPH SKAQIGMT	4383	KTVNGHDSP HKSGQTQQA	4479
KTINGSGSPHSK RYIQQT	4192	KAINGSGSPHSK STFQQT	4288	KTINGRGSPH SKAQNQVL	4384	KTVNGHESP HSKAQNQQT	4480
KTINGSGSPHSK RYNQQT	4193	KAITGHDSPHSK AQNQQT	4289	KTINGRGSPH SKAQSPTT	4385	KTVNGSGSP HSKAQGLST	4481
KTINGSGSPHSK RYPQQT	4194	KDVMGSGSPHS KAQNQQT	4290	KTINGRGSPH SKATSFQT	4386	KTVNGSGSP HSKAQNVTS	4482
KTINGSGSPHSK RYSQQT	4195	KEIVGSGSPHSK AQNQQT	4291	KTINGSGSPH FVVQNQQT	4387	KTVPASGSP HSKAQNQQT	4483
NTINGSGSPHSK AHNQQT	4484	TTINGGGSPHSK AQNQQT	4485	GHDSPHKS	4487	NGHDSPHKS G	4489
INGHDSPHKSGQ	4490	TINGHDSPHKSG QN	4491	KTINGHDSPH KSGQNQ	4492	GSPHSKAQ	4493
SGSPHSKAQN	4494	GSGSPHSKAQN Q	4495	NGSGSPHSK AQNQQ	4496	INGSGSPHSK AQNQQT	4497
LYYLSKTINGHD SPHKSGQNQQT LKF	4518	RLMNPLIDQYL YYLSKTINGHDS PHKSGQNQQTL KFSVAGPSNMA V	4519				

Table 2A. Exemplary Peptide Sequences

SEQ ID NO:	Amino Acid Sequence	SEQ ID NO:	Nucleotide Sequence
941	SPHSKA	942	AGCCCACACAGCAAAGCA
943	HDSPHKSG	944	CACGACAGCCCACACAAAAGCGGA
2	HDSPHK	3	CACGACAGCCCACACAAA

Table 2B. Exemplary Peptide Sequences

Amino Acid Sequence	SEQ ID NO:	Amino Acid Sequence	SEQ ID NO:	Amino Acid Sequence	SEQ ID NO:	Amino Acid Sequence	SEQ ID NO:
SPHSKA	945	SPHKKN	954	SPHKTS	963	SPHTRG	972
SPHKSG	946	SPHVRM	955	SPHKTT	964	SPHVRG	973
SPHARM	947	SPHRKA	956	SPHKTY	965	SPHKRG	974
SPHVKS	948	SPHKFG	957	SPHKYG	966	SPHGAR	975
SPHASR	949	SPHKIG	958	SPHSKD	967	SPHRSG	976
SPHVKI	950	SPHKLG	959	SPHSKP	968	SPHKSA	977
SPHKSR	951	SPHSKL	960	SPHSRA	969	SPHSKR	978
SPHSLR	952	SPHSRG	961	SPHSSR	970	SPHFLR	979
SPHSKW	953	SPHSKS	962	SPHWKA	971	SPHVRR	980
STHASR	985	SQHKSG	986				

Table 2C. Exemplary Phosphorylated Peptide Sequences

TTM-002 Derived Phospho Peptide Sequence	SEQ ID NO
NGHD**pS**PHKSG	4515
KTINGHD**pS**PHKSGQNQ	4516

YLSKTINGHD**pS**PHKSGQNQQTLKFS	4517

[0051] In some embodiments, a ligand described herein comprises a protein or a peptide comprising an amino acid sequence having the formula [N1]-[N2]-[N3], wherein [N2] comprises the amino acid sequence of SPH and [N3] comprises X4, X5, and X6, wherein at least one of X4, X5, or X6 is a basic amino acid, e.g., a K or R. In some embodiments, position X4 of [N2] is K. In some embodiments, position X5 of [N2] is K.

[0052] In some embodiments, [N1] comprises X1, X2, and X3, wherein at least one of X1, X2, or X3 is G. In some embodiments, position X1 of [N1] is independently chosen from G, V, R, D, E, M, T, I, S, A, N, L, K, H, P, W, or C. In some embodiments, position X2 of [N1] is independently chosen from: S, V, L, N, D, H, R, P, G, T, I, A, E, Y, M, or Q. In some embodiments, position X3 of [N1] is independently chosen from: G, C, L, D, E, Y, H, V, A, N, P, or S. In some embodiments, [N1] comprises GS, SG, GH, HD, GQ, QD, VS, CS, GR, RG, QS, SH, MS, RN, TS, IS, GP, ES, SS, GN, AS, NS, LS, GG, KS, GT, PS, RS, GI, WS, DS, ID, GL, DA, DG, ME, EN, KN, KE, AI, NG, PG, TG, SV, IG, LG, AG, EG, SA, YD, HE, HG, RD, ND, PD, MG, QV, DD, HN, HP, GY, GM, GD, or HS. In some embodiments, [N1] comprises GS, SG, GH, or HD. In some embodiments [N1] is or comprises GSG, GHD, GQD, VSG, CSG, GRG, CSH, GQS, GSH, RVG, GSC, GLL, GDD, GHE, GNY, MSG, RNG, TSG, ISG, GPG, ESG, SSG, GNG, ASG, NSG, LSG, GGG, KSG, HSG, GTG, PSG, GSV, RSG, GIG, WSG, DSG, IDG, GLG, DAG, DGG, MEG, ENG, GSA, KNG, KEG, AIG, GYD, GHG, GRD, GND, GPD, GMG, GQV, GHN, GHP, or GHS. In some embodiments, [N1] is or comprises GSG. In some embodiments, [N1] is or comprises GHD. In some embodiments, [N1]-[N2] comprises SGSPH (SEQ ID NO: 4752), HDSPH (SEQ ID NO: 4703), QDSPH (SEQ ID NO: 4753), RGSPH (SEQ ID NO: 4754), SHSPH (SEQ ID NO: 4755), QSSPH (SEQ ID NO: 4756), DDSPH (SEQ ID NO: 4757), HESPH (SEQ ID NO: 4758), NYSPH (SEQ ID NO: 4759), VGSPH (SEQ ID NO: 4760), SCSPH (SEQ ID NO: 4761), LLSPH (SEQ ID NO: 4762), NGSPH (SEQ ID NO: 4763), PGSPH (SEQ ID NO: 4764), GGSPH (SEQ ID NO: 4765), TGSPH (SEQ ID NO: 4766), SVSPH (SEQ ID NO: 4767), IGSPH (SEQ ID NO: 4768), DGSPH (SEQ ID NO: 4769), LGSPH (SEQ ID NO: 4770), AGSPH (SEQ ID NO: 4771), EGSPH (SEQ ID NO: 4772), SASPH (SEQ ID NO: 4773), YDSPH (SEQ ID NO: 4774), HGSPH (SEQ ID NO: 4775), RDSPH (SEQ ID NO: 4776), NDSPH (SEQ ID NO: 4777), PDSPH (SEQ ID NO: 4778), MGSPH (SEQ ID NO: 4779), QVSPH (SEQ ID NO: 4780), HNSPH (SEQ ID NO: 4781), HPSPH (SEQ ID NO: 4782), or HSSPH (SEQ ID NO: 4783); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, or 4 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-

[N2] is or comprises GSGSPH (SEQ ID NO: 4695), GHDSPH (SEQ ID NO: 4784), GQDSPH (SEQ ID NO: 4785), VSGSPH (SEQ ID NO: 4786), CSGSPH (SEQ ID NO: 4787), GRGSPH (SEQ ID NO: 4788), CSHSPH (SEQ ID NO: 4789), GQSSPH (SEQ ID NO: 4790), GSHSPH (SEQ ID NO: 4791), GDDSPH (SEQ ID NO: 4792), GHESPH (SEQ ID NO: 4793), GNYSPH (SEQ ID NO: 4794), RVGSPH (SEQ ID NO: 4795), GSCSPH (SEQ ID NO: 4796), GLLSPH (SEQ ID NO: 4797), MSGSPH (SEQ ID NO: 4798), RNGSPH (SEQ ID NO: 4799), TSGSPH (SEQ ID NO: 4800), ISGSPH (SEQ ID NO: 4801), GPGSPH (SEQ ID NO: 4802), ESGSPH (SEQ ID NO: 4803), SSGSPH (SEQ ID NO: 4804), GNGSPH (SEQ ID NO: 4805), ASGSPH (SEQ ID NO: 4806), NSGSPH (SEQ ID NO: 4807), LSGSPH (SEQ ID NO: 4808), GGGSPH (SEQ ID NO: 4809), KSGSPH (SEQ ID NO: 4810), HSGSPH (SEQ ID NO: 4811), GTGSPH (SEQ ID NO: 4812), PSGSPH (SEQ ID NO: 4813), GSVSPH (SEQ ID NO: 4814), RSGSPH (SEQ ID NO: 4815), GIGSPH (SEQ ID NO: 4816), WSGSPH (SEQ ID NO: 4817), DSGSPH (SEQ ID NO: 4818), IDGSPH (SEQ ID NO: 4819), GLGSPH (SEQ ID NO: 4820), DAGSPH (SEQ ID NO: 4821), DGGSPH (SEQ ID NO: 4822), MEGSPH (SEQ ID NO: 4823), ENGSPH (SEQ ID NO: 4824), GSASPH (SEQ ID NO: 4825), KNGSPH (SEQ ID NO: 4826), KEGSPH (SEQ ID NO: 4827), AIGSPH (SEQ ID NO: 4828), GYDSPH (SEQ ID NO: 4829), GHGSPH (SEQ ID NO: 4830), GRDSPH (SEQ ID NO: 4831), GNDSPH (SEQ ID NO: 4832), GPDSPH (SEQ ID NO: 4833), GMGSPH (SEQ ID NO: 4834), GQVSPH (SEQ ID NO: 4835), GHNSPH (SEQ ID NO: 4836), GHPSPH (SEQ ID NO: 4837), or GHSSPH (SEQ ID NO: 4838); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, or 5 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2] is or comprises GSGSPH (SEQ ID NO: 4695). In some embodiments, [N1]-[N2] is or comprises GHDSPH (SEQ ID NO: 4784).

[0053] In some embodiments, X4, X5, or both of [N3] are K. In some embodiments, X4, X5, or X6 of [N3] is R. In some embodiments, position X4 of [N3] is independently chosen from: A, K, V, S, T, G, F, W, V, N, or R. In some embodiments, position X5 of [N3] is independently chosen from: S, K, T, F, I, L, Y, H, M, or R. In some embodiments, position X6 of [N3] is independently chosen from: G, R, A, M, I, N, T, Y, D, P, V, L, E, W, N, Q, K, or S. In some embodiments, [N3] comprises SK, KA, KS, AR, RM, VK, AS, SR, VK, KR, KK, KN, VR, RS, RK, KT, TS, KF, FG, KI, IG, KL, LG, TT, TY, KY, YG, KD, KP, TR, RG, VR, GA, SL, SS, FL, WK, SA, RA, LR, KW, RR, GK, TK, NK, AK, KV, KG, KH, KM, TG, SE, SV, SW, SN, HG, SQ, LW, MG, MA, or SG. In some embodiments, [N3] comprises SK, KA, KS, or SG. In some embodiments, [N3] is or comprises SKA, KSG, ARM, VKS, ASR, VKI, KKN, VRM, RKA, KTS, KFG, KIG, KLG, KTT, KTY, KYG, SKD, SKP, TRG, VRG, KRG, GAR, KSA, KSR, SKL, SRA, SKR, SLR, SRG, SSR, FLR, SKW, SKS,

WKA, VRR, SKV, SKT, SKG, GKA, TKA, NKA, SKL, SKN, AKA, KTG, KSL, KSE, KSV, KSW, KSN, KHG, KSQ, KSK, KLW, WKG, KMG, KMA, or RSG. In some embodiments, [N3] is or comprises SKA. In some embodiments, [N3] is or comprises KSG. In some embodiments, [N2]-[N3] comprises SPHSK (SEQ ID NO: 4701), SPHKS (SEQ ID NO: 4704), SPHAR (SEQ ID NO: 4705), SPHVK (SEQ ID NO: 4706), SPHAS (SEQ ID NO: 4707), SPHKK (SEQ ID NO: 4708), SPHVR (SEQ ID NO: 4709), SPHRK (SEQ ID NO: 4710), SPHKT (SEQ ID NO: 4711), SPHKF (SEQ ID NO: 4712), SPHKI (SEQ ID NO: 4713), SPHKL (SEQ ID NO: 4714), SPHKY (SEQ ID NO: 4715), SPHTR (SEQ ID NO: 4716), SPHKR (SEQ ID NO: 4717), SPHGA (SEQ ID NO: 4718), SPHSR (SEQ ID NO: 4719), SPHSL (SEQ ID NO: 4720), SPHSS (SEQ ID NO: 4721), SPHFL (SEQ ID NO: 4722), SPHWK (SEQ ID NO: 4723), SPHGK (SEQ ID NO: 4724), SPHTK (SEQ ID NO: 4725), SPHNK (SEQ ID NO: 4726), SPHAK (SEQ ID NO: 4727), SPHKH (SEQ ID NO: 4728), SPHKM (SEQ ID NO: 4729), or SPHRS (SEQ ID NO: 4730). In some embodiments [N2]-[N3] comprises (SEQ ID NO: 4701) or SPHKS (SEQ ID NO: 4704). In some embodiments, [N2]-[N3] is or comprises SPHSKA (SEQ ID NO: 941), SPHKSG (SEQ ID NO: 946), SPHARM (SEQ ID NO: 947), SPHVKS (SEQ ID NO: 948), SPHASR (SEQ ID NO: 949), SPHVKI (SEQ ID NO: 950), SPHKKN (SEQ ID NO: 954), SPHVRM (SEQ ID NO: 955), SPHRKA (SEQ ID NO: 956), SPHKFG (SEQ ID NO: 957), SPHKIG (SEQ ID NO: 958), SPHKLG (SEQ ID NO: 959), SPHKTS (SEQ ID NO: 963), SPHKTT (SEQ ID NO: 964), SPHKTY (SEQ ID NO: 965), SPHKYG (SEQ ID NO: 966), SPHSKD (SEQ ID NO: 967), SPHSKP (SEQ ID NO: 968), SPHTRG (SEQ ID NO: 972), SPHVRG (SEQ ID NO: 973), SPHKRG (SEQ ID NO: 974), SPHGAR (SEQ ID NO: 975), SPHKSA (SEQ ID NO: 977), SPHKSR (SEQ ID NO: 951), SPHSKL (SEQ ID NO: 960), SPHSRA (SEQ ID NO: 969), SPHSKR (SEQ ID NO: 978), SPHSLR (SEQ ID NO: 952), SPHSRG (SEQ ID NO: 961), SPHSSR (SEQ ID NO: 970), SPHFLR (SEQ ID NO: 979), SPHSKW (SEQ ID NO: 953), SPHSKS (SEQ ID NO: 962), SPHWKA (SEQ ID NO: 971), SPHVRR (SEQ ID NO: 980), SPHSKT (SEQ ID NO: 4731), SPHSKG (SEQ ID NO: 4732), SPHGKA (SEQ ID NO: 4733), SPHNKA (SEQ ID NO: 4734), SPHSKN (SEQ ID NO: 4735), SPHAKA (SEQ ID NO: 4736), SPHSKV (SEQ ID NO: 4737), SPHKTG (SEQ ID NO: 4738), SPHTKA (SEQ ID NO: 4739), SPHKSL (SEQ ID NO: 4740), SPHKSE (SEQ ID NO: 4741), SPHKSV (SEQ ID NO: 4742), SPHKSW (SEQ ID NO: 4743), SPHKSN (SEQ ID NO: 4744), SPHKHG (SEQ ID NO: 4745), SPHKSQ (SEQ ID NO: 4746), SPHKSK (SEQ ID NO: 4747), SPHKLW (SEQ ID NO: 4748), SPHWKG (SEQ ID NO: 4749), SPHKMG (SEQ ID NO: 4750), SPHKMA (SEQ ID NO: 4751), or SPHRSG (SEQ ID NO: 976). In some embodiments, [N2]-[N3] is or comprises SPHSKA (SEQ ID NO: 941). In some embodiments, [N2]-[N3] is or comprises SPHKSG (SEQ ID NO: 946).

[0054] In some embodiments, [N1]-[N2]-[N3] comprises SGSPHSK (SEQ ID NO: 4839), HDSPHKS (SEQ ID NO: 4840), SGSPHAR (SEQ ID NO: 4841), SGSPHVK (SEQ ID NO: 4842), QDSPHKS (SEQ ID NO: 4843), SGSPHKK (SEQ ID NO: 4844), SGSPHVR (SEQ ID NO: 4845), SGSPHAS (SEQ ID NO: 4846), SGSPHRK (SEQ ID NO: 4847), SGSPHKT (SEQ ID NO: 4848),

SHSPHKS (SEQ ID NO: 4849), QSSPHRS (SEQ ID NO: 4850), RGSPHAS (SEQ ID NO: 4851), RGSPHSK (SEQ ID NO: 4852), SGSPHKF (SEQ ID NO: 4853), SGSPHKI (SEQ ID NO: 4854), SGSPHKL (SEQ ID NO: 4855), SGSPHKY (SEQ ID NO: 4856), SGSPHTR (SEQ ID NO: 4857), SHSPHKR (SEQ ID NO: 4858), SGSPHGA (SEQ ID NO: 4859), HDSPHKR (SEQ ID NO: 4860), DDSPHKS (SEQ ID NO: 4861), HESPHKS (SEQ ID NO: 4862), NYSPHKI (SEQ ID NO: 4863), SGSPHSR (SEQ ID NO: 4864), SGSPHSL (SEQ ID NO: 4865), SGSPHSS (SEQ ID NO: 4866), VGSPHSK (SEQ ID NO: 4867), SCSPHRK (SEQ ID NO: 4868), SGSPHFL (SEQ ID NO: 4869), LLSPHWK (SEQ ID NO: 4870), NGSPHSK (SEQ ID NO: 4871), PGSPHSK (SEQ ID NO: 4872), GGSPHSK (SEQ ID NO: 4873), TGSPHSK (SEQ ID NO: 4874), SVSPHGK (SEQ ID NO: 4875), SGSPHTK (SEQ ID NO: 4876), IGSPHSK (SEQ ID NO: 4877), DGSPHSK (SEQ ID NO: 4878), SGSPHNK (SEQ ID NO: 4879), LGSPHSK (SEQ ID NO: 4880), AGSPHSK (SEQ ID NO: 4881), EGSPHSK (SEQ ID NO: 4882), SASPHSK (SEQ ID NO: 4883), SGSPHAK (SEQ ID NO: 4884), HDSPHKI (SEQ ID NO: 4885), YDSPHKS (SEQ ID NO: 4886), HDSPHKT (SEQ ID NO: 4887), RGSPHKR (SEQ ID NO: 4888), HGSPHSK (SEQ ID NO: 4889), RDSPHKS (SEQ ID NO: 4890), NDSPHKS (SEQ ID NO: 4891), QDSPHKI (SEQ ID NO: 4892), PDSPHKI (SEQ ID NO: 4893), PDSPHKS (SEQ ID NO: 4894), MGSPHSK (SEQ ID NO: 4895), HDSPHKH (SEQ ID NO: 4896), QVSPHKS (SEQ ID NO: 4897), HNSPHKS (SEQ ID NO: 4898), NGSPHKR (SEQ ID NO: 4899), HDSPHKY (SEQ ID NO: 4900), NDSPHKI (SEQ ID NO: 4901), HDSPHKL (SEQ ID NO: 4902), HPSPHWK (SEQ ID NO: 4903), HDSPHKM (SEQ ID NO: 4904), or HSSPHRS (SEQ ID NO: 4905). In some embodiments, [N1]-[N2]-[N3] is or comprises GSGSPHSKA (SEQ ID NO: 4697), GHDSPHKSG (SEQ ID NO: 4698), GSGSPHARM (SEQ ID NO: 4906), GSGSPHVKS (SEQ ID NO: 4907), GQDSPHKSG (SEQ ID NO: 4908), GSGSPHASR (SEQ ID NO: 4909), GSGSPHVKI (SEQ ID NO: 4910), GSGSPHKKN (SEQ ID NO: 4911), GSGSPHVRM (SEQ ID NO: 4912), VSGSPHSKA (SEQ ID NO: 4913), CSGSPHSKA (SEQ ID NO: 4914), GSGSPHRKA (SEQ ID NO: 4915), CSGSPHKTS (SEQ ID NO: 4916), CSHSPHKSG (SEQ ID NO: 4917), GQSSPHRSG (SEQ ID NO: 4918), GRGSPHASR (SEQ ID NO: 4919), GRGSPHSKA (SEQ ID NO: 4920), GSGSPHKFG (SEQ ID NO: 4921), GSGSPHKIG (SEQ ID NO: 4922), GSGSPHKLG (SEQ ID NO: 4923), GSGSPHKTS (SEQ ID NO: 4924), GSGSPHKTT (SEQ ID NO: 4925), GSGSPHKTY (SEQ ID NO: 4926), GSGSPHKYG (SEQ ID NO: 4927), GSGSPHSKD (SEQ ID NO: 4928), GSGSPHSKP (SEQ ID NO: 4929), GSGSPHTRG (SEQ ID NO: 4930), GSGSPHVRG (SEQ ID NO: 4931), GSHSPHKRG (SEQ ID NO: 4932), GSHSPHKSG (SEQ ID NO: 4933), VSGSPHASR (SEQ ID NO: 4934), VSGSPHGAR (SEQ ID NO: 4935), VSGSPHKFG (SEQ ID NO: 4936), GHDSPHKRG (SEQ ID NO: 4937), GDDSPHKSG (SEQ ID NO: 4938), GHESPHKSA (SEQ ID NO: 4939), GHDSPHKSA (SEQ ID NO: 4940), GNYSPHKIG (SEQ ID NO: 4941), GHDSPHKSR (SEQ ID NO: 4942), GSGSPHSKL (SEQ ID NO: 4943), GSGSPHSRA (SEQ ID NO: 4944), GSGSPHSKR (SEQ ID NO: 4945), GSGSPHSLR (SEQ ID NO: 4946), GSGSPHSRG (SEQ ID NO: 4947), GSGSPHSSR (SEQ ID NO: 4948), RVGSPHSKA (SEQ ID NO: 4949), GSCSPHRKA (SEQ

ID NO: 4950), GSGSPHFLR (SEQ ID NO: 4951), GSGSPHSKW (SEQ ID NO: 4952), GSGSPHSKS (SEQ ID NO: 4953), GLLSPHWKA (SEQ ID NO: 4954), GSGSPHVRR (SEQ ID NO: 4955), GSGSPHSKV (SEQ ID NO: 4956), MSGSPHSKA (SEQ ID NO: 4957), RNGSPHSKA (SEQ ID NO: 4958), TSGSPHSKA (SEQ ID NO: 4959), ISGSPHSKA (SEQ ID NO: 4960), GPGSPHSKA (SEQ ID NO: 4961), GSGSPHSKT (SEQ ID NO: 4962), ESGSPHSKA (SEQ ID NO: 4963), SSGSPHSKA (SEQ ID NO: 4964), GNGSPHSKA (SEQ ID NO: 4965), ASGSPHSKA (SEQ ID NO: 4966), NSGSPHSKA (SEQ ID NO: 4967), LSGSPHSKA (SEQ ID NO: 4968), GGGSPHSKA (SEQ ID NO: 4969), KSGSPHSKA (SEQ ID NO: 4970), GGGSPHSKS (SEQ ID NO: 4971), GSGSPHSKG (SEQ ID NO: 4972), HSGSPHSKA (SEQ ID NO: 4973), GTGSPHSKA (SEQ ID NO: 4974), PSGSPHSKA (SEQ ID NO: 4975), GSVSPHGKA (SEQ ID NO: 4976), RSGSPHSKA (SEQ ID NO: 4977), GSGSPHTKA (SEQ ID NO: 4978), GIGSPHSKA (SEQ ID NO: 4979), WSGSPHSKA (SEQ ID NO: 4980), DSGSPHSKA (SEQ ID NO: 4981), IDGSPHSKA (SEQ ID NO: 4982), GSGSPHNKA (SEQ ID NO: 4983), GLGSPHSKS (SEQ ID NO: 4984), DAGSPHSKA (SEQ ID NO: 4985), DGGSPHSKA (SEQ ID NO: 4986), MEGSPHSKA (SEQ ID NO: 4987), ENGSPHSKA (SEQ ID NO: 4988), GSASPHSKA (SEQ ID NO: 4989), GNGSPHSKS (SEQ ID NO: 4990), KNGSPHSKA (SEQ ID NO: 4991), KEGSPHSKA (SEQ ID NO: 4992), AIGSPHSKA (SEQ ID NO: 4993), GSGSPHSKN (SEQ ID NO: 4994), GSGSPHAKA (SEQ ID NO: 4995), GHDSPHKIG (SEQ ID NO: 4996), GYDSPHKSG (SEQ ID NO: 4997), GHESPHKSG (SEQ ID NO: 4998), GHDSPHKTG (SEQ ID NO: 4999), GRGSPHKRG (SEQ ID NO: 5000), GQDSPHKSG (SEQ ID NO: 4908), GHDSPHKSL (SEQ ID NO: 5001), GHGSPHSKA (SEQ ID NO: 5002), GHDSPHKSE (SEQ ID NO: 5003), VSGSPHSKA (SEQ ID NO: 4913), GRDSPHKSG (SEQ ID NO: 5004), GNDSPHKSV (SEQ ID NO: 5005), GQDSPHKIG (SEQ ID NO: 5006), GHDSPHKSV (SEQ ID NO: 5007), GPDSPHKIG (SEQ ID NO: 5008), GPDSPHKSG (SEQ ID NO: 5009), GHDSPHKSW (SEQ ID NO: 5010), GHDSPHKSN (SEQ ID NO: 5011), GMGSPHSKT (SEQ ID NO: 5012), GHDSPHKHG (SEQ ID NO: 5013), GQVSPHKSG (SEQ ID NO: 5014), GDDSPHKSV (SEQ ID NO: 5015), GHNSPHKSG (SEQ ID NO: 5016), GNGSPHKRG (SEQ ID NO: 5017), GHDSPHKYG (SEQ ID NO: 5018), GHDSPHKSQ (SEQ ID NO: 5019), GNDSPHKIG (SEQ ID NO: 5020), GHDSPHKSK (SEQ ID NO: 5021), GHDSPHKLW (SEQ ID NO: 5022), GHPSPHWKG (SEQ ID NO: 5023), GHDSPHKMG (SEQ ID NO: 5024), GHDSPHKMA (SEQ ID NO: 5025), or GHSSPHRSG (SEQ ID NO: 5026); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, or 8 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2]-[N3] is or comprises GSGSPHSKA (SEQ ID NO: 4697). In some embodiments, [N1]-[N2]-[N3] is or comprises GHDSPHKSG (SEQ ID NO: 4698).

[0055] In some embodiments, ligand comprising the protein or the peptide comprising an amino acid sequence having the formula [N1]-[N2]-[N3], further comprises [N4] which comprises X7 X8 X9 X10. In some embodiments, position X7 of [N4] is independently chosen from W, Q, K, R, G, L, V, S, P, H, K, I, M, A, E, or F. In some embodiments, position X8 of [N4] is independently chosen from N, Y, C, K, T, H, R, D, V, S, P, G, W, E, F, A, I, M, Q, or L. In some embodiments, position X9 of [N4] is independently chosen from Q, G, K, H, R, T, L, D, A, P, I, F, V, M, W, Y, S, E, N, or Y. In some embodiments, position X10 of [N4] is independently chosen from Q, H, L, R, W, K, A, P, E, M, I, S, G, N, Y, C, V, T, D, or V. In some embodiments [N4] is or comprises QNQQ (SEQ ID NO: 5028), WNQQ (SEQ ID NO: 5029), QYYV (SEQ ID NO: 5030), RRQQ (SEQ ID NO: 5031), QNQQ (SEQ ID NO: 5028), GCGQ (SEQ ID NO: 5032), LRQQ (SEQ ID NO: 5033), RNQQ (SEQ ID NO: 5034), VNQQ (SEQ ID NO: 5035), FRLQ (SEQ ID NO: 5036), FNQQ (SEQ ID NO: 5037), LLQQ (SEQ ID NO: 5038), SNQQ (SEQ ID NO: 5039), RLQQ (SEQ ID NO: 5040), LNQQ (SEQ ID NO: 5041), QRKL (SEQ ID NO: 5042), LRRQ (SEQ ID NO: 5043), QRLR (SEQ ID NO: 5044), QRRL (SEQ ID NO: 5045), RRLQ (SEQ ID NO: 5046), RLRQ (SEQ ID NO: 5047), SKRQ (SEQ ID NO: 5048), QLYR (SEQ ID NO: 5049), QLTV (SEQ ID NO: 5050), QNKQ (SEQ ID NO: 5051), KNQQ (SEQ ID NO: 5052), QKQQ (SEQ ID NO: 5053), QTQQ (SEQ ID NO: 5054), QNHQ (SEQ ID NO: 5055), QHQQ (SEQ ID NO: 5056), QNQH (SEQ ID NO: 5057), QHRQ (SEQ ID NO: 5058), LTQQ (SEQ ID NO: 5059), QNQW (SEQ ID NO: 5060), QNTH (SEQ ID NO: 5061), RRRQ (SEQ ID NO: 5062), QYQQ (SEQ ID NO: 5063), QNDQ (SEQ ID NO: 5064), QNRH (SEQ ID NO: 5065), RDQQ (SEQ ID NO: 5066), PNLQ (SEQ ID NO: 5067), HVRQ (SEQ ID NO: 5068), PNQH (SEQ ID NO: 5069), HNQQ (SEQ ID NO: 5070), QSQQ (SEQ ID NO: 5071), QPAK (SEQ ID NO: 5072), QNLA (SEQ ID NO: 5073), QNQL (SEQ ID NO: 5074), QGQQ (SEQ ID NO: 5075), LNRQ (SEQ ID NO: 5076), QNPP (SEQ ID NO: 5077), QNLQ (SEQ ID NO: 5078), QDQE (SEQ ID NO: 5079), QDQQ (SEQ ID NO: 5080), HWQQ (SEQ ID NO: 5081), PNQQ (SEQ ID NO: 5082), PEQQ (SEQ ID NO: 5083), QRTM (SEQ ID NO: 5084), LHQH (SEQ ID NO: 5085), QHRI (SEQ ID NO: 5086), QYIH (SEQ ID NO: 5087), QKFE (SEQ ID NO: 5088), QFPS (SEQ ID NO: 5089), QNPL (SEQ ID NO: 5090), QAIK (SEQ ID NO: 5091), QNRQ (SEQ ID NO: 5092), QYQH (SEQ ID NO: 5093), QNPQ (SEQ ID NO: 5094), QHQL (SEQ ID NO: 5095), QSPP (SEQ ID NO: 5096), QAKL (SEQ ID NO: 5097), KSQQ (SEQ ID NO: 5098), QDRP (SEQ ID NO: 5099), QNLG (SEQ ID NO: 5100), QAFH (SEQ ID NO: 5101), QNAQ (SEQ ID NO: 5102), HNQL (SEQ ID NO: 5103), QKLN (SEQ ID NO: 5104), QNVQ (SEQ ID NO: 5105), QAQQ (SEQ ID NO: 5106), QTPP (SEQ ID NO: 5107), QPPA (SEQ ID NO: 5108), QERP (SEQ ID NO: 5109), QDLQ (SEQ ID NO: 5110), QAMH (SEQ ID NO: 5111), QHPS (SEQ ID NO: 5112), PGLQ (SEQ ID NO: 5113), QGIR (SEQ ID NO: 5114), QAPA (SEQ ID NO: 5115), QIPP (SEQ ID NO: 5116), QTQL (SEQ ID NO: 5117), QAPS (SEQ ID NO: 5118), QNTY (SEQ ID NO: 5119), QDKQ (SEQ ID NO: 5120), QNHL (SEQ ID NO: 5121), QIGM (SEQ ID NO: 5122), LNKQ (SEQ ID NO: 5123), PNQL (SEQ ID NO: 5124), QLQQ (SEQ ID NO: 5125), QRMS (SEQ ID NO: 5126), QGIL (SEQ ID NO: 5127), QDRQ (SEQ ID NO: 5128),

RDWQ (SEQ ID NO: 5129), QERS (SEQ ID NO: 5130), QNYQ (SEQ ID NO: 5131), QRTC (SEQ ID NO: 5132), QIGH (SEQ ID NO: 5133), QGAI (SEQ ID NO: 5134), QVPP (SEQ ID NO: 5135), QVQQ (SEQ ID NO: 5136), LMRQ (SEQ ID NO: 5137), QYSV (SEQ ID NO: 5138), QAIT (SEQ ID NO: 5139), QKTL (SEQ ID NO: 5140), QLHH (SEQ ID NO: 5141), QNII (SEQ ID NO: 5142), QGHH (SEQ ID NO: 5143), QSKV (SEQ ID NO: 5144), QLPS (SEQ ID NO: 5145), IGKQ (SEQ ID NO: 5146), QAIH (SEQ ID NO: 5147), QHGL (SEQ ID NO: 5148), QFMC (SEQ ID NO: 5149), QNQM (SEQ ID NO: 5150), QHLQ (SEQ ID NO: 5151), QPAR (SEQ ID NO: 5152), QSLQ (SEQ ID NO: 5153), QSQL (SEQ ID NO: 5154), HSQQ (SEQ ID NO: 5155), QMPS (SEQ ID NO: 5156), QGSL (SEQ ID NO: 5157), QVPA (SEQ ID NO: 5158), HYQQ (SEQ ID NO: 5159), QVPS (SEQ ID NO: 5160), RGEQ (SEQ ID NO: 5161), PGQQ (SEQ ID NO: 5162), LEQQ (SEQ ID NO: 5163), QNQS (SEQ ID NO: 5164), QKVI (SEQ ID NO: 5165), QNND (SEQ ID NO: 5166), QSVH (SEQ ID NO: 5167), QPLG (SEQ ID NO: 5168), HNQE (SEQ ID NO: 5169), QIQQ (SEQ ID NO: 5170), QVRN (SEQ ID NO: 5171), PSNQ (SEQ ID NO: 5172), QVGH (SEQ ID NO: 5173), QRDI (SEQ ID NO: 5174), QMPN (SEQ ID NO: 5175), RGLQ (SEQ ID NO: 5176), PSLQ (SEQ ID NO: 5177), QRDQ (SEQ ID NO: 5178), QAKG (SEQ ID NO: 5179), QSAH (SEQ ID NO: 5180), QSTM (SEQ ID NO: 5181), QREM (SEQ ID NO: 5182), QYRA (SEQ ID NO: 5183), QRQQ (SEQ ID NO: 5184), QWQQ (SEQ ID NO: 5185), QRMN (SEQ ID NO: 5186), GDSQ (SEQ ID NO: 5187), QKIS (SEQ ID NO: 5188), PSMQ (SEQ ID NO: 5189), SPRQ (SEQ ID NO: 5190), MEQQ (SEQ ID NO: 5191), QYQN (SEQ ID NO: 5192), QIRQ (SEQ ID NO: 5193), QSVQ (SEQ ID NO: 5194), RSQQ (SEQ ID NO: 5195), QNKL (SEQ ID NO: 5196), QIQH (SEQ ID NO: 5197), PRQQ (SEQ ID NO: 5198), HTQQ (SEQ ID NO: 5199), QRQH (SEQ ID NO: 5200), RNQE (SEQ ID NO: 5201), QSKQ (SEQ ID NO: 5202), QNQP (SEQ ID NO: 5203), QSPQ (SEQ ID NO: 5204), QTRQ (SEQ ID NO: 5205), QNLH (SEQ ID NO: 5206), QNQE (SEQ ID NO: 5207), LNQP (SEQ ID NO: 5208), QNQD (SEQ ID NO: 5209), QNLL (SEQ ID NO: 5210), QLVI (SEQ ID NO: 5211), RTQE (SEQ ID NO: 5212), QTHQ (SEQ ID NO: 5213), QDQH (SEQ ID NO: 5214), QSQH (SEQ ID NO: 5215), VRQQ (SEQ ID NO: 5216), AWQQ (SEQ ID NO: 5217), QSVP (SEQ ID NO: 5218), QNIQ (SEQ ID NO: 5219), LDQQ (SEQ ID NO: 5220), PDQQ (SEQ ID NO: 5221), ESQQ (SEQ ID NO: 5222), QRQL (SEQ ID NO: 5223), QIIV (SEQ ID NO: 5224), QKQS (SEQ ID NO: 5225), QSHQ (SEQ ID NO: 5226), QFVV (SEQ ID NO: 5227), QSQP (SEQ ID NO: 5228), QNEQ (SEQ ID NO: 5229), INQQ (SEQ ID NO: 5230), RNRQ (SEQ ID NO: 5231), RDQK (SEQ ID NO: 5232), QWKR (SEQ ID NO: 5233), ENRQ (SEQ ID NO: 5234), QTQP (SEQ ID NO: 5235), QKQL (SEQ ID NO: 5236), RNQL (SEQ ID NO: 5237), ISIQ (SEQ ID NO: 5238), QTVC (SEQ ID NO: 5239), QQIM (SEQ ID NO: 5240), LNHQ (SEQ ID NO: 5241), QNQA (SEQ ID NO: 5242), QMIH (SEQ ID NO: 5243), RNHQ (SEQ ID NO: 5244), or QKMN (SEQ ID NO: 5245), or any dipeptide or tripeptide thereof. In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises: the amino acid sequence of any of SEQ ID NOs: 1800-2241; an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids, e.g., consecutive amino acids)

thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises GSGSPHSKAQNQQ (SEQ ID NO: 1801). In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises GHDSPHKSGQNQQ (SEQ ID NO: 1800).

[0056] In some embodiments, the ligand comprising the protein or peptide comprising an amino acid sequence having the formula [N1]-[N2]-[N3], further comprises [N0], which comprises XA XB and XC. In some embodiments, XA of [N0] is independently chosen from T, S, Y, M, A, C, I, R, L, D, F, V, Q, N, H, E, or G. In some embodiments, XB of [N0] is independently chosen from I, M, P, E, N, D, S, A, T, G, Q, F, V, L, C, H, R, W, or L. In some embodiments, XC of [N0] is independently chosen from N, M, E, G, Y, W, T, I, Q, F, V, A, L, I, P, K, R, H, S, D, or S. In some embodiments, [N0] is or comprises TIN, SMN, TIM, YLS, GLS, MPE, MEG, MEY, AEW, CEW, ANN, IPE, ADM, IEY, ADY, IET, MEW, CEY, RIN, MEI, LEY, ADW, IEI, DIM, FEQ, MEF, CDQ, LPE, IEN, MES, AEI, VEY, IIN, TSN, IEV, MEM, AEV, MDA, VEW, AEQ, LEW, MEL, MET, MEA, IES, MEV, CEI, ATN, MDG, QEV, ADQ, NMN, IEM, ISN, TGN, QQQ, HDW, IEG, TII, TFP, TEK, EIN, TVN, TFN, SIN, TER, TSY, ELH, AIN, SVN, TDN, TFH, TVH, TEN, TSS, TID, TCN, NIN, TEH, AEM, AIK, TDK, TFK, SDQ, TEI, NTN, TET, SIK, TEL, TEA, TAN, TIY, TFS, TES, TTN, TED, TNN, EVH, TIS, TVR, TDR, TIK, NHI, TIP, ESD, TDL, TVP, TVI, AEH, NCL, TVK, NAD, TIT, NCV, TIR, NAL, VIN, TIQ, TEF, TRE, QGE, SEK, NVN, GGE, EFV, SDK, TEQ, EVQ, TEY, NCW, TDV, SDI, NSI, NSL, EVV, TEP, SEL, TWQ, TEV, AVN, GVL, TLN, TEG, TRD, NAI, AEN, AET, ETA, NNL, or any dipeptide thereof. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises the amino acid sequence of any one of SEQ ID NOs: 2242-2886; an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGSGSPHSKAQNQQ (SEQ ID NO: 2242). In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGHDSPHKSGQNQQ (SEQ ID NO: 2243).

[0057] In some embodiments, [N3] is present immediately subsequent to [N2]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [N2]-[N3]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [N1]-[N2]-[N3]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [N1]-[N2]-[N3]-[N4]. In some

embodiments, the peptide comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]-[N4].

[0058] In some embodiments, a ligand comprises a protein or a peptide comprising an amino acid sequence having the formula [A][B] (SEQ ID NO: 6410), wherein [A] comprises the amino acid sequence of GSGSPH (SEQ ID NO: 4695) and [B] comprises X1 X2 X3 X4 X5 X6 X7. In some embodiments, position X1 of [B] is independently chosen from S, C, F, or V. In some embodiments, position X2 of [B] is independently chosen from K, L, R, I, E, Y, V, or S. In some embodiments, X3 of [B] is independently chosen from A, R, L, G, I, Y, S, F, or W. In some embodiments X4 of [B] is independently chosen from W, Q, R, G, L, V, S, or F. In some embodiments, position X5 of [B] is independently chosen from N, Y, R, C, K, or L. In some embodiments, position X6 of [B] is independently chosen from Q, G, K, R, T, L, or Y. In some embodiment, position X7 of [B] is independently chosen from Q, L, R, or V. In some embodiments, [B] comprises SLLWNQQ (SEQ ID NO: 5247), SKAQYYV (SEQ ID NO: 5248), SKLRRQQ (SEQ ID NO: 5249), SIWQNQQ (SEQ ID NO: 5250), SKAGCGQ (SEQ ID NO: 5251), SRAQNQQ (SEQ ID NO: 5252), SKRLRQQ (SEQ ID NO: 5253), SLRRNQQ (SEQ ID NO: 5254), SRGRNQQ (SEQ ID NO: 5255), SEIVNQQ (SEQ ID NO: 5256), SSRRNQQ (SEQ ID NO: 5257), CLLQNQQ (SEQ ID NO: 5258), SKAFRLQ (SEQ ID NO: 5259), CLAQNQQ (SEQ ID NO: 5260), FLRQNQQ (SEQ ID NO: 5261), SLRFNQQ (SEQ ID NO: 5262), SYLRNQQ (SEQ ID NO: 5263), CSLQNQQ (SEQ ID NO: 5264), VLWQNQQ (SEQ ID NO: 5265), SKWLLQQ (SEQ ID NO: 5266), SLWSNQQ (SEQ ID NO: 5267), SKRRLQQ (SEQ ID NO: 5268), SVYLNQQ (SEQ ID NO: 5269), SLWLNQQ (SEQ ID NO: 5270), SKAQRKL (SEQ ID NO: 5271), SKALRRQ (SEQ ID NO: 5272), SKAQRLR (SEQ ID NO: 5273), SKAQNQQ (SEQ ID NO: 5274), SKAQRRL (SEQ ID NO: 5275), SKARRQQ (SEQ ID NO: 5276), SKARRLQ (SEQ ID NO: 5277), SKSRRQQ (SEQ ID NO: 5278), SKARLRQ (SEQ ID NO: 5279), SKASKRQ (SEQ ID NO: 5280), VRRQNQQ (SEQ ID NO: 5281), SKAQLYR (SEQ ID NO: 5282), SLFRNQQ (SEQ ID NO: 5283), SKAQLTV (SEQ ID NO: 5284), or any dipeptide, tripeptide, tetrapeptide, pentapeptide, or hexapeptide thereof. In some embodiments, [A][B] comprises GSGSPHSLLWNQQ (SEQ ID NO: 5285), GSGSPHSKAQYYV (SEQ ID NO: 2060), GSGSPHSKLRRQQ (SEQ ID NO: 2061), GSGSPHSIWQNQQ (SEQ ID NO: 5286), GSGSPHSKAGCGQ (SEQ ID NO: 2062), GSGSPHSRAQNQQ (SEQ ID NO: 2063), GSGSPHSKRLRQQ (SEQ ID NO: 2064), GSGSPHSLRRNQQ (SEQ ID NO: 2065), GSGSPHSRGRNQQ (SEQ ID NO: 2066), GSGSPHSEIVNQQ (SEQ ID NO: 5287), GSGSPHSSRRNQQ (SEQ ID NO: 2067), GSGSPHCLLQNQQ (SEQ ID NO: 5288), GSGSPHSKAFRLQ (SEQ ID NO: 2068), GSGSPHCLAQNQQ (SEQ ID NO: 5289), GSGSPHFLRQNQQ (SEQ ID NO: 2070), GSGSPHSLRFNQQ (SEQ ID NO: 2071), GSGSPHSYLRNQQ (SEQ ID NO: 5290), GSGSPHCSLQNQQ (SEQ ID NO: 5291), GSGSPHVLWQNQQ (SEQ ID NO: 5292), GSGSPHSKWLLQQ (SEQ ID NO: 2072), GSGSPHSLWSNQQ (SEQ ID NO: 5293), GSGSPHSKRRLQQ (SEQ ID NO: 2073), GSGSPHSVYLNQQ (SEQ ID NO: 5294),

GSGSPHSLWLNQQ (SEQ ID NO: 5295), GSGSPHSKAQRKL (SEQ ID NO: 2074), GSGSPHSKALRRQ (SEQ ID NO: 2075), GSGSPHSKAQRLR (SEQ ID NO: 2076), GSGSPHSKAQNQQ (SEQ ID NO: 1801), GSGSPHSKAQRRL (SEQ ID NO: 2077), GSGSPHSKARRQQ (SEQ ID NO: 2078), GSGSPHSKARRLQ (SEQ ID NO: 2079), GSGSPHSKSRRQQ (SEQ ID NO: 2080), GSGSPHSKARLRQ (SEQ ID NO: 2082), GSGSPHSKASKRQ (SEQ ID NO: 2083), GSGSPHVRRQNQQ (SEQ ID NO: 2084), GSGSPHSKAQLYR (SEQ ID NO: 2085), GSGSPHSLFRNQQ (SEQ ID NO: 5296), GSGSPHSKAQLTV (SEQ ID NO: 2086), or any portion thereof, e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids, e.g., consecutive amino acids, thereof. In some embodiments, [B] is present immediately subsequent to [A]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [A][B].

[0059] In some embodiments, the ligand comprises a protein or peptide comprising an amino acid sequence having the formula [A][B] (SEQ ID NO: 6411), wherein [A] comprises X1 X2 X3 X4 X5 X6 and [B] comprises SPHKSG (SEQ ID NO: 946). In some embodiments, position X1 of [A] is independently chosen from T, M, A, C, I, R, L, D, F, V, Q, N, or H. In some embodiments, position X2 of [A] is independently chosen from I, P, E, N, D, S, A, T, M, or Q. In some embodiments, position X3 of [A] is independently chosen from N, E, G, Y, W, M, T, I, K, Q, F, S, V, A, or L. In some embodiments, position X4 of [A] is independently chosen from G, D, R, or E. In some embodiments, position X5 of [A] is independently chosen from H, Q, N, or D. In some embodiments, position X6 of [A] is independently chosen from D or R. In some embodiments, [A] comprises TINGHD (SEQ ID NO: 5297), MPEGHD (SEQ ID NO: 5298), MEGGHD (SEQ ID NO: 5299), MEYGHD (SEQ ID NO: 5300), AEWGHD (SEQ ID NO: 5301), CEWGHD (SEQ ID NO: 5302), ANNGQD (SEQ ID NO: 5303), IPEGHD (SEQ ID NO: 5304), ADMGHD (SEQ ID NO: 5305), IEYGHD (SEQ ID NO: 5306), ADYGHD (SEQ ID NO: 5307), IETGHD (SEQ ID NO: 5308), MEWGHD (SEQ ID NO: 5309), CEYGHD (SEQ ID NO: 5310), RINGHD (SEQ ID NO: 5311), MEIGHD (SEQ ID NO: 5312), LEYGHD (SEQ ID NO: 5313), ADWGHD (SEQ ID NO: 5314), IEIGHD (SEQ ID NO: 5315), TIKDND (SEQ ID NO: 5316), DIMGHD (SEQ ID NO: 5317), FEQGHD (SEQ ID NO: 5318), MEFGHD (SEQ ID NO: 5319), CDQGHD (SEQ ID NO: 5320), LPEGHD (SEQ ID NO: 5321), IENGHD (SEQ ID NO: 5322), MESGHD (SEQ ID NO: 5323), AEIGHD (SEQ ID NO: 5324), VEYGHD (SEQ ID NO: 5325), TSNGDD (SEQ ID NO: 5326), IEVGHD (SEQ ID NO: 5327), MEMGHD (SEQ ID NO: 5328), AEVGHD (SEQ ID NO: 5329), MDAGHD (SEQ ID NO: 5330), VEWGHD (SEQ ID NO: 5331), AEQGHD (SEQ ID NO: 5332), LEWGHD (SEQ ID NO: 5333), MELGHD (SEQ ID NO: 5334), METGHD (SEQ ID NO: 5335), MEAGHD (SEQ ID NO: 5336), TINRQR (SEQ ID NO: 5337), IESGHD (SEQ ID NO: 5338), TAKDHD (SEQ ID NO: 5339), MEVGHD (SEQ ID NO: 5340), CEIGHD (SEQ ID NO: 5341), ATNGHD (SEQ ID NO: 5342), MDGGHD (SEQ ID NO: 5343), QEVGHD (SEQ ID NO: 5344), ADQGHD (SEQ ID NO: 5345), NMNGHD (SEQ ID NO: 5346), TPWEHD (SEQ ID NO: 5347),

IEMGHD (SEQ ID NO: 5348), TANEHD (SEQ ID NO: 5349), QQQGHD (SEQ ID NO: 5350), TPQDHD (SEQ ID NO: 5351), HDWGHD (SEQ ID NO: 5352), IEGGHD (SEQ ID NO: 5353), or any dipeptide, tripeptide, tetrapeptide, or pentapeptide thereof. In some embodiments, [A][B] comprises TINGHDSPHKR (SEQ ID NO: 5354), MPEGHDSPHKS (SEQ ID NO: 5355), MEGGHDSPHKS (SEQ ID NO: 5356), MEYGHDSPHKS (SEQ ID NO: 5357), AEWGHDSPHKS (SEQ ID NO: 5358), CEWGHDSPHKS (SEQ ID NO: 5359), ANNGQDSPHKS (SEQ ID NO: 5360), IPEGHDSPHKS (SEQ ID NO: 5361), ADMGHDSPHKS (SEQ ID NO: 5362), IEYGHDSPHKS (SEQ ID NO: 5363), ADYGHDSPHKS (SEQ ID NO: 5364), IETGHDSPHKS (SEQ ID NO: 5365), MEWGHDSPHKS (SEQ ID NO: 5366), CEYGHDSPHKS (SEQ ID NO: 5367), RINGHDSPHKS (SEQ ID NO: 5368), MEIGHDSPHKS (SEQ ID NO: 5369), LEYGHDSPHKS (SEQ ID NO: 5370), ADWGHDSPHKS (SEQ ID NO: 5371), IEIGHDSPHKS (SEQ ID NO: 5372), TIKDNDSPHKS (SEQ ID NO: 5373), DIMGHDSPHKS (SEQ ID NO: 5374), FEQGHDSPHKS (SEQ ID NO: 5375), MEFGHDSPHKS (SEQ ID NO: 5376), CDQGHDSPHKS (SEQ ID NO: 5377), LPEGHDSPHKS (SEQ ID NO: 5378), IENGHDSPHKS (SEQ ID NO: 5379), MESGHDSPHKS (SEQ ID NO: 5380), AEIGHDSPHKS (SEQ ID NO: 5381), VEYGHDSPHKS (SEQ ID NO: 5382), TSNGDDSPHKS (SEQ ID NO: 5383), IEVGHDSPHKS (SEQ ID NO: 5384), MEMGHDSPHKS (SEQ ID NO: 5385), AEVGHDSPHKS (SEQ ID NO: 5386), MDAGHDSPHKS (SEQ ID NO: 5387), VEWGHDSPHKS (SEQ ID NO: 5388), AEQGHDSPHKS (SEQ ID NO: 5389), LEWGHDSPHKS (SEQ ID NO: 5390), MELGHDSPHKS (SEQ ID NO: 5391), METGHDSPHKS (SEQ ID NO: 5392), MEAGHDSPHKS (SEQ ID NO: 5393), TINRQRSPHKS (SEQ ID NO: 5394), IESGHDSPHKS (SEQ ID NO: 5395), TAKDHDSPHKS (SEQ ID NO: 5396), MEVGHDSPHKS (SEQ ID NO: 5397), CEIGHDSPHKS (SEQ ID NO: 5398), ATNGHDSPHKS (SEQ ID NO: 5399), MDGGHDSPHKS (SEQ ID NO: 5400), QEVGHDSPHKS (SEQ ID NO: 5401), ADQGHDSPHKS (SEQ ID NO: 5402), NMNGHDSPHKS (SEQ ID NO: 5403), TPWEHDSPHKS (SEQ ID NO: 5404), IEMGHDSPHKS (SEQ ID NO: 5405), TANEHDSPHKS (SEQ ID NO: 5406), TINGHDSPHKS (SEQ ID NO: 5407), QQQGHDSPHKS (SEQ ID NO: 5408), TPQDHDSPHKS (SEQ ID NO: 5409), HDWGHDSPHKS (SEQ ID NO: 5410), IEGGHDSPHKS (SEQ ID NO: 5411), or any portion thereof, e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids, e.g., consecutive amino acids, thereof. In some embodiments, [B] is present immediately subsequent to [A]. In some embodiments, the peptide comprises from N-terminus to C-terminus, [A][B].

[0060] In some embodiments, a ligand described herein comprises a protein or peptide comprising an amino acid sequence comprising at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 consecutive amino acids from any one of the sequences provided in Tables 1, 2A, 2B, 2C, 13-19. In some embodiments, the peptide comprises an amino acid sequence comprising at least 3, 4, or 5 consecutive amino acids from any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the peptide comprises an amino acid sequence comprising at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 consecutive amino acids from any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-

909. In some embodiments, the peptide comprises a modification. In some embodiments, the peptide comprises a phosphate group. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue.

[0061] In some embodiments, the 3 consecutive amino acids comprise SPH. In some embodiments, the 4 consecutive amino acids comprise SPHS (SEQ ID NO: 4700). In some embodiments, the 5 consecutive amino acids comprise SPHSK (SEQ ID NO: 4701). In some embodiments, the 6 consecutive amino acids comprise SPHSKA (SEQ ID NO: 941). In some embodiments, the peptide comprises a modification. In some embodiments, the peptide comprises a phosphate group. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue present at position one, numbered according to SEQ ID NO: 941.

[0062] In some embodiments, 3 consecutive amino acids comprise HDS. In some embodiments, the 4 consecutive amino acids comprise HDSP (SEQ ID NO: 4702). In some embodiments, the 5 consecutive amino acids comprise HDSPH (SEQ ID NO: 4703). In some embodiments, the 6 consecutive amino acids comprise HDSPHK (SEQ ID NO: 2). In some embodiments, the 7 consecutive amino acids comprise HDSPHKS (SEQ ID NO: 4840). In some embodiments, the 8 consecutive amino acids comprise HDSPHKSG (SEQ ID NO: 943).

[0063] In some embodiments, 3 consecutive amino acids comprise HDS. In some embodiments, the 4 consecutive amino acids comprise HDSP (SEQ ID NO: 4702). In some embodiments, the 5 consecutive amino acids comprise HDSPH (SEQ ID NO: 4703). In some embodiments, the 6 consecutive amino acids comprise HDSPHK (SEQ ID NO: 2). In some embodiments, the peptide comprises a modification. In some embodiments, the peptide comprises a phosphate group. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue present at position three, numbered according to SEQ ID NO: 2.

[0064] In some embodiments, the 3 consecutive amino acids comprise SPH. In some embodiments, the 4 consecutive amino acids comprise SPHK (SEQ ID NO: 6398). In some embodiments, the 5 consecutive amino acids comprise SPHKY (SEQ ID NO: 4715). In some embodiments, the 6 consecutive amino acids comprise SPHKYG (SEQ ID NO: 966).

[0065] In some embodiments, a ligand described herein comprises a protein or a peptide comprising an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids, relative to the amino acid sequence of any one of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the peptide comprises an

amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids, relative to the amino acid sequence of any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SEQ ID NO: 3589. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of SEQ ID NO: 3589. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SEQ ID NO: 1754. In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of SEQ ID NO: 1754.

[0066] In some embodiments, a ligand described herein comprises a protein or a peptide comprising an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SPHSKA (SEQ ID NO: 941). In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of SPHSKA (SEQ ID NO: 941).

[0067] In some embodiments, a ligand described herein comprises a protein or a peptide comprising comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of HDSPHKSG (SEQ ID NO: 943). In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of HDSPHKSG (SEQ ID NO: 943).

[0068] In some embodiments, a ligand described herein comprises a protein or a peptide comprising comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions,

relative to the amino acid sequence of HDSPHK (SEQ ID NO: 2). In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of HDSPHK (SEQ ID NO: 2).

[0069] In some embodiments, a ligand described herein comprises a protein or a peptide comprising comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SPHKYG (SEQ ID NO: 966). In some embodiments, the peptide comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids relative to the amino acid sequence of SPHKYG (SEQ ID NO: 966).

[0070] In some embodiments, a ligand described herein comprises a protein or a peptide comprising the amino acid sequence of any of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the peptide comprises the amino acid sequence of any of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the peptide comprises the amino acid sequence of any of S EQ ID NOs: 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the peptide comprises the amino acid sequence of SEQ ID NO: 941. In some embodiments, the peptide comprises the amino acid sequence of SEQ ID NO: 943. In some embodiments, the peptide comprises the amino acid sequence of SEQ ID NO: 2. In some embodiments, the peptide comprises the amino acid sequence of SEQ ID NO: 3589. In some embodiments, the peptide comprises the amino acid sequence of SEQ ID NO: 1754.

[0071] In some embodiments, a ligand described herein comprises a protein or a peptide comprising comprises an amino acid sequence encoded by a nucleotide sequence described herein, e.g., a nucleotide sequence of Table 2A. In some embodiments, the peptide comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 942. In some embodiments, the peptide comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides, relative to the nucleotide sequence of SEQ ID NO: 942. In some embodiments, the peptide comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 942, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the peptide comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 944. In some embodiments, the peptide comprises an amino

acid sequence encoded by a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides, relative to the nucleotide sequence of SEQ ID NO: 944. In some embodiments, the peptide comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 944, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto.

[0072] In some embodiments, a ligand described herein comprises a protein or a peptide, which comprises a modification. In some embodiments, the peptide comprises a phosphate group. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue present at position three, numbered according to SEQ ID NO: 2. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue present at position one, numbered according to SEQ ID NO: 941. In some embodiments, the peptide comprises a modification, e.g., a phosphate group, on a serine residue present in the amino acid sequence of SPH.

[0073] In some embodiments, the nucleotide sequence encoding a peptide of a ligand described herein comprises a nucleotide sequence described herein, e.g., as described in Table 2A. In some embodiments, the nucleotide sequence encoding a peptide described herein is codon optimized. In some embodiments, the nucleotide sequence encoding a peptide described herein is isolated, e.g., recombinant.

[0074] In some embodiments the nucleotide sequence encoding a peptide of a ligand described herein comprises the nucleotide sequence of SEQ ID NO: 942, or a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 942. In some embodiments, the nucleotide sequence encoding a peptide described herein comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides, relative to the nucleotide sequence of SEQ ID NO: 942. In some embodiments the nucleic acid sequence encoding a peptide described herein comprises a nucleotide sequence comprising the nucleotide sequence of SEQ ID NO: 942, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto.

[0075] In some embodiments, the nucleic acid encoding a peptide of a ligand described herein comprises the nucleotide sequence of SEQ ID NO: 944, or a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 944. In some embodiments, the nucleotide sequence encoding a peptide described herein

comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides, relative to the nucleotide sequence of SEQ ID NO: 944. In some embodiments the nucleic acid encoding a peptide described herein comprises a nucleotide sequence comprising the nucleotide sequence of SEQ ID NO: 944, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto.

[0076] The present disclosure also provides a nucleic acid or polynucleotide encoding any of the peptides described herein and ligands, compositions, AAV capsid variants, AAV particles, vectors, and cells comprising the same.

Antibody Molecules

[0077] In some embodiments, a ligand described herein is or comprises an antibody molecule. In other embodiments, an active agent described herein, e.g., a therapeutic agent or a diagnostic agent, is or comprises an antibody molecule.

[0078] As used herein, the term "antibody molecule" refers to a protein, e.g., an immunoglobulin chain or fragment thereof, comprising at least one immunoglobulin variable domain sequence. The term “antibody molecule” includes, for example, a monoclonal antibody (including a full length antibody which has an immunoglobulin Fc region). In an embodiment, an antibody molecule comprises a full length antibody, or a full length immunoglobulin chain. In an embodiment, an antibody molecule comprises an antigen binding or functional fragment of a full length antibody, or a full length immunoglobulin chain.

[0079] In an embodiment, an antibody molecule is a monospecific antibody molecule and binds a single epitope, e.g., a monospecific antibody molecule having a plurality of immunoglobulin variable domain sequences, each of which binds the same epitope.

[0080] In an embodiment an antibody molecule is a multispecific antibody molecule, e.g., it comprises a plurality of immunoglobulin variable domains sequences, wherein a first immunoglobulin variable domain sequence of the plurality has binding specificity for a first epitope and a second immunoglobulin variable domain sequence of the plurality has binding specificity for a second epitope. In an embodiment the first and second epitopes are on the same antigen, e.g., the same protein (or subunit of a multimeric protein). In an embodiment the first and second epitopes overlap. In an embodiment the first and second epitopes do not overlap. In an embodiment the first and second epitopes are on different antigens, e.g., the different proteins (or different subunits of a multimeric protein). In an embodiment a multispecific antibody molecule comprises a third, fourth or fifth immunoglobulin variable domain. In an embodiment, a multispecific antibody molecule is a bispecific antibody molecule, a trispecific antibody molecule, or tetraspecific antibody molecule.

[0081] In an embodiment a multispecific antibody molecule is a bispecific antibody molecule. A bispecific antibody has specificity for no more than two antigens. A bispecific antibody molecule is

characterized by a first immunoglobulin variable domain sequence which has binding specificity for a first epitope and a second immunoglobulin variable domain sequence that has binding specificity for a second epitope. In an embodiment the first and second epitopes are on the same antigen, *e.g.*, the same protein (or subunit of a multimeric protein). In an embodiment the first and second epitopes overlap. In an embodiment the first and second epitopes do not overlap. In an embodiment the first and second epitopes are on different antigens, *e.g.*, the different proteins (or different subunits of a multimeric protein). In an embodiment a bispecific antibody molecule comprises a heavy chain variable domain sequence and a light chain variable domain sequence which have binding specificity for a first epitope and a heavy chain variable domain sequence and a light chain variable domain sequence which have binding specificity for a second epitope. In an embodiment a bispecific antibody molecule comprises a half antibody having binding specificity for a first epitope and a half antibody having binding specificity for a second epitope. In an embodiment a bispecific antibody molecule comprises a half antibody, or fragment thereof, having binding specificity for a first epitope and a half antibody, or fragment thereof, having binding specificity for a second epitope. In an embodiment a bispecific antibody molecule comprises a scFv, or fragment thereof, have binding specificity for a first epitope and a scFv, or fragment thereof, have binding specificity for a second epitope.

[0082] In some embodiments, the antibody molecule comprises at least one immunoglobulin variable domain sequence. An antibody molecule may include, for example, full-length, mature antibodies and antigen-binding fragments of an antibody. For example, an antibody molecule can include a heavy (H) chain variable domain sequence (abbreviated herein as VH), and a light (L) chain variable domain sequence (abbreviated herein as VL). In another example, an antibody molecule includes two heavy (H) chain variable domain sequences and two light (L) chain variable domain sequence, thereby forming two antigen binding sites, such as Fab, Fab', F(ab')2, Fc, Fd, Fd', Fv, single chain antibodies (scFv for example), single variable domain antibodies, diabodies (Dab) (bivalent and bispecific), and chimeric (e.g., humanized) antibodies, which may be produced by the modification of whole antibodies or those synthesized de novo using recombinant DNA technologies. These functional antibody fragments retain the ability to selectively bind with their respective antigen or receptor. Antibodies and antibody fragments can be from any class of antibodies including, but not limited to, IgG, IgA, IgM, IgD, and IgE, and from any subclass (e.g., IgG1, IgG2, IgG3, and IgG4) of antibodies. The antibody molecules can be monoclonal or polyclonal. The encoded antibody can also be a human, humanized, CDR-grafted, or in vitro generated antibody. The antibody can have a heavy chain constant region chosen from, e.g., IgG1, IgG2, IgG3, or IgG4. The antibody can also have a light chain chosen from, e.g., kappa or lambda.

[0083] Examples of antigen-binding fragments include: (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a F(ab')2 fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment

consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a diabody (dAb) fragment, which consists of a VH domain; (vi) a camelid or camelized variable domain; (vii) a single chain Fv (scFv), see e.g., Bird et al. (1988) Science 242:423-426; and Huston et al. (1988) Proc. Natl. Acad. Sci. USA 85:5879-5883); and (viii) a single domain antibody. These antibody fragments are obtained using conventional techniques known to those with skill in the art, and the fragments are screened for utility in the same manner as are intact antibodies. An antibody fragment can also be incorporated into single domain antibodies, maxibodies, minibodies, nanobodies, intrabodies, diabodies, triabodies, tetrabodies, v-NAR and bis-scFv (see, for example, Hollinger and Hudson, Nature Biotechnology 23:1126-1136, 2005).

[0084] The term "antibody" includes intact molecules as well as functional fragments thereof. Constant regions of the antibodies can be altered, *e.g.*, mutated, to modify the properties of the antibody (*e.g.*, to increase or decrease one or more of: Fc receptor binding, antibody glycosylation, the number of cysteine residues, effector cell function, or complement function).

[0085] In some embodiments, the antibody molecule can be single domain antibody. Single domain antibodies can include antibodies whose complementary determining regions are part of a single domain polypeptide. Examples include, but are not limited to, heavy chain antibodies, antibodies naturally devoid of light chains, single domain antibodies derived from conventional 4-chain antibodies, engineered antibodies and single domain scaffolds other than those derived from antibodies. Single domain antibodies may be any of the art, or any future single domain antibodies. Single domain antibodies may be derived from any species including, but not limited to mouse, human, camel, llama, fish, shark, goat, rabbit, and bovine. According to another aspect of the invention, a single domain antibody is a naturally occurring single domain antibody known as heavy chain antibody devoid of light chains. Such single domain antibodies are disclosed in WO 9404678, for example. For clarity reasons, this variable domain derived from a heavy chain antibody naturally devoid of light chain is known herein as a VHH or nanobody to distinguish it from the conventional VH of four chain immunoglobulins. Such a VHH molecule can be derived from antibodies raised in Camelidae species, for example in camel, llama, dromedary, alpaca and guanaco. Other species besides Camelidae may produce heavy chain antibodies naturally devoid of light chain; such VHHs are within the scope of the invention.

[0086] In some embodiments, the VH and VL regions of the antibody molecule can be subdivided into regions of hypervariability, termed "complementarity determining regions" (CDR), interspersed with regions that are more conserved, termed "framework regions" (FR or FW).

[0087] The extent of the framework region and CDRs has been precisely defined by a number of methods (see, Kabat, E. A., et al. (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242; Chothia, C. et al. (1987) J. Mol. Biol. 196:901-917; and the AbM definition used by Oxford Molecular's AbM

antibody modeling software. See, generally, e.g., Protein Sequence and Structure Analysis of Antibody Variable Domains. In: Antibody Engineering Lab Manual (Ed.: Duebel, S. and Kontermann, R., Springer-Verlag, Heidelberg).

[0088] "Complementarity determining region", and "CDR", as used herein refer to the sequences of amino acids within antibody variable regions which confer antigen specificity and binding affinity. In general, there are three CDRs in each heavy chain variable region (HCDR1, HCDR2, HCDR3) and three CDRs in each light chain variable region (LCDR1, LCDR2, LCDR3).

[0089] The precise amino acid sequence boundaries of a given CDR can be determined using any of a number of well-known schemes, including those described by Kabat et al. (1991), "Sequences of Proteins of Immunological Interest," 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD (Kabat numbering scheme), Al-Lazikani et al., (1997) JMB 273,927-948 (Chothia numbering scheme). In some embodiments, the CDRs defined according the Chothia number scheme are also sometimes referred to as hypervariable loops.

[0090] For example, under Kabat, the CDR amino acid residues in the heavy chain variable domain (VH) are numbered 31-35 (HCDR1), 50-65 (HCDR2), and 95-102 (HCDR3); and the CDR amino acid residues in the light chain variable domain (VL) are numbered 24-34 (LCDR1), 50-56 (LCDR2), and 89-97 (LCDR3). Under Chothia the CDR amino acids in the VH are numbered 26-32 (HCDR1), 52-56 (HCDR2), and 95-102 (HCDR3); and the amino acid residues in VL are numbered 26-32 (LCDR1), 50-52 (LCDR2), and 91-96 (LCDR3). By combining the CDR definitions of both Kabat and Chothia, the CDRs consist of amino acid residues 26-35 (HCDR1), 50-65 (HCDR2), and 95-102 (HCDR3) in human VH and amino acid residues 24-34 (LCDR1), 50-56 (LCDR2), and 89-97 (LCDR3) in human VL.

[0091] In some embodiments, the antigen binding domain of the antibody molecules of the present disclosure is the part of the antibody molecule that comprises determinants that form an interface that binds a therapeutic protein or an epitope thereof. With respect to proteins (or protein mimetics), the antigen-binding site typically includes one or more loops (of at least four amino acids or amino acid mimics) that form an interface that binds to the therapeutic protein. Typically, the antigen-binding site of an antibody molecule includes at least one or two CDRs and/or hypervariable loops, or more typically at least three, four, five or six CDRs and/or hypervariable loops.

[0092] The antibody molecule can be a monoclonal antibody molecule or a polyclonal antibody molecule. In some embodiments, a monoclonal antibody or a monoclonal antibody composition refer to a preparation of antibody molecules of single molecular composition. A monoclonal antibody composition displays a single binding specificity and affinity for a particular epitope. A monoclonal antibody can be made by hybridoma technology or by methods that do not use hybridoma technology (e.g., recombinant methods).

[0093] In some embodiments, the sequences of an antibody molecule to be included in an encoded payload described herein can be generated by recombinant libraries, e.g., generated by phage display or by combinatorial methods.

[0094] Phage display and combinatorial methods for generating antibodies are known in the art (as described in, e.g., Ladner et al. U.S. Patent No. 5,223,409; Kang et al. International Publication No. WO 92/18619; Dower et al. International Publication No. WO 91/17271; Winter et al. International Publication WO 92/20791; Markland et al. International Publication No. WO 92/15679; Breitling et al. International Publication WO 93/01288; McCafferty et al. International Publication No. WO 92/01047; Garrard et al. International Publication No. WO 92/09690; Ladner et al. International Publication No. WO 90/02809; Fuchs et al. (1991) Bio/Technology 9:1370-1372; Hay et al. (1992) Hum Antibod Hybridomas 3:81-85; Huse et al. (1989) Science 246:1275-1281; Griffths et al. (1993) EMBO J 12:725-734; Hawkins et al. (1992) J Mol Biol 226:889-896; Clackson et al. (1991) Nature 352:624-628; Gram et al. (1992) PNAS 89:3576-3580; Garrad et al. (1991) Bio/Technology 9:1373-1377; Hoogenboom et al. (1991) Nuc Acid Res 19:4133-4137; and Barbas et al. (1991) PNAS 88:7978-7982, the contents of all of which are incorporated by reference herein).

[0095] In some embodiments, the sequences of an antibody molecule to be included in an encoded payload described herein can be generated from an antibody molecule that is designed using the VERSITOPE™ Antibody Generation or BIOATLA®, e.g., in US20130303399, US20130281303, WO2012009026, WO2016033331, WO2016036916, and US8859467, the contents of which are herein incorporated by reference in their entirety. In some embodiments, the sequences of an antibody molecule to be included in an encoded payload described herein can be derived from an antibody molecule that is designed and/or produced using the methods described, e.g., in WO2017189959 and WO2020223276, the contents of which are herein incorporated by reference in their entirety.

[0096] In some embodiments, the antibody molecule comprises an amino acid sequence of a fully human antibody (e.g., an antibody made in a mouse which has been genetically engineered to produce an antibody from a human immunoglobulin sequence), or a non-human antibody, e.g., a rodent (mouse or rat), goat, primate (e.g., monkey), camel antibody. Preferably, the non-human antibody is a rodent (mouse or rat antibody). Methods of producing rodent antibodies are known in the art.

[0097] Human monoclonal antibodies can be generated using transgenic mice carrying the human immunoglobulin genes rather than the mouse system. Splenocytes from these transgenic mice immunized with the antigen of interest are used to produce hybridomas that secrete human mAbs with specific affinities for epitopes from a human protein (see, e.g., Wood et al. International Application WO 91/00906, Kucherlapati et al. PCT publication WO 91/10741; Lonberg et al. International Application WO 92/03918; Kay et al. International Application 92/03917; Lonberg, N. et al. 1994 Nature 368:856-859; Green, L.L. et al. 1994 Nature Genet. 7:13-21; Morrison, S.L. et al. 1994 Proc.

Natl. Acad. Sci. USA 81:6851-6855; Bruggeman et al. 1993 Year Immunol 7:33-40; Tuaillon et al. 1993 PNAS 90:3720-3724; Bruggeman et al. 1991 Eur J Immunol 21:1323-1326).

[0098] In some embodiments, the antibody comprises an amino acid sequence of an antibody in which the variable region, or a portion thereof, e.g., the CDRs, are generated in a non-human organism, e.g., a rat or mouse. Antibody molecules comprising chimeric, CDR-grafted, and humanized antibodies are within the invention. Antibody molecules comprising the sequences of antibodies generated in a non-human organism, e.g., a rat or mouse, and then modified, e.g., in the variable framework or constant region, to decrease antigenicity in a human are within the invention.

[0099] An effectively human protein is a protein that does substantially not evoke a neutralizing antibody response, e.g., the human anti-murine antibody (HAMA) response. HAMA can be problematic in a number of circumstances, e.g., if the antibody molecule is administered repeatedly, e.g., in treatment of a chronic or recurrent disease condition. A HAMA response can make repeated antibody administration potentially ineffective because of an increased antibody clearance from the serum (see, e.g., Saleh et al., Cancer Immunol. Immunother., 32:180-190 (1990)) and also because of potential allergic reactions (see, e.g., LoBuglio et al., Hybridoma, 5:5117-5123 (1986)).

[0100] Chimeric antibodies can be produced by recombinant DNA techniques known in the art (see Robinson et al., International Patent Publication PCT/US86/02269; Akira, et al., European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al., European Patent Application 173,494; Neuberger et al., International Application WO 86/01533; Cabilly et al. U.S. Patent No. 4,816,567; Cabilly et al., European Patent Application 125,023; Better et al. (1988 Science 240:1041-1043); Liu et al. (1987) PNAS 84:3439-3443; Liu et al., 1987, J. Immunol. 139:3521-3526; Sun et al. (1987) PNAS 84:214-218; Nishimura et al., 1987, Canc. Res. 47:999-1005; Wood et al. (1985) Nature 314:446-449; and Shaw et al., 1988, J. Natl Cancer Inst. 80:1553-1559).

[0101] A humanized or CDR-grafted antibody will have at least one or two but generally all three recipient CDRs (of heavy and or light immuoglobulin chains) replaced with a donor CDR. The antibody may be replaced with at least a portion of a non-human CDR or only some of the CDRs may be replaced with non-human CDRs. Preferably, the donor will be a rodent antibody, e.g., a rat or mouse antibody, and the recipient will be a human framework or a human consensus framework. Typically, the immunoglobulin providing the CDRs is called the donor and the immunoglobulin providing the framework is called the acceptor. In some embodiments, the donor immunoglobulin is a non-human (e.g., rodent). The acceptor framework is a naturally-occurring (e.g., a human) framework or a consensus framework, or a sequence about 85% or higher, preferably 90%, 95%, 99% or higher identical thereto.

[0102] In some embodiments, the consensus sequence refers to the sequence formed from the most frequently occurring amino acids (or nucleotides) in a family of related sequences (See e.g., Winnaker, From Genes to Clones (Verlagsgesellschaft, Weinheim, Germany 1987). In a family of

proteins, each position in the consensus sequence is occupied by the amino acid occurring most frequently at that position in the family. If two amino acids occur equally frequently, either can be included in the consensus sequence. In some embodiments, the consensus framework refers to the framework region in the consensus immunoglobulin sequence.

[0103] An antibody can be humanized by methods known in the art (see e.g., Morrison, S. L., 1985, Science 229:1202-1207, by Oi et al., 1986, BioTechniques 4:214, and by Queen et al. US 5,585,089, US 5,693,761 and US 5,693,762, the contents of all of which are hereby incorporated by reference).

[0104] Humanized or CDR-grafted antibodies can be produced by CDR-grafting or CDR substitution, wherein one, two, or all CDRs of an immunoglobulin chain can be replaced. See e.g., U.S. Patent 5,225,539; Jones et al. 1986 Nature 321:552-525; Verhoeyan et al. 1988 Science 239:1534; Beidler et al. 1988 J. Immunol. 141:4053-4060; Winter US 5,225,539, the contents of all of which are hereby expressly incorporated by reference. Winter describes a CDR-grafting method which may be used to prepare the humanized antibodies of the present invention (UK Patent Application GB 2188638A, filed on March 26, 1987; Winter US 5,225,539), the contents of which is expressly incorporated by reference.

[0105] In some embodiments, the antibodies comprise the sequences of humanized antibodies in which specific amino acids have been substituted, deleted or added. Criteria for selecting amino acids from the donor are described in US 5,585,089, e.g., columns 12-16 of US 5,585,089, e.g., columns 12-16 of US 5,585,089, the contents of which are hereby incorporated by reference. Other techniques for humanizing antibodies are described in Padlan et al. EP 519596 A1, published on December 23, 1992.

[0106] In some embodiments, the antibody molecule can be a single chain antibody. A single-chain antibody (scFV) may be engineered (see, for example, Colcher, D. et al. (1999) Ann N Y Acad Sci 880:263-80; and Reiter, Y. (1996) Clin Cancer Res 2:245-52). The single chain antibody can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target protein.

[0107] In yet other embodiments, the antibody molecule has a heavy chain constant region chosen from, e.g., the heavy chain constant regions of IgG1, IgG2, IgG3, IgG4, IgM, IgA1, IgA2, IgD, and IgE; particularly, chosen from, e.g., the (e.g., human) heavy chain constant regions of IgG1, IgG2, IgG3, and IgG4. In another embodiment, the antibody molecule has a light chain constant region chosen from, e.g., the (e.g., human) light chain constant regions of kappa or lambda. The constant region can be altered, e.g., mutated, to modify the properties of the antibody (e.g., to increase or decrease one or more of: Fc receptor binding, antibody glycosylation, the number of cysteine residues, effector cell function, and/or complement function). In some embodiments the antibody has: effector function; and can fix complement. In other embodiments the antibody does not recruit effector cells;

or fix complement. In other embodiments, the antibody has reduced or no ability to bind an Fc receptor. For example, it is an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor, e.g., it has a mutagenized or deleted Fc receptor binding region.

[0108] Methods for altering an antibody constant region are known in the art. Antibodies with altered function, e.g. altered affinity for an effector ligand, such as FcR on a cell, or the C1 component of complement can be produced by replacing at least one amino acid residue in the constant portion of the antibody with a different residue (see e.g., EP 388,151 A1, U.S. Pat. No. 5,624,821 and U.S. Pat. No. 5,648,260, the contents of all of which are hereby incorporated by reference). Similar type of alterations could be described which if applied to the murine, or other species immunoglobulin would reduce or eliminate these functions.

[0109] An antibody molecule can be derivatized or linked to another functional molecule (e.g., another peptide or protein). As used herein, a "derivatized" antibody molecule is one that has been modified. Methods of derivatization include but are not limited to the addition of a fluorescent moiety, a radionucleotide, a toxin, an enzyme or an affinity ligand such as biotin. Accordingly, the antibody molecules of the invention are intended to include derivatized and otherwise modified forms of the antibodies described herein, including immunoadhesion molecules. For example, an antibody molecule can be functionally linked (by chemical coupling, genetic fusion, noncovalent association or otherwise) to one or more other molecular entities, such as another antibody (e.g., a bispecific antibody or a diabody), a detectable agent, a cytotoxic agent, a pharmaceutical agent, and/or a protein or peptide that can mediate association of the antibody or antibody portion with another molecule (such as a streptavidin core region or a polyhistidine tag).

[0110] One type of derivatized antibody molecule is produced by crosslinking two or more antibodies (of the same type or of different types, e.g., to create bispecific antibodies). Suitable crosslinkers include those that are heterobifunctional, having two distinctly reactive groups separated by an appropriate spacer (e.g., m-maleimidobenzoyl-N-hydroxysuccinimide ester) or homobifunctional (e.g., disuccinimidyl suberate). Such linkers are available from Pierce Chemical Company, Rockford, Ill.

[0111] Useful detectable agents with which an antibody molecule of the invention may be derivatized (or labeled) to include fluorescent compounds, various enzymes, prosthetic groups, luminescent materials, bioluminescent materials, fluorescent emitting metal atoms, e.g., europium (Eu), and other anthanides, and radioactive materials (described below). Exemplary fluorescent detectable agents include fluorescein, fluorescein isothiocyanate, rhodamine, 5dimethylamine-1-napthalenesulfonyl chloride, phycoerythrin and the like. An antibody may also be derivatized with detectable enzymes, such as alkaline phosphatase, horseradish peroxidase, β-galactosidase, acetylcholinesterase, glucose oxidase and the like. When an antibody is derivatized with a detectable enzyme, it is detected by adding additional reagents that the enzyme uses to produce a detectable

reaction product. For example, when the detectable agent horseradish peroxidase is present, the addition of hydrogen peroxide and diaminobenzidine leads to a colored reaction product, which is detectable. An antibody molecule may also be derivatized with a prosthetic group (e.g., streptavidin/biotin and avidin/biotin). For example, an antibody may be derivatized with biotin, and detected through indirect measurement of avidin or streptavidin binding. Examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; and examples of bioluminescent materials include luciferase, luciferin, and aequorin.

[0112] Labeled antibody molecule can be used, for example, diagnostically and/or experimentally in a number of contexts, including (i) to isolate a predetermined antigen by standard techniques, such as affinity chromatography or immunoprecipitation; (ii) to detect a predetermined antigen (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the protein; (iii) to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen.

[0113] An antibody molecules may be conjugated to another molecular entity, typically a label or a therapeutic (e.g., a cytotoxic or cytostatic) agent or moiety. Radioactive isotopes can be used in diagnostic or therapeutic applications. Radioactive isotopes that can be coupled to the antibodies described herein include, but are not limited to α-, β-, or γ-emitters, or β-and γ-emitters. Such radioactive isotopes include, but are not limited to iodine (^{131}I or ^{125}I), yttrium (^{90}Y), lutetium (^{177}Lu), actinium (^{225}Ac), praseodymium, astatine (^{211}At), rhenium (^{186}Re), bismuth (^{212}Bi or ^{213}Bi), indium (^{111}In), technetium (^{99}mTc), phosphorus (^{32}P), rhodium (^{188}Rh), sulfur (^{35}S) , carbon (^{14}C), tritium (^{3}H), chromium (^{51}Cr), chlorine (^{36}Cl), cobalt (^{57}Co or ^{58}Co), iron (^{59}Fe), selenium (^{75}Se), or gallium (^{67}Ga). Radioisotopes useful as therapeutic agents include yttrium (^{90}Y), lutetium (^{177}Lu), actinium (^{225}Ac), praseodymium, astatine (^{211}At), rhenium (^{186}Re), bismuth (^{212} Bi or ^{213}Bi), and rhodium (^{188}Rh). Radioisotopes useful as labels, *e.g.*, for use in diagnostics, include iodine (^{131}I or ^{125}I), indium (^{111}In), technetium (^{99}mTc), phosphorus (^{32}P), carbon (^{14}C), and tritium (^{3}H), or one or more of the therapeutic isotopes listed above.

[0114] The invention provides radiolabeled antibody molecules and methods of labeling the same. In one embodiment, a method of labeling an antibody molecule is disclosed. The method includes contacting an antibody molecule, with a chelating agent, to thereby produce a conjugated antibody. The conjugated antibody is radiolabeled with a radioisotope, e.g., ^{111}Indium, ^{90}Yttrium and ^{177}Lutetium, to thereby produce a labeled antibody molecule.

[0115] As is discussed above, the antibody molecule can be conjugated to a therapeutic agent. Therapeutically active radioisotopes have already been mentioned. Examples of other therapeutic agents include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide,

tenoposide, vincristine, vinblastine, colchicine, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, puromycin, maytansinoids, e.g., maytansinol (see U.S. Pat. No. 5,208,020), CC-1065 (see U.S. Pat. Nos. 5,475,092, 5,585,499, 5,846, 545) and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, CC-1065, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclinies (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine, vinblastine, taxol and maytansinoids).

[0116] In some embodiments, a ligand described herein is or comprises an antibody molecule that binds to a GPI-anchored protein. In some embodiments, the antibody molecule binds to ALPL, e.g., human or murine ALPL. In some embodiments, the antibody molecule is F2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, 2F4, or a variant thereof. In some embodiments the antibody molecule is an antibody provided in Table 40 or a variant thereof, e.g., Ab 9 of Table 40.

Multispecific Antibody Molecules

[0117] In some embodiments, the antibody molecule is a multispecific antibody molecule, e.g., it comprises a plurality of immunoglobulin variable domains sequences, wherein a first immunoglobulin variable domain sequence of the plurality has binding specificity for a first epitope and a second immunoglobulin variable domain sequence of the plurality has binding specificity for a second epitope. In some embodiments, the first and second epitopes are on the same antigen, e.g., the same protein (or subunit of a multimeric protein). In some embodiments, the first and second epitopes overlap. In some embodiments, the first and second epitopes do not overlap. In some embodiments, the first and second epitopes are on different antigens, e.g., the different proteins (or different subunits of a multimeric protein). In some embodiments, a multispecific antibody molecule comprises a third, fourth or fifth immunoglobulin variable domain. In some embodiments, a multispecific antibody molecule is a bispecific antibody molecule, a trispecific antibody molecule, or tetraspecific antibody molecule. In some embodiments, an antibody molecule described herein is a multispecific antibody molecule.

[0118] In some embodiments, a multispecific antibody molecule is a bispecific antibody molecule. A bispecific antibody has specificity for no more than two antigens. A bispecific antibody molecule is characterized by a first immunoglobulin variable domain sequence which has binding specificity for a first epitope and a second immunoglobulin variable domain sequence that has binding specificity for a second epitope. In some embodiments, the first and second epitopes are on the same

antigen, e.g., the same protein (or subunit of a multimeric protein). In some embodiments, the first and second epitopes overlap. In some embodiments, the first and second epitopes do not overlap. In some embodiments, the first and second epitopes are on different antigens, e.g., the different proteins (or different subunits of a multimeric protein). In some embodiments, a bispecific antibody molecule comprises a heavy chain variable domain sequence and a light chain variable domain sequence which have binding specificity for a first epitope and a heavy chain variable domain sequence and a light chain variable domain sequence which have binding specificity for a second epitope. In some embodiments, a bispecific antibody molecule comprises a half antibody having binding specificity for a first epitope and a half antibody having binding specificity for a second epitope. In some embodiments, a bispecific antibody molecule comprises a half antibody, or fragment thereof, having binding specificity for a first epitope and a half antibody, or fragment thereof, having binding specificity for a second epitope. In some embodiments, a bispecific antibody molecule comprises a scFv, or fragment thereof, have binding specificity for a first epitope and a scFv, or fragment thereof, have binding specificity for a second epitope. In some embodiment, an antibody molecule as described herein is a bispecific antibody molecule.

[0119] In some embodiments, the sequences of the antibody molecules can be generated from bispecific or heterodimeric antibody molecules produced using protocols known in the art; including but not limited to, for example, the "knob in a hole" approach described in, e.g., US5731168; the electrostatic steering Fc pairing as described in, e.g., WO 09/089004, WO 06/106905 and WO 2010/129304; Strand Exchange Engineered Domains (SEED) heterodimer formation as described in, e.g., WO 07/110205; Fab arm exchange as described in, e.g., WO 08/119353, WO 2011/131746, and WO 2013/060867; double antibody conjugate, e.g., by antibody cross-linking to generate a bi-specific structure using a heterobifunctional reagent having an amine-reactive group and a sulfhydryl reactive group as described in, e.g., US4433059; bispecific antibody determinants generated by recombining half antibodies (heavy-light chain pairs or Fabs) from different antibodies through cycle of reduction and oxidation of disulfide bonds between the two heavy chains, as described in, e.g., US 4444878; trifunctional antibodies, e.g., three Fab' fragments cross-linked through sulfhdryl reactive groups, as described in, e.g., US5273743; biosynthetic binding proteins, e.g., pair of scFvs cross-linked through C-terminal tails preferably through disulfide or amine-reactive chemical cross-linking, as described in, e.g., US5534254; bifunctional antibodies, e.g., Fab fragments with different binding specificities dimerized through leucine zippers (e.g., c-fos and c-jun) that have replaced the constant domain, as described in, e.g., US5582996; bispecific and oligospecific mono-and oligovalent receptors, e.g., VH-CH1 regions of two antibodies (two Fab fragments) linked through a polypeptide spacer between the CH1 region of one antibody and the VH region of the other antibody typically with associated light chains, as described in, e.g., US5591828; bispecific DNA-antibody conjugates, e.g., crosslinking of antibodies or Fab fragments through a double stranded piece of DNA, as described in, e.g., US5635602; bispecific fusion proteins, e.g., an expression construct containing two scFvs with a

hydrophilic helical peptide linker between them and a full constant region, as described in, e.g., US5637481; multivalent and multispecific binding proteins, e.g., dimer of polypeptides having first domain with binding region of Ig heavy chain variable region, and second domain with binding region of Ig light chain variable region, generally termed diabodies (higher order structures are also disclosed creating bispecific, trispecific, or tetraspecific molecules, as described in, e.g., US5837242; minibody constructs with linked VL and VH chains further connected with peptide spacers to an antibody hinge region and CH3 region, which can be dimerized to form bispecific/multivalent molecules, as described in, e.g., US5837821; VH and VL domains linked with a short peptide linker (e.g., 5 or 10 amino acids) or no linker at all in either orientation, which can form dimers to form bispecific diabodies; trimers and tetramers, as described in, e.g., US5844094; String of VH domains (or VL domains in family members) connected by peptide linkages with crosslinkable groups at the C-terminus further associated with VL domains to form a series of FVs (or scFvs), as described in, e.g., US5864019; and single chain binding polypeptides with both a VH and a VL domain linked through a peptide linker are combined into multivalent structures through non-covalent or chemical crosslinking to form, e.g., homobivalent, heterobivalent, trivalent, and tetravalent structures using both scFv or diabody type format, as described in, e.g., US5869620. Additional exemplary multispecific and bispecific molecules and methods of making the same are found, for example, in US5910573, US5932448, US5959083, US5989830, US6005079, US6239259, US6294353, US6333396, US6476198, US6511663, US6670453, US6743896, US6809185, US6833441, US7129330, US7183076, US7521056, US7527787, US7534866, US7612181, US2002/004587A1, US2002/076406A1, US2002/103345A1, US2003/207346A1, US2003/211078A1, US2004/219643A1, US2004/220388A1, US2004/242847A1, US2005/003403A1, US2005/004352A1, US2005/069552A1, US2005/079170A1, US2005/100543A1, US2005/136049A1, US2005/136051A1, US2005/163782A1, US2005/266425A1, US2006/083747A1, US2006/120960A1, US2006/204493A1, US2006/263367A1, US2007/004909A1, US2007/087381A1, US2007/128150A1, US2007/141049A1, US2007/154901A1, US2007/274985A1, US2008/050370A1, US2008/069820A1, US2008/152645A1, US2008/171855A1, US2008/241884A1, US2008/254512A1, US2008/260738A1, US2009/130106A1, US2009/148905A1, US2009/155275A1, US2009/162359A1, US2009/162360A1, US2009/175851A1, US2009/175867A1, US2009/232811A1, US2009/234105A1, US2009/263392A1, US2009/274649A1, EP346087A2, WO00/06605A2, WO02/072635A2, WO04/081051A1, WO06/020258A2, WO2007/044887A2, WO2007/095338A2, WO2007/137760A2, WO2008/119353A1, WO2009/021754A2, WO2009/068630A1, WO91/03493A1, WO93/23537A1, WO94/09131A1, WO94/12625A2, WO95/09917A1, WO96/37621A2, WO99/64460A1. The contents of the above-referenced applications are incorporated herein by reference in their entireties.

[0120] In some embodiments, a ligand described herein comprises a multispecific, e.g., bispecific, antibody molecule comprising a first binding domain that binds to ALPL (e.g., an anti-ALPL binding domain) and a second binding domain that binds to a therapeutic target.

Fc Polypeptides

[0121] In some embodiments, a ligand described herein comprises an Fc polypeptide. In some embodiments, the ligand is or comprises a first Fc polypeptide. In some embodiments, the ligand is a first Fc polypeptide and the active agent is a second Fc polypeptide.

[0122] In some embodiments, the first Fc polypeptide and the second Fc polypeptide form a dimer. In some embodiments, the first Fc polypeptide and the second Fc polypeptide comprise a dimerization domain, e.g., an interface of a first and second Fc polypeptides. In some embodiments, the dimerization domain is engineered, e.g., mutated, to increase or decrease dimerization, e.g., relative to a non-engineered interface. In some embodiments, the dimerization of the first Fc polypeptide and the second Fc polypeptide is enhanced by providing an Fc interface of the first and a second Fc polypeptides with one or more of: a paired cavity-protuberance ("knob-in-a hole"), an electrostatic interaction, or a strand-exchange, such that a greater ratio of heteromultimer:homomultimer forms, e.g., relative to a non-engineered interface. In some embodiments, the first Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof). In some embodiments, the second Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob). In some embodiments, the first Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof); and the second Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob). In some embodiments, the second Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof). In some embodiments, the first Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob). In some embodiments, the second Fc polypeptide comprises an amino acid substitution chosen from: T366S, L368A, or Y407V (e.g., corresponding to a cavity or hole) (or a combination thereof); and the first Fc polypeptide comprises the amino acid substitution T366W (e.g., corresponding to a protuberance or knob).

[0123] In some embodiments, the first Fc polypeptide, the second Fc polypeptide, or both (i) has reduced affinity, e.g., ablated, affinity for an Fc receptor, e.g., as compared to a reference, wherein the reference is a wild-type Fc receptor; (ii) comprises a mutation at one, two, or all of positions I253 (e.g., I253A), H310 (e.g., H310A or H310Q), and/or H435 (e.g., H435A or H435Q), numbered according to the EU index as in Kabat; (iii) has reduced effector function (e.g., reduced ADCC), compared to a reference wherein the reference is a wild-type Fc receptor; (iv) comprises a mutation at

one, two, three, four, or all of positions L235 (e.g., L235V), F243 (e.g., F243L), R292 (e.g., R292P), Y300 (e.g., Y300L), and P396 (e.g., P396L), numbered according to the EU index as in Kabat. In some embodiments, the first Fc polypeptide, the second Fc polypeptide, or both comprises a half-life extender or an amino acid modification that increases serum half-life (e.g., (i) a Leu at position 428 and a Ser at position 434, or (ii) a Ser or Ala at position 434, according to EU numbering).

[0124] In some embodiments, the ligand comprises a first Fc polypeptide, wherein the first Fc polypeptide comprises a protein or peptide sequence provided herein, e.g., as set forth in any of Tables 1, 2A, 2B, 13-19. In some embodiments, the protein or peptide sequence is present in the CH3 domain of the first Fc polypeptide. In some embodiments, the CH3 domain is modified from a human IgG1, IgG2, IgG3, or IgG4 CH3 domain. In some embodiments, the CH3 domain comprises one, two, three, four, five, six, seven, eight, nine, ten, or eleven substitutions in a set of amino acid positions comprising 380, 384, 386, 387, 388, 389, 390, 413, 415, 416, and 421, according to EU numbering. In some embodiments, the protein or peptide is present at or near the C-terminus of the first Fc polypeptide (e.g., within 20, 30, 40, 50, 60, 70, 80, 90, 100, or more amino acids from the C-terminus of the therapeutic protein, enzyme, or antibody molecule). In some embodiments, the first Fc polypeptide, the second Fc polypeptide or both the first Fc polypeptide and the second Fc polypeptide does not comprise an immunoglobulin heavy and/or light chain variable region sequence or an antigen-binding portion thereof.

[0125] In some embodiments, the second Fc polypeptide is fused or coupled (e.g., directly or indirectly via a linker) to a therapeutic protein or variant thereof (e.g., an enzyme).

Other Exemplary Ligands

[0126] In some embodiments, a ligand described herein comprises a nucleic acid molecule. In some embodiments a ligand described herein comprises an aptamer. In some embodiments the aptamer binds to a GPI anchored protein. In some embodiments, the aptamer binds to ALPL, e.g., human or murine ALPL. In some embodiments the aptamer is or comprises DNA, RNA, modified DNA, modified RNA, or a combination thereof. In some embodiments, the aptamer is fused or coupled to a therapeutic agent chosen from a protein (e.g., an enzyme), an antibody molecule, a nucleic acid molecule (e.g., an RNAi agent), or a small molecule.

[0127] In some embodiments a ligand described herein is or comprises a small molecule. In some embodiments, the small molecule is an inhibitor of ALPL, e.g., a small molecule that interferes with ALPL dimerization. In some embodiments, the small molecule is an aryl sulfonamide, a phosphonate derivative, a pyrazole, a triazole, or an imidazole. In some embodiments, the small molecule is 5-((5-chloro-2-methoxyphenyl)sulfonamido)nicotinamide (SBI-425). In some embodiments, the small molecule is 2,5-Dimethoxy-N-(quinolin-3-yl)benzenesulfonamide (Tissue-Nonspecific Alkaline Phosphatase Inhibitor (TNAPi)).

[0128] In some embodiments, a ligand described herein is present or coupled to a carrier, e.g., an exosome, a microvesicle, or a lipid nanoparticle (LNP). In some embodiments, the carrier is an exosome or LNP. In some embodiments, the ligand is present on the surface of the carrier. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80% of the surface of the carrier comprises at least 1-5, e.g., at least 1, 2, 3, 4, or 5, proteins or peptides comprising an amino acid sequence provided herein, e.g., as set forth in any one of Tables 1, 2A, 2B, or 13-19. In some embodiments, the ligand is conjugated to the surface of the carrier by post-insertion. In some embodiments, the ligand is conjugated to the surface of the carrier via a covalent bond (e.g., using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) chemistry or thiol-maleimide linkage reactions). In some embodiments, the carrier is coupled to a therapeutic agent. In some embodiments, the carrier comprises an RNAi agent, an mRNA, a ribonucleoprotein complex (e.g., a Cas9/gRNA complex), or a circRNA.

AAV serotypes and capsids

[0129] In some embodiments, a ligand described herein is a component of a viral particle, e.g., an AAV particle or a lentivirus. In some embodiments, the ligand is a component of a capsid protein, e.g., and AAV capsid protein described herein.

[0130] In some embodiments, an AAV particle may comprise a capsid protein or variant thereof any natural or recombinant AAV serotype. AAV serotypes may differ in characteristics such as, but not limited to, packaging, tropism, transduction and immunogenic profiles. While not wishing to be bound by theory, it is believed in some embodiments, that the AAV capsid protein, e.g., an AAV capsid variant, can modulate, e.g., direct, AAV particle tropism to a particular tissue.

[0131] In some embodiments, an AAV comprises a small non-enveloped icosahedral capsid virus of the Parvoviridae family and is characterized by a single stranded DNA viral genome. Parvoviridae family viruses consist of two subfamilies: Parvovirinae, which infect vertebrates, and Densovirinae, which infect invertebrates. The Parvoviridae family comprises the Dependovirus genus which includes AAV, capable of replication in vertebrate hosts including, but not limited to, human, primate, bovine, canine, equine, and ovine species.

[0132] In some embodiments, AAV are used as a biological tool due to a relatively simple structure, their ability to infect a wide range of cells (including quiescent and dividing cells) without integration into the host genome and without replicating, and their relatively benign immunogenic profile. The genome of the virus may be manipulated to contain a minimum of components for the assembly of a functional recombinant virus, or viral particle, which is loaded with or engineered to target a particular tissue and express or deliver a desired payload.

[0133] In some embodiments, the AAV, is a naturally occurring (e.g., wild-type) AAV or a recombinant AAV. In some embodiments, the wild-type AAV vector genome is a linear, single-stranded DNA (ssDNA) molecule approximately 5,000 nucleotides (nt) in length. In some

embodiments, inverted terminal repeats (ITRs) cap the viral genome at both the 5' and the 3' end, providing origins of replication for the viral genome. In some embodiments, an AAV viral genome typically comprises two ITR sequences. These ITRs have a characteristic T-shaped hairpin structure defined by a self-complementary region (145nt in wild-type AAV) at the 5' and 3' ends of the ssDNA which form an energetically stable double stranded region. The double stranded hairpin structures comprise multiple functions including, but not limited to, acting as an origin for DNA replication by functioning as primers for the endogenous DNA polymerase complex of the host viral replication cell.

[0134] In some embodiments, the wild-type AAV viral genome further comprises nucleotide sequences for two open reading frames, one for the four non-structural Rep proteins (Rep78, Rep68, Rep52, Rep40, encoded by Rep genes) and one for the three capsid, or structural, proteins (VP1, VP2, VP3, encoded by capsid genes or Cap genes). The Rep proteins are used for replication and packaging, while the capsid proteins are assembled to create the protein shell of the AAV, or AAV capsid polypeptide, e.g., an AAV capsid variant. Alternative splicing and alternate initiation codons and promoters result in the generation of four different Rep proteins from a single open reading frame and the generation of three capsid proteins from a single open reading frame. Though it varies by AAV serotype, as a non-limiting example, for AAV9/hu.14 (SEQ ID NO: 123 of US 7,906,111, the contents of which are herein incorporated by reference in their entirety) VP1 refers to amino acids 1-736, VP2 refers to amino acids 138-736, and VP3 refers to amino acids 203-736. In some embodiments, for any one of the amino acid sequences of SEQ ID NO: 981 or 982, VP1 comprises amino acids 1-742, VP2 comprises amino acids 138-742, and VP3 comprises amino acids 203-742. In other words, VP1 is the full-length capsid sequence, while VP2 and VP3 are shorter components of the whole. As a result, changes in the sequence in the VP3 region, are also changes to VP1 and VP2, however, the percent difference as compared to the parent sequence will be greatest for VP3 since it is the shortest sequence of the three. Though described here in relation to the amino acid sequence, the nucleic acid sequence encoding these proteins can be similarly described. Together, the three capsid proteins assemble to create the AAV capsid protein. While not wishing to be bound by theory, the AAV capsid protein typically comprises a molar ratio of 1:1:10 of VP1:VP2:VP3.

[0135] AAV vectors of the present disclosure may be produced recombinantly and may be based on adeno-associated virus (AAV) reference sequences. In addition to single stranded AAV viral genomes (e.g., ssAAVs), the present disclosure also provides for self-complementary AAV (scAAVs) viral genomes. scAAV vector genomes contain DNA strands which anneal together to form double stranded DNA. By skipping second strand synthesis, scAAVs allow for rapid expression in the transduced cell. In some embodiments, the AAV particle of the present disclosure is an scAAV. In some embodiments, the AAV particle of the present disclosure is an ssAAV.

[0136] Methods for producing and/or modifying AAV particles are disclosed in the art such as pseudotyped AAV vectors (PCT Patent Publication Nos. WO200028004; WO200123001;

WO2004112727; WO2005005610; and WO2005072364, the content of each of which is incorporated herein by reference in its entirety).

[0137] As described herein, the AAV particles of the disclosure comprising an AAV capsid variant, and a viral genome, have enhanced tropism for a cell-type or a tissue, e.g., a CNS cell-type, region, or tissue.

[0138] In some embodiments, an AAV capsid variant described herein allows for blood brain barrier penetration following intravenous administration. In some embodiments, the AAV capsid variant allows for blood brain barrier penetration following intravenous administration, focused ultrasound (FUS), e.g., coupled with the intravenous administration of microbubbles (FUS-MB), or MRI-guided FUS coupled with intravenous administration. In some embodiments the AAV capsid variant allows for increased distribution to a brain region. In some embodiments, the brain region comprises a frontal cortex, sensory cortex, motor cortex, caudate, dentate nucleus, cerebellar cortex, cerebral cortex, brain stem, hippocampus, thalamus, putamen, or a combination thereof. In some embodiments, the AAV capsid variant allows for preferential transduction in a brain region relative to the transduction in the dorsal root ganglia (DRG). In some embodiments, the AAV capsid variant allows for transduction in a non-neuronal cell, e.g., a glial cell (e.g., an astrocyte, an oligodendrocyte, or a combination thereof).

[0139] In some embodiments, an AAV capsid variant allows for increased distribution to a spinal cord region. In some embodiments, the spinal region comprises a cervical spinal cord region, thoracic spinal cord region, and/or lumbar spinal cord region.

[0140] In some embodiments, the AAV capsid variant, is suitable for intramuscular administration and/or transduction of muscle fibers. In some embodiments the AAV capsid variant, allows for increased distribution to a muscle region. In some embodiments, the muscle region comprises a heart muscle, quadriceps muscle, a diaphragm muscle region, or a combination thereof. In some embodiments, the muscle region comprises a heart muscle region, e.g., a heart atrium muscle region or a heart ventricle muscle region.

[0141] In some embodiments, the initiation codon for translation of the AAV VP1 capsid protein, e.g., a capsid variant, described herein may be CTG, TTG, or GTG as described in US Patent No. US8163543, the contents of which are herein incorporated by reference in its entirety.

[0142] The present disclosure refers to structural capsid proteins (including VP1, VP2 and VP3) which are encoded by capsid (Cap) genes. These capsid proteins form an outer protein structural shell (e.g., capsid) of a viral vector such as AAV. VP capsid proteins synthesized from Cap polynucleotides generally include a methionine as the first amino acid in the peptide sequence (Met1), which is associated with the start codon (AUG or ATG) in the corresponding Cap nucleotide sequence. However, it is common for a first-methionine (Met1) residue or generally any first amino acid (AA1) to be cleaved off after or during polypeptide synthesis by protein processing enzymes such as Met-

aminopeptidases. This "Met/AA-clipping" process often correlates with a corresponding acetylation of the second amino acid in the polypeptide sequence (e.g., alanine, valine, serine, threonine, etc.). Met-clipping commonly occurs with VP1 and VP3 capsid proteins but can also occur with VP2 capsid proteins.

[0143] Where the Met/AA-clipping is incomplete, a mixture of one or more (one, two or three) VP capsid proteins comprising the viral capsid may be produced, some of which may include a Met1/AA1 amino acid (Met+/AA+) and some of which may lack a Met1/AA1 amino acid as a result of Met/AA-clipping (Met-/AA-). For further discussion regarding Met/AA-clipping in capsid proteins, see Jin, et al. Direct Liquid Chromatography/Mass Spectrometry Analysis for Complete Characterization of Recombinant Adeno-Associated Virus Capsid Proteins. *Hum Gene Ther Methods*. 2017 Oct. 28(5):255-267; Hwang, et al. N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals. *Science*. 2010 February 19. 327(5968): 973–977; the contents of which are each incorporated herein by reference in its entirety.

[0144] According to the present disclosure, references to capsid proteins, e.g., AAV capsid variants, is not limited to either clipped (Met-/AA-) or unclipped (Met+/AA+) and may, in context, refer to independent capsid proteins, viral capsids comprised of a mixture of capsid proteins, and/or polynucleotide sequences (or fragments thereof) which encode, describe, produce or result in capsid proteins of the present disclosure. A direct reference to a capsid protein or capsid polypeptide (such as VP1, VP2 or VP2) may also comprise VP capsid proteins which include a Met1/AA1 amino acid (Met+/AA+) as well as corresponding VP capsid proteins which lack the Met1/AA1 amino acid as a result of Met/AA-clipping (Met-/AA-).

[0145] Further according to the present disclosure, a reference to a specific SEQ ID NO: (whether a protein or nucleic acid) which comprises or encodes, respectively, one or more capsid proteins which include a Met1/AA1 amino acid (Met+/AA+) should be understood to teach the VP capsid proteins which lack the Met1/AA1 amino acid as upon review of the sequence, it is readily apparent any sequence which merely lacks the first listed amino acid (whether or not Met1/AA1).

[0146] As a non-limiting example, reference to a VP1 polypeptide sequence which is 736 amino acids in length, and which includes a "Met1" amino acid (Met+) encoded by the AUG/ATG start codon may also be understood to teach a VP1 polypeptide sequence which is 735 amino acids in length, and which does not include the "Met1" amino acid (Met-) of the 736 amino acid Met+ sequence. As a second non-limiting example, reference to a VP1 polypeptide sequence which is 736 amino acids in length, and which includes an "AA1" amino acid (AA1+) encoded by any NNN initiator codon may also be understood to teach a VP1 polypeptide sequence which is 735 amino acids in length, and which does not include the "AA1" amino acid (AA1-) of the 736 amino acid AA1+ sequence.

[0147] References to viral capsids formed from VP capsid proteins (such as reference to specific AAV capsid serotypes), can incorporate VP capsid proteins which include a Met1/AA1 amino acid (Met+/AA1+), corresponding VP capsid proteins which lack the Met1/AA1 amino acid as a result of Met/AA1-clipping (Met-/AA1-), and combinations thereof (Met+/AA1+ and Met-/AA1-).

[0148] As a non-limiting example, an AAV capsid serotype can include VP1 (Met+/AA1+), VP1 (Met-/AA1-), or a combination of VP1 (Met+/AA1+) and VP1 (Met-/AA1-). An AAV capsid serotype can also include VP3 (Met+/AA1+), VP3 (Met-/AA1-), or a combination of VP3 (Met+/AA1+) and VP3 (Met-/AA1-); and can also include similar optional combinations of VP2 (Met+/AA1) and VP2 (Met-/AA1-).

AAV Capsid Variant

[0149] In some embodiments, an AAV capsid variant disclosed herein comprises a modification in loop IV of AAV9, e.g., at positions between 449-460, e.g., at position 454 and/or 455, numbered relative to SEQ ID NO: 138, 981, or 982. In some embodiments, loop (e.g., loop IV) is used interchangeably herein with the term variable region (e.g., variable region IV), or VR (e.g., VR-IV). In some embodiments loop IV comprises positions 449-475 (e.g., amino acids KTINGSGQNQQTLKFSVAGPSNMAVQG (SEQ ID NO: 6404)), numbered according to SEQ ID NO: 138. In some embodiments loop IV comprises positions 449-460 (e.g., amino acids KTINGSGQNQQT (SEQ ID NO: 6405)), numbered according to SEQ ID NO: 138.

[0150] The AAV particles and payloads of the disclosure may be delivered to one or more target cells, tissues, organs, or organisms. In some embodiments, the AAV particles of the disclosure demonstrate enhanced tropism for a target cell type, tissue or organ. As a non-limiting example, the AAV particle may have enhanced tropism for cells and tissues of the central or peripheral nervous systems (CNS and PNS, respectively). In some embodiments, an AAV particle of the disclosure may, in addition, or alternatively, have decreased tropism for a cell-type, tissue or organ.

[0151] As demonstrated in the Examples herein below, certain AAV capsid variants described herein show multiple advantages over wild-type AAV9, including (i) increased penetrance through the blood brain barrier following intravenous administration, (ii) wider distribution throughout the multiple brain regions, e.g., frontal cortex, sensory cortex, motor cortex, putamen, thalamus, cerebellar cortex, dentate nucleus, caudate, and/or hippocampus, and/or (iii) elevated payload expression in multiple brain regions. Without wishing to be being bound by theory, it is believed that these advantages may be due, in part, to the dissemination of the AAV capsid variants through the brain vasculature. In some embodiments, the AAV capsids described herein enhance the delivery of a payload to multiple regions of the brain including for example, the frontal cortex, sensory cortex, motor cortex, putamen, thalamus, cerebellar cortex, dentate nucleus, caudate, and/or hippocampus.

[0152] In some embodiments, an AAV particle described herein comprises an AAV capsid variant, e.g., an AAV capsid variant described herein (e.g., an AAV capsid variant comprising a

peptide described herein). In some embodiments, an AAV capsid variant comprises a peptide as set forth in any of Tables 1, 2A, 2B, 13-19.

[0153] In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence having the formula [N1]-[N2]-[N3], wherein [N2] comprises the amino acid sequence of SPH and [N3] comprises X4, X5, and X6, wherein at least one of X4, X5, or X6 is a basic amino acid, e.g., a K or R. In some embodiments, position X4 of [N2] is K. In some embodiments, position X5 of [N2] is K.

[0154] In some embodiments, [N1] comprises X1, X2, and X3, wherein at least one of X1, X2, or X3 is G. In some embodiments, position X1 of [N1] is independently chosen from G, V, R, D, E, M, T, I, S, A, N, L, K, H, P, W, or C. In some embodiments, position X2 of [N1] is independently chosen from: S, V, L, N, D, H, R, P, G, T, I, A, E, Y, M, or Q. In some embodiments, position X3 of [N1] is independently chosen from: G, C, L, D, E, Y, H, V, A, N, P, or S. In some embodiments, [N1] comprises GS, SG, GH, HD, GQ, QD, VS, CS, GR, RG, QS, SH, MS, RN, TS, IS, GP, ES, SS, GN, AS, NS, LS, GG, KS, GT, PS, RS, GI, WS, DS, ID, GL, DA, DG, ME, EN, KN, KE, AI, NG, PG, TG, SV, IG, LG, AG, EG, SA, YD, HE, HG, RD, ND, PD, MG, QV, DD, HN, HP, GY, GM, GD, or HS. In some embodiments, [N1] comprises GS, SG, GH, or HD. In some embodiments [N1] is or comprises GSG, GHD, GQD, VSG, CSG, CSH, GQS, GRG, GSH, RVG, GSC, GLL, GDD, GHE, GNY, MSG, RNG, TSG, ISG, GPG, ESG, SSG, GNG, ASG, NSG, LSG, GGG, KSG, HSG, GTG, PSG, GSV, RSG, GIG, WSG, DSG, IDG, GLG, DAG, DGG, MEG, ENG, GSA, KNG, KEG, AIG, GYD, GHG, GRD, GND, GPD, GMG, GQV, GHN, GHP, or GHS. In some embodiments, [N1] is or comprises GSG. In some embodiments, [N1] is or comprises GHD. In some embodiments, [N1]-[N2] comprises SGSPH (SEQ ID NO: 4752), HDSPH (SEQ ID NO: 4703), QDSPH (SEQ ID NO: 4753), RGSPH (SEQ ID NO: 4754), SHSPH (SEQ ID NO: 4755), QSSPH (SEQ ID NO: 4756), DDSPH (SEQ ID NO: 4757), HESPH (SEQ ID NO: 4758), NYSPH (SEQ ID NO: 4759), VGSPH (SEQ ID NO: 4760), SCSPH (SEQ ID NO: 4761), LLSPH (SEQ ID NO: 4762), NGSPH (SEQ ID NO: 4763), PGSPH (SEQ ID NO: 4764), GGSPH (SEQ ID NO: 4765), TGSPH (SEQ ID NO: 4766), SVSPH (SEQ ID NO: 4767), IGSPH (SEQ ID NO: 4768), DGSPH (SEQ ID NO: 4769), LGSPH (SEQ ID NO: 4770), AGSPH (SEQ ID NO: 4771), EGSPH (SEQ ID NO: 4772), SASPH (SEQ ID NO: 4773), YDSPH (SEQ ID NO: 4774), HGSPH (SEQ ID NO: 4775), RDSPH (SEQ ID NO: 4776), NDSPH (SEQ ID NO: 4777), PDSPH (SEQ ID NO: 4778), MGSPH (SEQ ID NO: 4779), QVSPH (SEQ ID NO: 4780), HNSPH (SEQ ID NO: 4781), HPSPH (SEQ ID NO: 4782), or HSSPH (SEQ ID NO: 4783); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, or 4 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-

[N2] is or comprises GSGSPH (SEQ ID NO: 4695), GHDSPH (SEQ ID NO: 4784), GQDSPH (SEQ ID NO: 4785), VSGSPH (SEQ ID NO: 4786), CSGSPH (SEQ ID NO: 4787), GRGSPH (SEQ ID NO: 4788), CSHSPH (SEQ ID NO: 4789), GQSSPH (SEQ ID NO: 4790), GSHSPH (SEQ ID NO: 4791), GDDSPH (SEQ ID NO: 4792), GHESPH (SEQ ID NO: 4793), GNYSPH (SEQ ID NO: 4794), RVGSPH (SEQ ID NO: 4795), GSCSPH (SEQ ID NO: 4796), GLLSPH (SEQ ID NO: 4797), MSGSPH (SEQ ID NO: 4798), RNGSPH (SEQ ID NO: 4799), TSGSPH (SEQ ID NO: 4800), ISGSPH (SEQ ID NO: 4801), GPGSPH (SEQ ID NO: 4802), ESGSPH (SEQ ID NO: 4803), SSGSPH (SEQ ID NO: 4804), GNGSPH (SEQ ID NO: 4805), ASGSPH (SEQ ID NO: 4806), NSGSPH (SEQ ID NO: 4807), LSGSPH (SEQ ID NO: 4808), GGGSPH (SEQ ID NO: 4809), KSGSPH (SEQ ID NO: 4810), HSGSPH (SEQ ID NO: 4811), GTGSPH (SEQ ID NO: 4812), PSGSPH (SEQ ID NO: 4813), GSVSPH (SEQ ID NO: 4814), RSGSPH (SEQ ID NO: 4815), GIGSPH (SEQ ID NO: 4816), WSGSPH (SEQ ID NO: 4817), DSGSPH (SEQ ID NO: 4818), IDGSPH (SEQ ID NO: 4819), GLGSPH (SEQ ID NO: 4820), DAGSPH (SEQ ID NO: 4821), DGGSPH (SEQ ID NO: 4822), MEGSPH (SEQ ID NO: 4823), ENGSPH (SEQ ID NO: 4824), GSASPH (SEQ ID NO: 4825), KNGSPH (SEQ ID NO: 4826), KEGSPH (SEQ ID NO: 4827), AIGSPH (SEQ ID NO: 4828), GYDSPH (SEQ ID NO: 4829), GHGSPH (SEQ ID NO: 4830), GRDSPH (SEQ ID NO: 4831), GNDSPH (SEQ ID NO: 4832), GPDSPH (SEQ ID NO: 4833), GMGSPH (SEQ ID NO: 4834), GQVSPH (SEQ ID NO: 4835), GHNSPH (SEQ ID NO: 4836), GHPSPH (SEQ ID NO: 4837), or GHSSPH (SEQ ID NO: 4838); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, or 5 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2] is or comprises GSGSPH (SEQ ID NO: 4695). In some embodiments, [N1]-[N2] is or comprises GHDSPH (SEQ ID NO: 4784).

[0155] In some embodiments, X4, X5, or both of [N3] are K. In some embodiments, X4, X5, or X6 of [N3] is R. In some embodiments, position X4 of [N3] is independently chosen from: A, K, V, S, T, G, F, W, V, N, or R. In some embodiments, position X5 of [N3] is independently chosen from: S, K, T, F, I, L, Y, H, M, or R. In some embodiments, position X6 of [N3] is independently chosen from: G, R, A, M, I, N, T, Y, D, P, V, L, E, W, N, Q, K, or S. In some embodiments, [N3] comprises SK, KA, KS, AR, RM, VK, AS, SR, VK, KR, KK, KN, VR, RS, RK, KT, TS, KF, FG, KI, IG, KL, LG, TT, TY, KY, YG, KD, KP, TR, RG, VR, GA, SL, SS, FL, WK, SA, RA, LR, KW, RR, GK, TK, NK, AK, KV, KG, KH, KM, TG, SE, SV, SW, SN, HG, SQ, LW, MG, MA, or SG. In some embodiments, [N3] comprises SK, KA, KS, or SG. In some embodiments, [N3] is or comprises SKA, KSG, ARM, VKS, ASR, VKI, KKN, VRM, RKA, KTS, KFG, KIG, KLG, KTT, KTY, KYG, SKD, SKP, TRG, VRG, KRG, GAR, KSA, KSR, SKL, SRA, SKR, SLR, SRG, SSR, FLR, SKW, SKS,

WKA, VRR, SKV, SKT, SKG, GKA, TKA, NKA, SKL, SKN, AKA, KTG, KSL, KSE, KSV, KSW, KSN, KHG, KSQ, KSK, KLW, WKG, KMG, KMA, or RSG. In some embodiments, [N3] is or comprises SKA. In some embodiments, [N3] is or comprises KSG. In some embodiments, [N2]-[N3] comprises SPHSK (SEQ ID NO: 4701), SPHKS (SEQ ID NO: 4704), SPHAR (SEQ ID NO: 4705), SPHVK (SEQ ID NO: 4706), SPHAS (SEQ ID NO: 4707), SPHKK (SEQ ID NO: 4708), SPHVR (SEQ ID NO: 4709), SPHRK (SEQ ID NO: 4710), SPHKT (SEQ ID NO: 4711), SPHKF (SEQ ID NO: 4712), SPHKI (SEQ ID NO: 4713), SPHKL (SEQ ID NO: 4714), SPHKY (SEQ ID NO: 4715), SPHTR (SEQ ID NO: 4716), SPHKR (SEQ ID NO: 4717), SPHGA (SEQ ID NO: 4718), SPHSR (SEQ ID NO: 4719), SPHSL (SEQ ID NO: 4720), SPHSS (SEQ ID NO: 4721), SPHFL (SEQ ID NO: 4722), SPHWK (SEQ ID NO: 4723), SPHGK (SEQ ID NO: 4724), SPHTK (SEQ ID NO: 4725), SPHNK (SEQ ID NO: 4726), SPHAK (SEQ ID NO: 4727), SPHKH (SEQ ID NO: 4728), SPHKM (SEQ ID NO: 4729), or SPHRS (SEQ ID NO: 4730). In some embodiments [N2]-[N3] comprises SPHSK (SEQ ID NO: 4701) or SPHKS (SEQ ID NO: 4704). In some embodiments, [N2]-[N3] is or comprises SPHSKA (SEQ ID NO: 941), SPHKSG (SEQ ID NO: 946), SPHARM (SEQ ID NO: 947), SPHVKS (SEQ ID NO: 948), SPHASR (SEQ ID NO: 949), SPHVKI (SEQ ID NO: 950), SPHKKN (SEQ ID NO: 954), SPHVRM (SEQ ID NO: 955), SPHRKA (SEQ ID NO: 956), SPHKFG (SEQ ID NO: 957), SPHKIG (SEQ ID NO: 958), SPHKLG (SEQ ID NO: 959), SPHKTS (SEQ ID NO: 963), SPHKTT (SEQ ID NO: 964), SPHKTY (SEQ ID NO: 965), SPHKYG (SEQ ID NO: 966), SPHSKD (SEQ ID NO: 967), SPHSKP (SEQ ID NO: 968), SPHTRG (SEQ ID NO: 972), SPHVRG (SEQ ID NO: 973), SPHKRG (SEQ ID NO: 974), SPHGAR (SEQ ID NO: 975), SPHKSA (SEQ ID NO: 977), SPHKSR (SEQ ID NO: 951), SPHSKL (SEQ ID NO: 960), SPHSRA (SEQ ID NO: 969), SPHSKR (SEQ ID NO: 978), SPHSLR (SEQ ID NO: 952), SPHSRG (SEQ ID NO: 961), SPHSSR (SEQ ID NO: 970), SPHFLR (SEQ ID NO: 979), SPHSKW (SEQ ID NO: 953), SPHSKS (SEQ ID NO: 962), SPHWKA (SEQ ID NO: 971), SPHVRR (SEQ ID NO: 980), SPHSKT (SEQ ID NO: 4731), SPHSKG (SEQ ID NO: 4732), SPHGKA (SEQ ID NO: 4733), SPHNKA (SEQ ID NO: 4734), SPHSKN (SEQ ID NO: 4735), SPHAKA (SEQ ID NO: 4736), SPHSKV (SEQ ID NO: 4737), SPHKTG (SEQ ID NO: 4738), SPHTKA (SEQ ID NO: 4739), SPHKSL (SEQ ID NO: 4740), SPHKSE (SEQ ID NO: 4741), SPHKSV (SEQ ID NO: 4742), SPHKSW (SEQ ID NO: 4743), SPHKSN (SEQ ID NO: 4744), SPHKHG (SEQ ID NO: 4745), SPHKSQ (SEQ ID NO: 4746), SPHKSK (SEQ ID NO: 4747), SPHKLW (SEQ ID NO: 4748), SPHWKG (SEQ ID NO: 4749), SPHKMG (SEQ ID NO: 4750), SPHKMA (SEQ ID NO: 4751), or SPHRSG (SEQ ID NO: 976). In some embodiments, [N2]-[N3] is SPHSKA (SEQ ID NO: 941). In some embodiments, [N2]-[N3] is or comprises SPHKSG (SEQ ID NO: 946).

[0156] In some embodiments, [N1]-[N2]-[N3] comprises SGSPHSK (SEQ ID NO: 4839), HDSPHKS (SEQ ID NO: 4840), SGSPHAR (SEQ ID NO: 4841), SGSPHVK (SEQ ID NO: 4842), QDSPHKS (SEQ ID NO: 4843), SGSPHKK (SEQ ID NO: 4844), SGSPHVR (SEQ ID NO: 4845), SGSPHAS (SEQ ID NO: 4846), SGSPHRK (SEQ ID NO: 4847), SGSPHKT (SEQ ID NO: 4848),

SHSPHKS (SEQ ID NO: 4849), QSSPHRS (SEQ ID NO: 4850), RGSPHAS (SEQ ID NO: 4851), RGSPHSK (SEQ ID NO: 4852), SGSPHKF (SEQ ID NO: 4853), SGSPHKI (SEQ ID NO: 4854), SGSPHKL (SEQ ID NO: 4855), SGSPHKY (SEQ ID NO: 4856), SGSPHTR (SEQ ID NO: 4857), SHSPHKR (SEQ ID NO: 4858), SGSPHGA (SEQ ID NO: 4859), HDSPHKR (SEQ ID NO: 4860), DDSPHKS (SEQ ID NO: 4861), HESPHKS (SEQ ID NO: 4862), NYSPHKI (SEQ ID NO: 4863), SGSPHSR (SEQ ID NO: 4864), SGSPHSL (SEQ ID NO: 4865), SGSPHSS (SEQ ID NO: 4866), VGSPHSK (SEQ ID NO: 4867), SCSPHRK (SEQ ID NO: 4868), SGSPHFL (SEQ ID NO: 4869), LLSPHWK (SEQ ID NO: 4870), NGSPHSK (SEQ ID NO: 4871), PGSPHSK (SEQ ID NO: 4872), GGSPHSK (SEQ ID NO: 4873), TGSPHSK (SEQ ID NO: 4874), SVSPHGK (SEQ ID NO: 4875), SGSPHTK (SEQ ID NO: 4876), IGSPHSK (SEQ ID NO: 4877), DGSPHSK (SEQ ID NO: 4878), SGSPHNK (SEQ ID NO: 4879), LGSPHSK (SEQ ID NO: 4880), AGSPHSK (SEQ ID NO: 4881), EGSPHSK (SEQ ID NO: 4882), SASPHSK (SEQ ID NO: 4883), SGSPHAK (SEQ ID NO: 4884), HDSPHKI (SEQ ID NO: 4885), YDSPHKS (SEQ ID NO: 4886), HDSPHKT (SEQ ID NO: 4887), RGSPHKR (SEQ ID NO: 4888), HGSPHSK (SEQ ID NO: 4889), RDSPHKS (SEQ ID NO: 4890), NDSPHKS (SEQ ID NO: 4891), QDSPHKI (SEQ ID NO: 4892), PDSPHKI (SEQ ID NO: 4893), PDSPHKS (SEQ ID NO: 4894), MGSPHSK (SEQ ID NO: 4895), HDSPHKH (SEQ ID NO: 4896), QVSPHKS (SEQ ID NO: 4897), HNSPHKS (SEQ ID NO: 4898), NGSPHKR (SEQ ID NO: 4899), HDSPHKY (SEQ ID NO: 4900), NDSPHKI (SEQ ID NO: 4901), HDSPHKL (SEQ ID NO: 4902), HPSPHWK (SEQ ID NO: 4903), HDSPHKM (SEQ ID NO: 4904), or HSSPHRS (SEQ ID NO: 4905). In some embodiments, [N1]-[N2]-[N3] is GSGSPHSKA (SEQ ID NO: 4697), GHDSPHKSG (SEQ ID NO: 4698), GSGSPHARM (SEQ ID NO: 4906), GSGSPHVKS (SEQ ID NO: 4907), GQDSPHKSG (SEQ ID NO: 4908), GSGSPHASR (SEQ ID NO: 4909), GSGSPHVKI (SEQ ID NO: 4910), GSGSPHKKN (SEQ ID NO: 4911), GSGSPHVRM (SEQ ID NO: 4912), VSGSPHSKA (SEQ ID NO: 4913), CSGSPHSKA (SEQ ID NO: 4914), GSGSPHRKA (SEQ ID NO: 4915), CSGSPHKTS (SEQ ID NO: 4916), CSHSPHKSG (SEQ ID NO: 4917), GQSSPHRSG (SEQ ID NO: 4918), GRGSPHASR (SEQ ID NO: 4919), GRGSPHSKA (SEQ ID NO: 4920), GSGSPHKFG (SEQ ID NO: 4921), GSGSPHKIG (SEQ ID NO: 4922), GSGSPHKLG (SEQ ID NO: 4923), GSGSPHKTS (SEQ ID NO: 4924), GSGSPHKTT (SEQ ID NO: 4925), GSGSPHKTY (SEQ ID NO: 4926), GSGSPHKYG (SEQ ID NO: 4927), GSGSPHSKD (SEQ ID NO: 4928), GSGSPHSKP (SEQ ID NO: 4929), GSGSPHTRG (SEQ ID NO: 4930), GSGSPHVRG (SEQ ID NO: 4931), GSHSPHKRG (SEQ ID NO: 4932), GSHSPHKSG (SEQ ID NO: 4933), VSGSPHASR (SEQ ID NO: 4934), VSGSPHGAR (SEQ ID NO: 4935), VSGSPHKFG (SEQ ID NO: 4936), GHDSPHKRG (SEQ ID NO: 4937), GDDSPHKSG (SEQ ID NO: 4938), GHESPHKSA (SEQ ID NO: 4939), GHDSPHKSA (SEQ ID NO: 4940), GNYSPHKIG (SEQ ID NO: 4941), GHDSPHKSR (SEQ ID NO: 4942), GSGSPHSKL (SEQ ID NO: 4943), GSGSPHSRA (SEQ ID NO: 4944), GSGSPHSKR (SEQ ID NO: 4945), GSGSPHSLR (SEQ ID NO: 4946), GSGSPHSRG (SEQ ID NO: 4947), GSGSPHSSR (SEQ ID NO: 4948), RVGSPHSKA (SEQ ID NO: 4949), GSCSPHRKA (SEQ ID NO: 4950),

GSGSPHFLR (SEQ ID NO: 4951), GSGSPHSKW (SEQ ID NO: 4952), GSGSPHSKS (SEQ ID NO: 4953), GLLSPHWKA (SEQ ID NO: 4954), GSGSPHVRR (SEQ ID NO: 4955), GSGSPHSKV (SEQ ID NO: 4956), MSGSPHSKA (SEQ ID NO: 4957), RNGSPHSKA (SEQ ID NO: 4958), TSGSPHSKA (SEQ ID NO: 4959), ISGSPHSKA (SEQ ID NO: 4960), GPGSPHSKA (SEQ ID NO: 4961), GSGSPHSKT (SEQ ID NO: 4962), ESGSPHSKA (SEQ ID NO: 4963), SSGSPHSKA (SEQ ID NO: 4964), GNGSPHSKA (SEQ ID NO: 4965), ASGSPHSKA (SEQ ID NO: 4966), NSGSPHSKA (SEQ ID NO: 4967), LSGSPHSKA (SEQ ID NO: 4968), GGGSPHSKA (SEQ ID NO: 4969), KSGSPHSKA (SEQ ID NO: 4970), GGGSPHSKS (SEQ ID NO: 4971), GSGSPHSKG (SEQ ID NO: 4972), HSGSPHSKA (SEQ ID NO: 4973), GTGSPHSKA (SEQ ID NO: 4974), PSGSPHSKA (SEQ ID NO: 4975), GSVSPHGKA (SEQ ID NO: 4976), RSGSPHSKA (SEQ ID NO: 4977), GSGSPHTKA (SEQ ID NO: 4978), GIGSPHSKA (SEQ ID NO: 4979), WSGSPHSKA (SEQ ID NO: 4980), DSGSPHSKA (SEQ ID NO: 4981), IDGSPHSKA (SEQ ID NO: 4982), GSGSPHNKA (SEQ ID NO: 4983), GLGSPHSKS (SEQ ID NO: 4984), DAGSPHSKA (SEQ ID NO: 4985), DGGSPHSKA (SEQ ID NO: 4986), MEGSPHSKA (SEQ ID NO: 4987), ENGSPHSKA (SEQ ID NO: 4988), GSASPHSKA (SEQ ID NO: 4989), GNGSPHSKS (SEQ ID NO: 4990), KNGSPHSKA (SEQ ID NO: 4991), KEGSPHSKA (SEQ ID NO: 4992), AIGSPHSKA (SEQ ID NO: 4993), GSGSPHSKN (SEQ ID NO: 4994), GSGSPHAKA (SEQ ID NO: 4995), GHDSPHKIG (SEQ ID NO: 4996), GYDSPHKSG (SEQ ID NO: 4997), GHESPHKSG (SEQ ID NO: 4998), GHDSPHKTG (SEQ ID NO: 4999), GRGSPHKRG (SEQ ID NO: 5000), GQDSPHKSG (SEQ ID NO: 4908), GHDSPHKSL (SEQ ID NO: 5001), GHGSPHSKA (SEQ ID NO: 5002), GHDSPHKSE (SEQ ID NO: 5003), VSGSPHSKA (SEQ ID NO: 4913), GRDSPHKSG (SEQ ID NO: 5004), GNDSPHKSV (SEQ ID NO: 5005), GQDSPHKIG (SEQ ID NO: 5006), GHDSPHKSV (SEQ ID NO: 5007), GPDSPHKIG (SEQ ID NO: 5008), GPDSPHKSG (SEQ ID NO: 5009), GHDSPHKSW (SEQ ID NO: 5010), GHDSPHKSN (SEQ ID NO: 5011), GMGSPHSKT (SEQ ID NO: 5012), GHDSPHKHG (SEQ ID NO: 5013), GQVSPHKSG (SEQ ID NO: 5014), GDDSPHKSV (SEQ ID NO: 5015), GHNSPHKSG (SEQ ID NO: 5016), GNGSPHKRG (SEQ ID NO: 5017), GHDSPHKYG (SEQ ID NO: 5018), GHDSPHKSQ (SEQ ID NO: 5019), GNDSPHKIG (SEQ ID NO: 5020), GHDSPHKSK (SEQ ID NO: 5021), GHDSPHKLW (SEQ ID NO: 5022), GHPSPHWKG (SEQ ID NO: 5023), GHDSPHKMG (SEQ ID NO: 5024), GHDSPHKMA (SEQ ID NO: 5025), or GHSSPHRSG (SEQ ID NO: 5026); an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, or 8 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2]-[N3] is or comprises GSGSPHSKA (SEQ ID NO: 4697). In some embodiments, [N1]-[N2]-[N3] is or comprises GHDSPHKSG (SEQ ID NO: 4698).

[0157] In some embodiments, the AAV capsid variant comprising an amino acid sequence having the formula [N1]-[N2]-[N3], further comprises [N4], wherein [N4] comprises X7 X8 X9 X10. In some embodiments, position X7 of [N4] is independently chosen from W, Q, K, R, G, L, V, S, P, H, K, I, M, A, E, or F. In some embodiments, position X8 of [N4] is independently chosen from N, Y, C, K, T, H, R, D, V, S, P, G, W, E, F, A, I, M, Q, or L. In some embodiments, position X9 of [N4] is independently chosen from Q, G, K, H, R, T, L, D, A, P, I, F, V, M, W, Y, S, E, N, or Y. In some embodiments, position X10 of [N4] is independently chosen from Q, H, L, R, W, K, A, P, E, M, I, S, G, N, Y, C, V, T, D, or V. In some embodiments [N4] comprises QNQQ (SEQ ID NO: 5028), WNQQ (SEQ ID NO: 5029), QYYV (SEQ ID NO: 5030), RRQQ (SEQ ID NO: 5031), GCGQ (SEQ ID NO: 5032), LRQQ (SEQ ID NO: 5033), RNQQ (SEQ ID NO: 5034), VNQQ (SEQ ID NO: 5035), FRLQ (SEQ ID NO: 5036), FNQQ (SEQ ID NO: 5037), LLQQ (SEQ ID NO: 5038), SNQQ (SEQ ID NO: 5039), RLQQ (SEQ ID NO: 5040), LNQQ (SEQ ID NO: 5041), QRKL (SEQ ID NO: 5042), LRRQ (SEQ ID NO: 5043), QRLR (SEQ ID NO: 5044), QRRL (SEQ ID NO: 5045), RRLQ (SEQ ID NO: 5046), RLRQ (SEQ ID NO: 5047), SKRQ (SEQ ID NO: 5048), QLYR (SEQ ID NO: 5049), QLTV (SEQ ID NO: 5050), QNKQ (SEQ ID NO: 5051), KNQQ (SEQ ID NO: 5052), QKQQ (SEQ ID NO: 5053), QTQQ (SEQ ID NO: 5054), QNHQ (SEQ ID NO: 5055), QHQQ (SEQ ID NO: 5056), QNQH (SEQ ID NO: 5057), QHRQ (SEQ ID NO: 5058), LTQQ (SEQ ID NO: 5059), QNQW (SEQ ID NO: 5060), QNTH (SEQ ID NO: 5061), RRRQ (SEQ ID NO: 5062), QYQQ (SEQ ID NO: 5063), QNDQ (SEQ ID NO: 5064), QNRH (SEQ ID NO: 5065), RDQQ (SEQ ID NO: 5066), PNLQ (SEQ ID NO: 5067), HVRQ (SEQ ID NO: 5068), PNQH (SEQ ID NO: 5069), HNQQ (SEQ ID NO: 5070), QSQQ (SEQ ID NO: 5071), QPAK (SEQ ID NO: 5072), QNLA (SEQ ID NO: 5073), QNQL (SEQ ID NO: 5074), QGQQ (SEQ ID NO: 5075), LNRQ (SEQ ID NO: 5076), QNPP (SEQ ID NO: 5077), QNLQ (SEQ ID NO: 5078), QDQE (SEQ ID NO: 5079), QDQQ (SEQ ID NO: 5080), HWQQ (SEQ ID NO: 5081), PNQQ (SEQ ID NO: 5082), PEQQ (SEQ ID NO: 5083), QRTM (SEQ ID NO: 5084), LHQH (SEQ ID NO: 5085), QHRI (SEQ ID NO: 5086), QYIH (SEQ ID NO: 5087), QKFE (SEQ ID NO: 5088), QFPS (SEQ ID NO: 5089), QNPL (SEQ ID NO: 5090), QAIK (SEQ ID NO: 5091), QNRQ (SEQ ID NO: 5092), QYQH (SEQ ID NO: 5093), QNPQ (SEQ ID NO: 5094), QHQL (SEQ ID NO: 5095), QSPP (SEQ ID NO: 5096), QAKL (SEQ ID NO: 5097), KSQQ (SEQ ID NO: 5098), QDRP (SEQ ID NO: 5099), QNLG (SEQ ID NO: 5100), QAFH (SEQ ID NO: 5101), QNAQ (SEQ ID NO: 5102), HNQL (SEQ ID NO: 5103), QKLN (SEQ ID NO: 5104), QNVQ (SEQ ID NO: 5105), QAQQ (SEQ ID NO: 5106), QTPP (SEQ ID NO: 5107), QPPA (SEQ ID NO: 5108), QERP (SEQ ID NO: 5109), QDLQ (SEQ ID NO: 5110), QAMH (SEQ ID NO: 5111), QHPS (SEQ ID NO: 5112), PGLQ (SEQ ID NO: 5113), QGIR (SEQ ID NO: 5114), QAPA (SEQ ID NO: 5115), QIPP (SEQ ID NO: 5116), QTQL (SEQ ID NO: 5117), QAPS (SEQ ID NO: 5118), QNTY (SEQ ID NO: 5119), QDKQ (SEQ ID NO: 5120), QNHL (SEQ ID NO: 5121), QIGM (SEQ ID NO: 5122), LNKQ (SEQ ID NO: 5123), PNQL (SEQ ID NO: 5124), QLQQ (SEQ ID NO: 5125), QRMS (SEQ ID NO: 5126), QGIL (SEQ ID NO: 5127), QDRQ (SEQ ID NO: 5128), RDWQ (SEQ ID NO: 5129), QERS (SEQ

ID NO: 5130), QNYQ (SEQ ID NO: 5131), QRTC (SEQ ID NO: 5132), QIGH (SEQ ID NO: 5133), QGAI (SEQ ID NO: 5134), QVPP (SEQ ID NO: 5135), QVQQ (SEQ ID NO: 5136), LMRQ (SEQ ID NO: 5137), QYSV (SEQ ID NO: 5138), QAIT (SEQ ID NO: 5139), QKTL (SEQ ID NO: 5140), QLHH (SEQ ID NO: 5141), QNII (SEQ ID NO: 5142), QGHH (SEQ ID NO: 5143), QSKV (SEQ ID NO: 5144), QLPS (SEQ ID NO: 5145), IGKQ (SEQ ID NO: 5146), QAIH (SEQ ID NO: 5147), QHGL (SEQ ID NO: 5148), QFMC (SEQ ID NO: 5149), QNQM (SEQ ID NO: 5150), QHLQ (SEQ ID NO: 5151), QPAR (SEQ ID NO: 5152), QSLQ (SEQ ID NO: 5153), QSQL (SEQ ID NO: 5154), HSQQ (SEQ ID NO: 5155), QMPS (SEQ ID NO: 5156), QGSL (SEQ ID NO: 5157), QVPA (SEQ ID NO: 5158), HYQQ (SEQ ID NO: 5159), QVPS (SEQ ID NO: 5160), RGEQ (SEQ ID NO: 5161), PGQQ (SEQ ID NO: 5162), LEQQ (SEQ ID NO: 5163), QNQS (SEQ ID NO: 5164), QKVI (SEQ ID NO: 5165), QNND (SEQ ID NO: 5166), QSVH (SEQ ID NO: 5167), QPLG (SEQ ID NO: 5168), HNQE (SEQ ID NO: 5169), QIQQ (SEQ ID NO: 5170), QVRN (SEQ ID NO: 5171), PSNQ (SEQ ID NO: 5172), QVGH (SEQ ID NO: 5173), QRDI (SEQ ID NO: 5174), QMPN (SEQ ID NO: 5175), RGLQ (SEQ ID NO: 5176), PSLQ (SEQ ID NO: 5177), QRDQ (SEQ ID NO: 5178), QAKG (SEQ ID NO: 5179), QSAH (SEQ ID NO: 5180), QSTM (SEQ ID NO: 5181), QREM (SEQ ID NO: 5182), QYRA (SEQ ID NO: 5183), QRQQ (SEQ ID NO: 5184), QWQQ (SEQ ID NO: 5185), QRMN (SEQ ID NO: 5186), GDSQ (SEQ ID NO: 5187), QKIS (SEQ ID NO: 5188), PSMQ (SEQ ID NO: 5189), SPRQ (SEQ ID NO: 5190), MEQQ (SEQ ID NO: 5191), QYQN (SEQ ID NO: 5192), QIRQ (SEQ ID NO: 5193), QSVQ (SEQ ID NO: 5194), RSQQ (SEQ ID NO: 5195), QNKL (SEQ ID NO: 5196), QIQH (SEQ ID NO: 5197), PRQQ (SEQ ID NO: 5198), HTQQ (SEQ ID NO: 5199), QRQH (SEQ ID NO: 5200), RNQE (SEQ ID NO: 5201), QSKQ (SEQ ID NO: 5202), QNQP (SEQ ID NO: 5203), QSPQ (SEQ ID NO: 5204), QTRQ (SEQ ID NO: 5205), QNLH (SEQ ID NO: 5206), QNQE (SEQ ID NO: 5207), LNQP (SEQ ID NO: 5208), QNQD (SEQ ID NO: 5209), QNLL (SEQ ID NO: 5210), QLVI (SEQ ID NO: 5211), RTQE (SEQ ID NO: 5212), QTHQ (SEQ ID NO: 5213), QDQH (SEQ ID NO: 5214), QSQH (SEQ ID NO: 5215), VRQQ (SEQ ID NO: 5216), AWQQ (SEQ ID NO: 5217), QSVP (SEQ ID NO: 5218), QNIQ (SEQ ID NO: 5219), LDQQ (SEQ ID NO: 5220), PDQQ (SEQ ID NO: 5221), ESQQ (SEQ ID NO: 5222), QRQL (SEQ ID NO: 5223), QIIV (SEQ ID NO: 5224), QKQS (SEQ ID NO: 5225), QSHQ (SEQ ID NO: 5226), QFVV (SEQ ID NO: 5227), QSQP (SEQ ID NO: 5228), QNEQ (SEQ ID NO: 5229), INQQ (SEQ ID NO: 5230), RNRQ (SEQ ID NO: 5231), RDQK (SEQ ID NO: 5232), QWKR (SEQ ID NO: 5233), ENRQ (SEQ ID NO: 5234), QTQP (SEQ ID NO: 5235), QKQL (SEQ ID NO: 5236), RNQL (SEQ ID NO: 5237), ISIQ (SEQ ID NO: 5238), QTVC (SEQ ID NO: 5239), QQIM (SEQ ID NO: 5240), LNHQ (SEQ ID NO: 5241), QNQA (SEQ ID NO: 5242), QMIH (SEQ ID NO: 5243), RNHQ (SEQ ID NO: 5244), or QKMN (SEQ ID NO: 5245), or any dipeptide or tripeptide thereof. In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises: the amino acid sequence of any of SEQ ID NOs: 1800-2241; an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one,

two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises GSGSPHSKAQNQQ (SEQ ID NO: 1801). In some embodiments, [N1]-[N2]-[N3]-[N4] is or comprises GHDSPHKSGQNQQ (SEQ ID NO: 1800).

[0158] In some embodiments, the AAV capsid variant comprising an amino acid sequence having the formula [N1]-[N2]-[N3], further comprises [N0], wherein [N0] comprises XA XB and XC. In some embodiments, XA of [N0] is independently chosen from T, S, Y, M, A, C, I, R, L, D, F, V, Q, N, H, E, or G. In some embodiments, XB of [N0] is independently chosen from I, M, P, E, N, D, S, A, T, G, Q, F, V, L, C, H, R, W, or L. In some embodiments, XC of [N0] is independently chosen from N, M, E, G, Y, W, T, I, Q, F, V, A, L, I, P, K, R, H, S, D, or S. In some embodiments, [N0] comprises TIN, SMN, TIM, YLS, GLS, MPE, MEG, MEY, AEW, CEW, ANN, IPE, ADM, IEY, ADY, IET, MEW, CEY, RIN, MEI, LEY, ADW, IEI, DIM, FEQ, MEF, CDQ, LPE, IEN, MES, AEI, VEY, IIN, TSN, IEV, MEM, AEV, MDA, VEW, AEQ, LEW, MEL, MET, MEA, IES, MEV, CEI, ATN, MDG, QEV, ADQ, NMN, IEM, ISN, TGN, QQQ, HDW, IEG, TII, TFP, TEK, EIN, TVN, TFN, SIN, TER, TSY, ELH, AIN, SVN, TDN, TFH, TVH, TEN, TSS, TID, TCN, NIN, TEH, AEM, AIK, TDK, TFK, SDQ, TEI, NTN, TET, SIK, TEL, TEA, TAN, TIY, TFS, TES, TTN, TED, TNN, EVH, TIS, TVR, TDR, TIK, NHI, TIP, ESD, TDL, TVP, TVI, AEH, NCL, TVK, NAD, TIT, NCV, TIR, NAL, VIN, TIQ, TEF, TRE, QGE, SEK, NVN, GGE, EFV, SDK, TEQ, EVQ, TEY, NCW, TDV, SDI, NSI, NSL, EVV, TEP, SEL, TWQ, TEV, AVN, GVL, TLN, TEG, TRD, NAI, AEN, AET, ETA, NNL, or any dipeptide thereof. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises the amino acid sequence of any one of SEQ ID NOs: 2242-2886; an amino acid sequence comprising any portion of any of the aforesaid amino acid sequences (e.g., any 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids, e.g., consecutive amino acids) thereof; an amino acid sequence comprising one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the aforesaid amino acid sequences; or an amino acid sequence comprising one, two, or three but no more than four different amino acids, relative to any one of the aforesaid amino acid sequences. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGSGSPHSKAQNQQ (SEQ ID NO: 2242). In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is or comprises TINGHDSPHKSGQNQQ (SEQ ID NO: 2243).

[0159] In some embodiments, [N1]-[N2]-[N3] is present in loop IV of the AAV capsid variant. In some embodiments [N0] and [N4] are present in loop IV of the AAV capsid variant. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is present in loop IV of the AAV capsid variant.

In some embodiments, [N0] is present immediately subsequent to position 449, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, wherein [N0] is present immediately subsequent to position 449, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 981 or 982. In some embodiments, [N0] replaces positions 450, 451, and 452 (e.g., amino acids T450, I451, and N452), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. wherein [N0] is present immediately subsequent to position 449 and wherein [N0] replaces positions 450-452 (e.g., T450, I451, and N452), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. In some embodiments, [N1] is present immediately subsequent to position 452, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138, 981 or 982. In some embodiments, wherein [N1] replaces positions 453- 455 (e.g., G453, S454, and G455), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. In some embodiments, [N1] is present immediately subsequent to position 452 and wherein [N1] replaces positions 453-455 (e.g., G453, S454, and G455), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. In some embodiments, [N2] is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138, 981, or 982. In some embodiments, [N2]-[N3] is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138, 981, or 982. In some embodiments [N1]-[N2]-[N3] is present immediately subsequent to position 452, numbered relative to SEQ ID NO: 138, 981, or 982. In some embodiments, [N1]-[N2]-[N3] replaces positions 453-455 (e.g., G453, S454, and G455), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. In some embodiments, [N1] is present immediately subsequent to position 452 and wherein [N1]-[N2]-[N3] replaces positions 453-455 (e.g., G453, S454, and G455), relative to a reference sequence numbered according to SEQ ID NO: 138, 981, or 982. In some embodiments, [N4] is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N4] replaces positions 456-459 (e.g., Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N4] is present immediately subsequent to position 455, and [N4] replaces positions 456-459 (e.g., Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N2]-[N3]-[N4] replaces positions 456-459 (e.g., Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N2]-[N3]-[N4] is present immediately subsequent to position 455, and wherein [N2]-[N3]-[N4] replaces positions 456-459 (e.g., Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N1]-[N2]-[N3]-[N4] replaces positions 453-459 (e.g., G453, S454, G455, Q456, N457, Q458, and Q459), relative to a reference

sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N1]-[N2]-[N3]-[N4] is present immediately subsequent to position 452, and wherein [N1]-[N2]-[N3]-[N4] replaces positions 453-459 (e.g., G453, S454, G455, Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] replaces positions 450-456 (e.g., T450, I451, N452, G453, S454, G455, Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, [N0]-[N1]-[N2]-[N3]-[N4] is present immediately subsequent to position 449, and wherein [N0]-[N1]-[N2]-[N3]-[N4] replaces positions 450-456 (e.g., T450, I451, N452, G453, S454, G455, Q456, N457, Q458, and Q459), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138.

[0160] In some embodiments, [N3] is present immediately subsequent to [N2].

[0161] In some embodiments, the AAV capsid variant comprises from N-terminus to C-terminus, [N2]-[N3]. In some embodiments, the AAV capsid variant comprises from N-terminus to C-terminus, [N1]-[N2]-[N3]. In some embodiments, the AAV capsid variant comprises from N-terminus to C-terminus, [N1]-[N2]-[N3]-[N4]. In some embodiments, the AAV capsid variant comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]. In some embodiments, the AAV capsid variant comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]-[N4].

[0162] In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence comprising at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, or 17 consecutive amino acids from any one of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least 3, 4, or 5 consecutive amino acids from any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 consecutive amino acids from any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the amino acid sequence is present in loop IV. In some embodiments, the amino acid sequence is present immediately subsequent to position 448, 452, 453, 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138, 981, or 982. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, numbered according to SEQ ID NO: 982. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, numbered according to SEQ ID NO: 981. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of positions 499 (e.g., K499), 450 (e.g., T450), 451 (e.g., I451), 452 (e.g., N452), 453 (e.g., G453), 454 (e.g., S454),

455 (e.g., G455), 456 (e.g., Q456), 457 (e.g., N457), 458 (e.g., Q458), 459 (e.g., Q459), and 460 (e.g., T460), numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises one or more amino acid substitutions at positions 499 (e.g., K499), 450 (e.g., T450), 451 (e.g., I451), 452 (e.g., N452), 453 (e.g., G453), 454 (e.g., S454), 455 (e.g., G455), 456 (e.g., Q456), 457 (e.g., N457), 458 (e.g., Q458), 459 (e.g., Q459), and/or 460 (e.g., T460), numbered according to SEQ ID NO: 138.

[0163] In some embodiments, the 3 consecutive amino acids comprise SPH. In some embodiments, the 4 consecutive amino acids comprise SPHS (SEQ ID NO: 4700). In some embodiments, the 5 consecutive amino acids comprise SPHSK (SEQ ID NO: 4701). In some embodiments, the 6 consecutive amino acids comprise SPHSKA (SEQ ID NO: 941).

[0164] In some embodiments, 3 consecutive amino acids comprise HDS. In some embodiments, the 4 consecutive amino acids comprise HDSP (SEQ ID NO: 4702). In some embodiments, the 5 consecutive amino acids comprise HDSPH (SEQ ID NO: 4703). In some embodiments, the 6 consecutive amino acids comprise HDSPHK (SEQ ID NO: 2). In some embodiments, the 7 consecutive amino acids comprise HDSPHKS (SEQ ID NO: 4840). In some embodiments, the 8 consecutive amino acids comprise HDSPHKSG (SEQ ID NO: 943).

[0165] In some embodiments, 3 consecutive amino acids comprise HDS. In some embodiments, the 4 consecutive amino acids comprise HDSP (SEQ ID NO: 4702). In some embodiments, the 5 consecutive amino acids comprise HDSPH (SEQ ID NO: 4703). In some embodiments, the 6 consecutive amino acids comprise HDSPHK (SEQ ID NO: 2).

[0166] In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids, relative to the amino acid sequence of any one of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four different amino acids, relative to the amino acid sequence of any one of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the AAV capsid variant comprises an

amino acid sequence comprising at least one, two, or three but no more than four different amino acids, from the amino acid sequence of any one of SEQ ID NOs: 2, 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the amino acid sequence is present in loop IV. In some embodiments, the amino acid sequence is present immediately subsequent to position 448, 452, 453, 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138, 981, or 982. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, numbered according to SEQ ID NO: 982. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, numbered according to SEQ ID NO: 981. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of positions 499 (e.g., K499), 450 (e.g., T450), 451 (e.g., I451), 452 (e.g., N452), 453 (e.g., G453), 454 (e.g., S454), 455 (e.g., G455), 456 (e.g., Q456), 457 (e.g., N457), 458 (e.g., Q458), 459 (e.g., Q459), and 460 (e.g., T460), numbered according to SEQ ID NO: 138.

[0167] In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SPHSKA (SEQ ID NO: 941). In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three, but no more than four different amino acids from the amino acid sequence of SPHSKA (SEQ ID NO: 941).

[0168] In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of HDSPHKSG (SEQ ID NO: 943). In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three, but no more than four different amino acids that relative to the amino acid sequence of HDSPHKSG (SEQ ID NO: 943).

[0169] In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of HDSPHK (SEQ ID NO: 2). In some embodiments, the AAV capsid variant comprises an amino acid sequence comprising at least one, two, or three, but no more than four different amino acids that relative to the amino acid sequence of HDSPHK (SEQ ID NO: 2).

[0170] In some embodiments, the AAV capsid variant, comprises the amino acid sequence of any of the sequences provided in Tables 1, 2A, 2B, 13-19. In some embodiments, the peptide comprises the amino acid sequence of any of SEQ ID NOs: 945-980 or 985-986. In some embodiments, the

AAV capsid variant comprises the amino acid sequence of any of SEQ ID NOs: 200, 201, 941, 943, 204, 208, 404, or 903-909. In some embodiments, the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 941. In some embodiments, the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 943. In some embodiments, the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 2. In some embodiments, the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 3589. In some embodiments, the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 1754. In some embodiments, the amino acid sequence is present in loop IV. In some embodiments, the amino acid sequence is present immediately subsequent to position 448, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 449-460 (e.g., K449, T450, I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 448 and replaces positions 449-460 (e.g., K449, T450, I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 449, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 450-460 (e.g., T450, I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 449, and replaces positions 450-460 (e.g., T450, I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 450, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 451-460 (e.g., I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 450 and replaces positions 451-460 (e.g., I451, N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 451, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 452-460 (e.g., N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 451 and replaces positions 452-460 (e.g., N452, G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 452, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 453-460 (e.g., G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative

to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 452, and replaces positions 453-460 (e.g., G453, S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 454 and 455 (e.g., S454 and G455), numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, and replaces positions 454 and 455 (e.g., S454 and G455), numbered according to SEQ ID NO: 138. In some embodiments, the amino acid sequence replaces positions 454-460 (e.g., S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 453, and replaces positions 454-460 (e.g., S454, G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 454, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 454, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 981. In some embodiments, the amino acid sequence replaces positions 455-460 (e.g., positions G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to positions 454, and replaces positions 455-460 (e.g., positions G455, Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 982. In some embodiments, the amino acid sequence replaces positions 456-460 (e.g., Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138. In some embodiments, the amino acid sequence is present immediately subsequent to position 455, and replaces positions 456-460 (e.g., Q456, N457, Q458, Q459, and T460), numbered relative to SEQ ID NO: 138.

[0171] In some embodiments, the AAV capsid variant (e.g., an AAV capsid variant described herein), comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 942 or 944, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the AAV capsid variant described herein, comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 942 or 944, or a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions),

insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 942 or 944. In some embodiments, the AAV capsid variant comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides relative to the nucleotide sequence of SEQ ID NO: 942 or 944.

[0172] In some embodiments, the nucleotide sequence encoding the AAV capsid variant (e.g., an AAV capsid variant described herein), comprises the nucleotide sequence of SEQ ID NO: 942, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the nucleic acid sequence encoding the AAV capsid variant comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequences of SEQ ID NO: 942. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides, relative to the nucleotide sequence of SEQ ID NO: 942.

[0173] In some embodiments, the nucleotide sequence encoding the AAV capsid variant (e.g., an AAV capsid variant described herein), comprises the nucleotide sequence of SEQ ID NO: 944, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the nucleic acid sequence encoding the AAV capsid variant comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but no more than ten modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequences of SEQ ID NO: 944. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein comprises a nucleotide sequence comprising at least one, two, three, four, five, six, or seven, but no more than ten different nucleotides relative to the nucleotide sequence of SEQ ID NO: 944.

[0174] In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), wherein the amino acid sequence is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), wherein the amino acid sequence is present immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 981.

[0175] In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of HDSPHKSG (SEQ ID NO: 943), wherein the amino acid sequence is present immediately subsequent to position 453, relative to a reference sequence numbered according to the

amino acid sequence of SEQ ID NO: 138. In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of HDSPHKSG (SEQ ID NO: 943), wherein the amino acid sequence is present immediately subsequent to position 453, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 982.

[0176] In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of HDSPHK (SEQ ID NO: 2), wherein the amino acid sequence is present immediately subsequent to position 453, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of HDSPHK (SEQ ID NO: 2), wherein the amino acid sequence is present immediately subsequent to position 453, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 982.

[0177] In some embodiments, an AAV capsid variant described herein comprises (i) the amino acid sequence of HDSPHKSG (SEQ ID NO: 943), which is present immediately subsequent to position 453; and (ii) a deletion of amino acids SG at position 454 and 455; wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0178] In some embodiments, an AAV capsid variant described herein comprises (i) the amino acid sequence of HDSPHSKA (SEQ ID NO: 4486), which is present immediately subsequent to position 453; and (ii) a deletion of amino acids SG at position 454 and 455; wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0179] In some embodiments, an AAV capsid variant described herein comprises an amino acid other than S at position 454 and/or an amino acid other than G at position 455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises the amino acid H at position 454 and the amino acid D at position 455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises the amino acid sequence of SPHKSG (SEQ ID NO: 946). In some embodiments, the AAV capsid variant comprises: (i) the amino acid H at position 454 and the amino acid D at position 455, and (ii) the amino acid sequence SPHKSG (SEQ ID NO: 946), wherein the amino acid sequence of SPHKSG (SEQ ID NO: 946) is present immediately subsequent to position 455, wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0180] In some embodiments, an AAV capsid variant described herein comprises an amino acid other than S at position 454 and/or an amino acid other than G at position 455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises the amino acid H at position 454 and the amino acid D at position 455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941). In some embodiments, the AAV capsid variant comprises: (i) the amino acid H at position 454 and the amino acid D at position 455, and (ii) the amino acid sequence SPHSKA (SEQ

ID NO: 941), wherein the amino acid sequence of SPHSKA (SEQ ID NO: 941) is present immediately subsequent to position 455, wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0181] In some embodiments, an AAV capsid variant described herein comprises a modification, e.g., substitution, relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a modification, e.g., substitution, at position S454 and/or G455, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a S454H substitution and/or G455D substitution, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a S454H substitution and a G455D substitution, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises the amino acid sequence of SPHKSG (SEQ ID NO: 946). In some embodiments, the AAV capsid variant comprises: (i) a S454H substitution and a G455D substitution, and (ii) the amino acid sequence SPHKSG (SEQ ID NO: 946), wherein the amino acid sequence of SPHKSG (SEQ ID NO: 946) is present immediately subsequent to position 455, wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0182] In some embodiments, an AAV capsid variant described herein comprises a modification, e.g., substitution, relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a modification, e.g., substitution, at position S454 and/or G455, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a S454H substitution and/or G455D substitution, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises a S454H substitution and a G455D substitution, numbered relative to SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941). In some embodiments, the AAV capsid variant comprises: (i) a S454H substitution and a G455D substitution, and (ii) the amino acid sequence SPHSKA (SEQ ID NO: 941), wherein the amino acid sequence of SPHSKA (SEQ ID NO: 941) is present immediately subsequent to position 455, wherein (i) and (ii) are numbered according to SEQ ID NO: 138.

[0183] In some embodiments, the AAV capsid variant further comprises one, two, or all of an amino acid other than T at position 450 (e.g., S, Y, or G), an amino acid other than I at position 451 (e.g., M or L), and/or an amino acid other than N at position 452 (e.g., S), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises an S at position 450 and an M at position 451, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises a Y at position 450, an L at position 451, and an S at position 452, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises a G at position 450, an L at position 451, and an S at position 452, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138.

[0184] In some embodiments, the AAV capsid variant further comprises one, two, three, four, or all of an amino acid other than Q at position 456 (e.g., R or L), N at position 457 (e.g., H, K, or R), Q at position 458 (e.g., R or T), Q at position 459 (H), and/or T at position 460 (N or S), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises an R at position 456, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises an L at position 456, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises an H at position 457 and an R at position 458, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises a K at position 457 and an N at position 460, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises a T at position 458, an H at position 459, and an S at position 460, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises an R at position 456, an R at position 457, and an R at position 458, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138.

[0185] In some embodiments, an AAV capsid variant described herein comprises an amino acid other than I at position 451, an amino acid other than N at position 452, and an amino acid other than G at position 453, numbered according to SEQ ID NO: 138 or 981. In some embodiments, the AAV capsid variant comprises E at position 451, R at position 452, and V at position 453, numbered according to SEQ ID NO: 138 or 981. In some embodiments, the AAV capsid variant comprises the substitutions I451E, N452R, and G453V, numbered according to SEQ ID NO: 138 or 981.

[0186] In some embodiments, the AAV capsid variant comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), wherein the amino acid sequence is present immediately subsequent to position 455 and wherein the AAV capsid variant comprises the E at position 451, R at position 452, and V at position 453, numbered according to the amino acid sequence of SEQ ID NO: 138 or 981. In some embodiments, the AAV capsid variant comprises the substitutions I451E, N452R, and G453V, and further comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), wherein the amino acid sequence is present immediately subsequent to position 455, all numbered according to SEQ ID NO: 138 or 981. In some embodiments, the AAV capsid variant comprises the amino acid sequence of ERVSGSPHSKA (SEQ ID NO: 6399), and wherein the amino acid sequence is present immediately subsequent to position 449 and replaces positions 450-455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises the amino acid sequence of KTERVSGSPHSKAQNQQT (SEQ ID NO: 3589), wherein the amino acid sequence is present

immediately subsequent to position 448 and replaces positions 449-460, numbered according to SEQ ID NO: 138.

[0187] In some embodiments, an AAV capsid variant described herein comprises an amino acid other than T at position 450, an amino acid other than I at position 451, and an amino acid other than N at position 452, numbered according to SEQ ID NO: 138 or 982. In some embodiments, the AAV capsid variant comprises A at position 450, E at position 451, and I at position 452, numbered according to SEQ ID NO: 138 or 982. In some embodiments, the AAV capsid variant comprises the substitutions T450A, I451E, and N452I, numbered according to SEQ ID NO: 138 or 982.

[0188] In some embodiments, the AAV capsid variant comprises the amino acid sequence of SPHKSG (SEQ ID NO: 946), which is present immediately subsequent to positions 455, and further comprises A at position 450, E at position 451, I at position 452, H at position 454, and D at position 455, all numbered according to SEQ ID NO: 138 or 982. In some embodiments, the AAV capsid variant comprises the substitutions T450A, I451E, N452I, S454H, and G455D, and further comprises the amino acid sequence SPHKSG (SEQ ID NO: 946) present immediately subsequent to position 455, all numbered according to SEQ ID NO: 138 or 982. In some embodiments, the AAV capsid variant comprises the amino acid sequence of AEIGHDSPHKSG (SEQ ID NO: 6400), wherein the amino acid sequence is present immediately subsequent to position 449 and replaces positions 450-455, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises the amino acid sequence of KAEIGHDSPHKSGQNQQT (SEQ ID NO: 1754), wherein the amino acid sequence is present immediately subsequent to position 448 and replaces positions 449-460, numbered according to SEQ ID NO: 138.

[0189] In some embodiments, the AAV capsid variant, further comprises a substitution at position K449, e.g., a K449R substitution, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant, further comprises an amino acid other than K at position 449 (e.g., R), relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises an R at position 449, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further comprises a modification, e.g., an insertion, substitution, and/or deletion in loop I, II, VI, and/or VIII.

[0190] In some embodiments, the AAV capsid variant, further comprises an amino acid sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, of the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant, further comprises an amino acid sequence comprising at least one, two or three, but not more than 30, 20 or 10 amino acids that differ from the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid variant further

comprises the amino acid sequence of SEQ ID NO: 138, or an amino acid sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto.

[0191] In some embodiments, the AAV capsid variant further comprises (a) a VP1 protein comprising the amino acid sequence of SEQ ID NO: 138, 981, or 982; (b) a VP2 protein comprising the amino acid sequence of positions 138-736 of SEQ ID NO: 138 or positions 138-742 of SEQ ID NO: 981 or 982; (c) a VP3 protein comprising the amino acid sequence of positions 203-736 of SEQ ID NO: 138 or positions 203-742 of SEQ ID NO: 981 or 982; or (d) an amino acid sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity to any of the amino acid sequences in (a)-(c), an amino acid sequence comprising at least one, two or three, but not more than 30, 20 or 10 different amino acids relative to any of the amino acid sequences in (a)-(c), or an amino acid sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to any of the amino acid sequences in (a)-(c).

[0192] In some embodiments, the AAV capsid variant further comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 137, or a sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the AAV capsid variant further comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 137. In some embodiments, the AAV capsid variant further comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two or three, but not more than 30, 20 or 10 different nucleotides, relative to the amino acid sequence of SEQ ID NO: 137.

[0193] In some embodiments, the nucleotide sequence encoding the AAV capsid variant further comprises the nucleotide sequence of SEQ ID NO: 137, or a sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the nucleotide sequence encoding the AAV capsid variant further comprises a nucleotide sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 137. In some embodiments, the nucleotide sequence encoding the AAV capsid variant further comprises a nucleotide sequence comprising at least one, two or three, but not more than 30, 20 or 10 different nucleotides, relative to the amino acid sequence of SEQ ID NO: 137.

[0194] In some embodiments, an AAV capsid variant of the present disclosure comprises an amino acid sequence as described herein, e.g., an amino acid sequence of an AAV capsid variant of TTM-001 or TTM-002, e.g., as described in Tables 3 and 4.

[0195] In some embodiments, an AAV capsid variant described herein comprises a VP1, VP2, and/or VP3 protein comprising an amino acid sequence described herein, e.g., an amino acid sequence of an AAV capsid variant of TTM-001 or TTM-002, e.g., as described in Tables 3 and 4.

[0196] In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence encoded by a nucleotide sequence as described herein, e.g., a nucleotide sequence of an AAV capsid variant of TTM-001 or TTM-002, e.g., as described in Tables 3 and 5.

[0197] In some embodiments, a polynucleotide or nucleic acid encoding an AAV capsid variant, of the present disclosure comprises a nucleotide sequence described herein, e.g., a nucleotide sequence of an AAV capsid variant of TTM-001 or TTM-002, e.g., as described in Tables 3 and 5.

Table 3. Exemplary full length capsid sequences

Name	VP1 DNA SEQ ID NO:	VP1 (amino acid) SEQ ID NO:	Peptide (amino acid) SEQ ID NO:	Peptide DNA SEQ ID NO:
TTM-001	983	981	941	942
TTM-002	984	982	2	3

Table 4. Exemplary full length capsid amino acid sequences

Name and Annotation	SEQ ID NO:	Amino Acid Sequence
TTM-001 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); 742 aa	981	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTINGSG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-002 6mer peptide underlined, starts at position 454 (immediately subsequent to position 453); 742 aa	982	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHK**SGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK

		LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-003 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 452, and 453 underlined; 742 aa	36	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**ERV**SG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-004 6mer peptide underlined, starts at position 454 (immediately subsequent to position 453); modifications at positions 450, 451, and 452 underlined; 742 aa	37	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS K**AEI**G**HDSPHK**SGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSLITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-005 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 452, 464, and 465 underlined; 742 aa	38	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTI**I**GSG**SPHSKA**QN**RH**TLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-006 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 452, and 453 underlined; 742 aa	39	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**EKM**SG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-007 6mer peptide underlined,	40	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK

starts at position 456 (immediately subsequent to position 455); modifications at positions 450 and 454 underlined; 742 aa		KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS K**E**ING**R**G**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-008 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 462, and 464 underlined; 742 aa	41	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**F**NGSG**SPHSKAP**N**L**QTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-009 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 452, and 453 underlined; 742 aa	42	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**EKT**SG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-010 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 454, and 455 underlined; 742 aa	43	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**M**NG**HDSPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-011 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455);	44	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS

modifications at positions 452, 454, and 455 underlined; 742 aa		KTI**D**G**HDSPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-012 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 452, 454, and 455 underlined; 742 aa	45	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTN**N**G**HDSPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-013 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 452, and 453 underlined; 742 aa	46	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**QRK**SG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-014 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 463, 464, and 465 underlined; 742 aa	47	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTINGSG**SPHSKA**Q**ARK**TLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-015 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 450 and 452 underlined; 742 aa	48	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS K**Y**I**V**GSG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK

		LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-016 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 452, 453, and 454 underlined; 742 aa	49	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTI**SKR**G**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-017 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 450, 451, and 452 underlined; 742 aa	50	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS K**GLG**GSG**SPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-018 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 454, 455, and 464 underlined; 742 aa	51	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHSKA**QN**L**QTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-019 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 451, 454, and 455 underlined; 742 aa	52	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KT**V**NG**HDSPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-020 6mer peptide underlined,	53	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK

starts at position 456 (immediately subsequent to position 455); modifications at positions 454, 455, and 462 underlined; 742 aa		KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHSKAL**NQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-021 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 454, 455, and 466 underlined; 742 aa	54	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHSKA**QNQQ**S**LKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-022 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modifications at positions 454, 455, and 466 underlined; 742 aa	55	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHSKA**QNQQ**I**LKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-023 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); 742 aa	56	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTINGSG**SPHFTR**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-024 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455);	57	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS

modifications at positions 451, 454, and 455 underlined; 742 aa		KT**S**NG**HDSPHSKA**QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-025 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modification at position 462 underlined; 742 aa	58	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTINGSG**SPHSLPW**NQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL
TTM-026 6mer peptide underlined, starts at position 456 (immediately subsequent to position 455); modification at positions 454, 455, and 464 underlined; 742 aa	59	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPG NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGG NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAK KRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSS SGNWHCDSQWLGDRVITTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTP WGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIAN NLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRS SFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLS KTING**HDSPHSKA**QN**H**QTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNN SEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVD ADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDR DVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDK LNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYS EPRPIGTRYLTRNL

Table 5. Exemplary full length capsid nucleic acid sequences

Name and Annotation	SEQ ID NO:	NT Sequence
TTM-001 9mer peptide underlined	983	ATGGCTGCCGATGGTTATCTTCCAGattggcTCGAGGACAACCTTAGTGAAGGAAT TCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAAC ATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGC AACGGACTCGACAAGGGGGAGCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCA CGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACA ACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGC AACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCT GGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTC CTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAA AAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCAACC AATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAG GTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCC TCGGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAG CACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCA ACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCC TGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCA GCGACTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCT TCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAAT AACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGT GCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGA TTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCG

		TCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACTT CCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCC AAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTAtTActTgagt AAaACaATTAACGGAAGCGGA**AGCCCACACAGCAAAGCA**CAAAACCAACAGACCtT gAAgTTttcgGTaGCtGGtCCtAGCAACATGGCTGTCCAGGGAAGAAACTACATAC CTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAAC AGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTT GATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTC CTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGAT GCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGC AACGGAGTCCTATGGACAAGtggccacaaaccaccagagtGCCCAAGCACAGGCGC AGaccggctgggttcaaaaccaAGGAATACTTCCGGGTATGGTTTGGCAGGACAGA GATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTT TCACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCC TCATCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAG CTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTG GGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCA ACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGT GAACCCCGCCCCATTGGCACgcGgTAttTAACgaGgAActTaTAA
TTM-002 7mer peptide underlined	984	ATGGCTGCCGATGGTTATCTTCCAGattggcTCGAGGACAACCTTAGTGAAGGAAT TCGCGAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAAC ATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGC AACGGACTCGACAAGGGGGAGCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCA CGACAAGGCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACA ACCACGCCGACGCCGAGTTCCAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGC AACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTCTTGGTCT GGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTC CTCAGGAACCGGACTCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGCTAAA AAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTCCCAGACCCTCAACC AATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCTCTTACAATGGCTTCAG GTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGGAGTGGGTAGTTCC TCGGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATCACCACCAG CACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATCTCCA ACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCC TGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCA GCGACTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCT TCAACATTCAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAAT AACCTTACCAGCACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGT GCTCGGGTCGGCTCACGAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGA TTCCTCAGTACGGGTATCTGACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCG TCCTTTTACTGCCTGGAATATTTCCCGTCGCAAATGCTAAGAACGGGTAACAACTT CCAGTTCAGCTACGAGTTTGAGAACGTACCTTTCCATAGCAGCTACGCTCACAGCC AAAGCCTGGACCGACTAATGAATCCACTCATCGACCAATACTTGTAtTActTgagt AAaACaATTAACGGA**CACGACAGCCCACACAAAAGCGGA**CAAAACCAACAGACCtT gAAgTTttcgGTaGCtGGtCCtAGCAACATGGCTGTCCAGGGAAGAAACTACATAC CTGGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAAC AGCGAATTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTT GATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTC CTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGAT GCGGACAAAGTCATGATAACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGC AACGGAGTCCTATGGACAAGtggccacaaaccaccagagtGCCCAAGCACAGGCGC AGaccggctgggttcaaaaccaAGGAATACTTCCGGGTATGGTTTGGCAGGACAGA GATGTGTACCTGCAAGGACCCATTTGGGCCAAAATTCCTCACACGGACGGCAACTT TCACCCTTCTCCGCTGATGGGAGGGTTTGGAATGAAGCACCCGCCTCCTCAGATCC TCATCAAAAACACACCTGTACCTGCGGATCCTCCAACGGCCTTCAACAAGGACAAG CTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGCGTGGAGATCGAGTG GGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAGTACACTTCCA ACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTATATAGT GAACCCCGCCCCATTGGCACgcGgTAttTAACgaGgAActTaTAA

[0198] In some embodiments, the polynucleotide encoding an AAV capsid variant, described herein comprises the nucleotide sequence of SEQ ID NO: 983 or 984, or a nucleotide sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto.

[0199] In some embodiments, the polynucleotide encoding an AAV capsid variant described herein comprises the nucleotide sequence of SEQ ID NO: 983, or a nucleotide sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein, comprises a nucleotide sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 983. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein, comprises a nucleotide sequence comprising at least one, two or three, but not more than 30, 20 or 10 different nucleotides relative to the amino acid sequence of SEQ ID NO: 983. In some embodiments, the nucleic acid sequence encoding an AAV capsid variant described herein is codon optimized.

[0200] In some embodiments, the polynucleotide encoding an AAV capsid variant described herein comprises the nucleotide sequence of SEQ ID NO: 984, or a nucleotide sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein, comprises a nucleotide sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 984. In some embodiments, the nucleotide sequence encoding an AAV capsid variant described herein, comprises a nucleotide sequence comprising at least one, two or three, but not more than 30, 20 or 10 different nucleotides, relative to the amino acid sequence of SEQ ID NO: 984. In some embodiments, the nucleic acid sequence encoding an AAV capsid variant described herein is codon optimized.

[0201] In some embodiments, an AAV capsid variant described herein, comprises the amino acid sequence of SEQ ID NO: 981, or an amino acid sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence comprising at least one, two, or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SEQ ID NO: 981. In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence comprising at least one, two or

three, but not more than 30, 20 or 10 different amino acids, relative to the amino acid sequence of SEQ ID NO: 981.

[0202] In some embodiments, an AAV capsid variant described herein comprises the amino acid sequence of SEQ ID NO: 982, or an amino acid sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence comprising at least one, two, or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the amino acid sequence of SEQ ID NO: 982. In some embodiments, the AAV capsid variant, comprises an amino acid sequence comprising at least one, two or three, but not more than 30, 20 or 10 different amino acids, relative to the amino acid sequence of SEQ ID NO: 982.

[0203] In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 983 or 984, or a nucleotide sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two or three, but not more than 30, 20 or 10 different nucleotides, relative to the amino acid sequence of SEQ ID NO: 983 or 984. In some embodiments, an AAV capsid variant described herein comprises an amino acid sequence encoded by a nucleotide sequence comprising at least one, two or three modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, but not more than 30, 20 or 10 modifications, e.g., substitutions (e.g., conservative substitutions), insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 983 or 984.

[0204] In some embodiments, an AAV capsid variant described herein comprises a VP1, VP2, VP3 protein, or a combination thereof. In some embodiments, an AAV capsid variant comprises the amino acid sequence corresponding to positions 138-742, e.g., a VP2, of SEQ ID NO: 981 or 982, or a sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the AAV capsid protein comprises the amino acid sequence corresponding to positions 203-742, e.g., a VP3, of SEQ ID NO: 981 or 982, or a sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto. In some embodiments, the AAV capsid variant comprises the amino acid sequence corresponding to positions 1-742, e.g., a VP1, of SEQ ID NO: 981 or 982, or an amino acid sequence with at least 70% (e.g., at least about 80, 85, 90, 95, 96, 97, 98, or 99%) sequence identity thereto.

[0205] In some embodiments, an AAV capsid variant, described herein has an increased tropism for a CNS cell or tissue, e.g., a brain cell, brain tissue, spinal cord cell, or spinal cord tissue, relative to the tropism of a reference sequence comprising the amino acid sequence of SEQ ID NO: 138.

[0206] In some embodiments, an AAV capsid variant described herein transduces a brain region, e.g., a midbrain region (e.g., the hippocampus, or thalamus) or the brain stem. In some embodiments, the level of transduction is at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, or 65-fold greater as compared to a reference sequence of SEQ ID NO: 138. In some embodiments, the level of transduction is at least 30, 35, 40, 45, 50, 55, 60, or 65-fold greater as compared to a reference sequence of SEQ ID NO: 138.

[0207] In some embodiments, an AAV capsid variant described herein is enriched at least about 3, 4, 5, 6, 7, 8, 9, or 10-fold in the brain compared to a reference sequence of SEQ ID NO: 138. In some embodiments, an AAV capsid variant described herein is enriched at least about 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or 85-fold in the brain compared to a reference sequence of SEQ ID NO: 138.

[0208] In some embodiments, an AAV capsid variant described herein is enriched in the brain of at least two to three species, e.g., a non-human primate and rodent (e.g., mouse) species, compared to a reference sequence of SEQ ID NO: 138. In some embodiments, an AAV capsid variant described herein is enriched at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100-fold in the brain of at least two to three species, e.g., a non-human primate and rodent (e.g., mouse) species, compared to a reference sequence of SEQ ID NO: 138. In some embodiments, the at least two to three species are *Macaca fascicularis*, *Chlorocebus sabaeus*, *Callithrix jacchus*, and/or mouse (e.g., BALB/c mice, C57Bl/6 mice, and/or CD-1 outbred mice).

[0209] In some embodiments, an AAV capsid variant described herein is enriched at least about 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, or 8-fold, in the brain compared to a reference sequence of SEQ ID NO: 981. In some embodiments, an AAV capsid variant described herein is enriched about 2, 2.5, 3, 3.5, 4, 4.5, 5, or 5.5-fold, in the brain compared to a reference sequence of SEQ ID NO: 982.

[0210] In some embodiments, an AAV capsid variant described herein delivers an increased level of viral genomes to a brain region. In some embodiments, the level of viral genomes is increased by at least 20, 25, 30, 35, 40, 45, or 50-fold, as compared to a reference sequence of SEQ ID NO: 138. In some embodiments, the brain region comprises a midbrain region (e.g., the hippocampus or thalamus) and/or the brainstem.

[0211] In some embodiments, an AAV capsid variant described herein delivers an increased level of a payload to a brain region. In some embodiments, the level of the payload is increased by at least 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, or 70-fold, as compared to a reference sequence of SEQ ID NO: 138. In some embodiments, the brain region comprises a midbrain region (e.g., the hippocampus or thalamus) and/or the brainstem.

[0212] In some embodiments, an AAV capsid variant described herein is enriched at least about 5, 10, 15, 20, 25, 30, or 35-fold, in the spinal cord compared to a reference sequence of SEQ ID NO: 138.

[0213] In some embodiments, an AAV capsid variant described herein shows preferential transduction in a brain region relative to the transduction in the dorsal root ganglia (DRG). In some embodiments, the AAV capsid variant shows preferential transduction in a brain region relative to the transduction in the liver. In some embodiments, the AAV capsid variant shows preferential transduction in a brain region relative to the transduction in the liver and the DRG. In some embodiments, the AAV capsid variant shows preferential transduction in a brain region relative to the transduction in the heart. In some embodiments, the AAV capsid variant shows preferential transduction in a brain region relative to the transduction in the heart. and DRG. In some embodiments, the AAV capsid variant shows preferential transduction in a brain region relative to the transduction in the heart, DRG, and liver.

[0214] In some embodiments, an AAV capsid variant described herein is capable of transducing non-neuronal cells, e.g., glial cells (e.g., oligodendrocytes or astrocytes). In some embodiments, the AAV capsid variant described herein is capable of transducing neuronal cells and non-neuronal cells, e.g., glial cells (e.g., oligodendrocytes or astrocytes). In some embodiments, the non-neuronal cells are glial cells, oligodendrocytes (e.g., Olig2 positive oligodendrocytes), or astrocytes (e.g., Olig2 positive astrocytes). In some embodiments, the AAV capsid variant is capable of transducing Olig2 positive cells, e.g., Olig2 positive astrocytes or Olig2 positive oligodendrocytes.

[0215] In some embodiments, an AAV capsid variant described herein is capable of binding to a glycosylphosphatidylinositol (GPI) anchored protein, e.g., alkaline phosphatase (ALPL). In some embodiments, the GPI anchored protein is conserved in at least two to three species, e.g., at least three species (e.g., mice, NHPs (e.g., *Macaca fascicularis*), and/or humans). In some embodiments, the GPI anchored protein is present on the surface of a cell in the blood brain barrier. In some embodiments, the GPI anchored protein is ALPL. In some embodiments, the AAV capsid variant is capable of binding N-linked galactose. In some embodiments, binding to ALPL results in increased cellular transduction, e.g., as compared to a reference sequence of SEQ ID NO: 138. In some embodiments, binding to ALPL results in increased crossing of the blood brain barrier, e.g., as compared to a reference sequence of SEQ ID NO: 138. Without wishing to be bound by theory, it is believed in some embodiments, that the binding of the AAV capsid variants described herein to ALPL is part of the mechanism leading to increased crossing of the blood brain barrier relative to the AAV9 control. Without wishing to be bound by theory, it is believed in some embodiments, that ALPL is upregulated in aging brain (e.g., as described in Yang et al. "Physiological blood–brain transport is impaired with age by a shift in transcytosis," *Nature*. 2020 583:425-430, the contents of which are hereby incorporated by reference in its entirety).

[0216] In some embodiments, an AAV capsid variant of the present disclosure is isolated, e.g., recombinant. In some embodiments, a polynucleotide encoding an AAV capsid polypeptide, e.g., an AAV capsid variant, of the present disclosure is isolated, e.g., recombinant.

[0217] Also provided herein are polynucleotide sequences encoding any of the AAV capsid variants described above and AAV particles, vectors, and cells comprising the same.

Additional AAV Sequences

[0218] In some embodiments, the AAV capsid variant, comprises immediately subsequent to position 448, 452, 453, 455, numbered relative to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety)), at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 consecutive amino acids of any of amino acid sequence provided in Tables 1, 2A, 2B, 2C, 13-19. In some embodiments, the amino acid sequence replaces at least one, two, three, four, five, six, seven, eight, nine or all of positions T450, I451, N452, G453, S454, G455, Q456, N457, Q458, and/or Q459, numbered according to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety)). In some embodiments, the amino acid sequence replaces positions S454, G455, or both positions S454 and G455, numbered according to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety)). In some embodiments, the AAV capsid variant comprises an amino acid other than the wild-type, e.g., native, amino acid, at one, two, three, four, five, six, seven, eight, nine or all of positions T450, I451, N452, G453, S454, G455, Q456, N457, Q458, and/or Q459, numbered according to SEQ ID NO: 138. In some embodiments, the AAV capsid variant comprises an amino acid other than the wild-type, e.g., native, amino acid, at position S454, G455, or both positions S454 and G455, numbered according to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety)). In some embodiments, the AAV capsid variant comprises a modification, e.g., substitution, at one, two, three, four, five, six, seven, eight, nine or all of positions T450, I451, N452, G453, S454, G455, Q456, N457, Q458, and/or Q459, numbered according to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of

which are hereby incorporated by reference in their entirety)). In some embodiments, the AAV capsid variant comprises a modification, e.g., substitution, at position S454, G455, or both positions S454 and G455, numbered according to SEQ ID NO: 138 or corresponding to equivalent positions in any other AAV serotype (e.g., AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAVrh10, AAVrh32.33, AAVrh74, SEQ ID NO: 1, SEQ ID NO: 11, PHP.N, PHP.B, or an AAV serotype as provided in Table 6 of WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety)).

[0219] In some embodiments, an AAV capsid polypeptide or AAV capsid variant described herein may comprise a VOY101 capsid polypeptide, an AAVPHP.B (PHP.B) capsid polypeptide, a AAVPHP.N (PHP.N) capsid polypeptide, an AAV1 capsid polypeptide, an AAV2 capsid polypeptide, an AAV5 capsid polypeptide, an AAV9 capsid polypeptide, an AAV9 K449R capsid polypeptide, an AAVrh10 capsid polypeptide, or a functional variant thereof. In some embodiments, the AAV capsid polypeptide, e.g., AAV capsid variant, comprises an amino acid sequence of any of the AAV capsid polypeptides in Table 6, or an amino acid sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the nucleotide sequence encoding the AAV capsid polypeptide comprises any one of the nucleotide sequences in Table 6, or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto.

[0220] In some embodiments, an AAV capsid polypeptide or an AAV capsid variant described herein comprises the amino acid sequence of SEQ ID NO: 138 or an amino acid sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments the AAV capsid polypeptide or the AAV capsid variant, comprises an amino acid sequence comprising at least one, two, or three modifications, e.g., substitutions (e.g., conservative substitutions), but no more than 30, 20, or 10 modifications, e.g., substitutions (e.g., conservative substitutions), relative to the amino acid sequence of SEQ ID NO: 138. In some embodiments, the AAV capsid polypeptide or the AAV capsid variant, comprises an amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 137 or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the nucleotide sequence encoding the AAV capsid polypeptide or the AAV capsid variant comprises the nucleotide sequence of SEQ ID NO: 137 or a nucleotide sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments, the AAV capsid polypeptide or the AAV capsid variant, comprises substitution at position K449, e.g., a K449R substitution, numbered according to SEQ ID NO: 138.

[0221] In some embodiments, the AAV capsid polypeptide or the AAV capsid variant, comprises a peptide comprising the amino acid sequence of TLAVPFK (SEQ ID NO: 4680). In some

embodiments, the peptide is present immediately subsequent to position 588, relative to a reference sequence numbered according to SEQ ID NO: 138. In some embodiments, the capsid polypeptide comprises the amino acid substitutions of A587D and Q588G, numbered according to SEQ ID NO: 138.

[0222] In some embodiments, the AAV capsid polypeptide or the AAV capsid variant comprises the amino acid substitution of K449R, numbered according to SEQ ID NO: 138; and a peptide comprising the amino acid sequence of TLAVPFK (SEQ ID NO: 4680), wherein the peptide is present immediately subsequent to position 588, relative to a reference sequence numbered according to SEQ ID NO: 138.

[0223] In some embodiments, the AAV capsid polypeptide or the AAV capsid variant comprises the amino acid substitution of K449R, numbered according to SEQ ID NO: 138; an peptide comprising the amino acid sequence of TLAVPFK (SEQ ID NO: 4680), wherein the insert is present immediately subsequent to position 588, relative to a reference sequence numbered according to SEQ ID NO: 138; and the amino acid substitutions of A587D and Q588G, numbered according to SEQ ID NO: 138.

[0224] In some embodiments, the AAV capsid polypeptide or the AAV capsid variant comprises a peptide comprising the amino acid sequence of TLAVPFK (SEQ ID NO: 4680), wherein the insert is present immediately subsequent to position 588, relative to a reference sequence numbered according to SEQ ID NO: 138; and the amino acid substitutions of A587D and Q588G, numbered according to SEQ ID NO: 138.

[0225] In some embodiments, the AAV capsid polypeptide or the AAV capsid variant comprises the amino acid sequence of SEQ ID NO: 11 or an amino acid sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments the AAV capsid polypeptide or the AAV capsid variant, comprises an amino acid sequence comprising at least one, two, or three modifications, e.g., substitutions (e.g., conservative substitutions), but no more than 30, 20, or 10 modifications, e.g., substitutions (conservative substitutions), relative to the amino acid sequence of SEQ ID NO: 11, optionally wherein position 449 is not R.

[0226] In some embodiments, the AAV capsid polypeptide or AAV capsid variant, comprises the amino acid sequence of SEQ ID NO: 1 or an amino acid sequence substantially identical (e.g., having at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, or 99% sequence identity) thereto. In some embodiments the AAV capsid polypeptide or the AAV capsid variant, comprises an amino acid sequence comprising at least one, two, or three modifications, e.g., substitutions (e.g., conservative substitutions), but no more than 30, 20, or 10 modifications, e.g., substitutions (e.g., conservative substitutions), relative to the amino acid sequence of SEQ ID NO: 1.

Table 6. AAV Sequences

Serotype	SEQ ID NO:	Sequence
VOY101	1	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE SVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI TTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR LINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAH EGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV PFHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIP GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSDGTLAVPFKAQAQT GWVQNQGILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTP VPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEF AVNTEGVYSEPRPIGTRYLTRNL
AAV9/hu.14 K449R	11	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE SVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI TTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR LINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAH EGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV PFHSSYAHSQSLDRLMNPLIDQYLYYLSRTINGSGQNQQTLKFSVAGPSNMAVQGRNYIP GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGV YSEPRPIGTRYLTRNL
AAV9/hu.14 WT (amino acid)	138	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE SVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI TTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR LINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAH EGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV PFHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIP GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGV YSEPRPIGTRYLTRNL
AAV9/hu.14 WT (DNA)	137	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCGC GAGTGGTGGGCTTTGAAACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGAC AACGCTCGAGGTCTTGTGCTTCCGGGTTACAAATACCTTGGACCCGGCAACGGACTCGAC AAGGGGGAGCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAGGCCTACGAC CAGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTTC CAGGAGCGGCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG GCCAAAAAGAGGCTTCTTGAACCTCTTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCT GGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGACTCCTCCGCGGGTATTGGC AAATCGGGTGCACAGCCCGCTAAAAAGAGACTCAATTTCGGTCAGACTGGCGACACAGAG TCAGTCCCAGACCCTCAACCAATCGGAGAACCTCCCGCAGCCCCCTCAGGTGTGGGATCT CTTACAATGGCTTCAGGTGGTGGCGCACCAGTGGCAGACAATAACGAAGGTGCCGATGGA GTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTCCCAATGGCTGGGGGACAGAGTCATC ACCACCAGCACCCGAACCTGGGCCCTGCCCACCTACAACAATCACCTCTACAAGCAAATC TCCAACAGCACATCTGGAGGATCTTCAAATGACAACGCCTACTTCGGCTACAGCACCCCC TGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGA CTCATCAACAACAACTGGGGATTCCGGCCTAAGCGACTCAACTTCAAGCTCTTCAACATT CAGGTCAAAGAGGTTACGGACAACAATGGAGTCAAGACCATCGCCAATAACCTTACCAGC

		ACGGTCCAGGTCTTCACGGACTCAGACTATCAGCTCCCGTACGTGCTCGGGTCGGCTCAC GAGGGCTGCCTCCCGCCGTTCCCAGCGGACGTTTTCATGATTCCTCAGTACGGGTATCTG ACGCTTAATGATGGAAGCCAGGCCGTGGGTCGTTCGTCCTTTTACTGCCTGGAATATTTC CCGTCGCAAATGCTAAGAACGGGTAACAACTTCCAGTTCAGCTACGAGTTTGAGAACGTA CCTTTCCATAGCAGCTACGCTCACAGCCAAAGCCTGGACCGACTAATGAATCCACTCATC GACCAATACTTGTACTATCTCTCAAAGACTATTAACGGTTCTGGACAGAATCAACAAACG CTAAAATTCAGTGTGGCCGGACCCAGCAACATGGCTGTCCAGGGAAGAAACTACATACCT GGACCCAGCTACCGACAACAACGTGTCTCAACCACTGTGACTCAAAACAACAACAGCGAA TTTGCTTGGCCTGGAGCTTCTTCTTGGGCTCTCAATGGACGTAATAGCTTGATGAATCCT GGACCTGCTATGGCCAGCCACAAAGAAGGAGAGGACCGTTTCTTTCCTTTGTCTGGATCT TTAATTTTTGGCAAACAAGGAACTGGAAGAGACAACGTGGATGCGGACAAAGTCATGATA ACCAACGAAGAAGAAATTAAAACTACTAACCCGGTAGCAACGGAGTCCTATGGACAAGTG GCCACAAACCACCAGAGTGCCCAAGCACAGGCGCAGACCGGCTGGGTTCAAAACCAAGGA ATACTTCCGGGTATGGTTTGGCAGGACAGAGATGTGTACCTGCAAGGACCCATTTGGGCC AAAATTCCTCACACGGACGGCAACTTTCACCCTTCTCCGCTGATGGGAGGGTTTGGAATG AAGCACCCGCCTCCTCAGATCCTCATCAAAAACACACCTGTACCTGCGGATCCTCCAACG GCCTTCAACAAGGACAAGCTGAACTCTTTCATCACCCAGTATTCTACTGGCCAAGTCAGC GTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCCGGAGATCCAG TACACTTCCAACTATTACAAGTCTAATAATGTTGAATTTGCTGTTAATACTGAAGGTGTA TATAGTGAACCCCGCCCCATTGGCACCAGATACCTGACTCGTAATCTGTAA

Viral Genome of the AAV particle

[0227] In some embodiments, an AAV particle as described herein comprising an AAV capsid variant described herein, may be used for the delivery of a viral genome to a tissue (e.g., CNS, DRG, and/or muscle). In some embodiments, an AAV particle comprising an AAV capsid variant described herein can be used for delivery of a viral genome to a tissue or cell, e.g., CNS, DRG, or muscle cell or tissue. In some embodiments, an AAV particle of the present disclosure is a recombinant AAV particle. In some embodiments, an AAV particle of the present disclosure is an isolated AAV particle.

[0228] The viral genome may encode any payload, such as but not limited to a polypeptide (e.g., a therapeutic polypeptide), an antibody, an enzyme, an RNAi agent and/or components of a gene editing system. In one embodiment, the AAV particles described herein are used to deliver a payload to cells of the CNS, after intravenous delivery. In another embodiment, the AAV particles described herein are used to deliver a payload to cells of the DRG, after intravenous delivery. In some embodiments, the AAV particles described herein are used to deliver a payload to cells of a muscle, e.g., a heart muscle, after intravenous delivery.

[0229] In some embodiments, a viral genome of an AAV particle comprising an AAV capsid variant, as described herein, comprises a nucleotide sequence comprising a transgene encoding a payload. In some embodiments, the viral genome comprises an inverted terminal repeat sequence (ITR). In some embodiments, the viral genome comprises two ITR sequences, one at the 5' end of the viral genome (e.g., 5' relative to the encoded payload) and one at the 3' end of the viral genome (e.g., 3' relative to the encoded payload). In some embodiments, a viral genome of an AAV particle, e.g., an AAV particle comprising an AAV capsid variant described herein, may comprise a regulatory element (e.g., promoter), untranslated regions (UTR), a miR binding site, a polyadenylation sequence

(polyA), a filler or stuffer sequence, an intron, and/or a linker sequence, e.g., for enhancing transgene expression.

[0230] In some embodiments, the viral genome components are selected and/or engineered for expression of the payload in a target tissue (e.g., CNS, muscle, or DRG).

Viral Genome Component: Inverted Terminal Repeats (ITRs)

[0231] In some embodiments, the AAV particle comprising an AAV capsid variant described herein comprises a viral genome comprising an ITR and a transgene encoding a payload. In some embodiments, the viral genome comprises two ITRs. In some embodiments, the two ITRs flank the nucleotide sequence encoding the payload at the 5' and 3' ends. In some embodiments, the ITRs function as origins of replication comprising recognition sites for replication. In some embodiments, the ITRs comprise sequence regions which can be complementary and symmetrically arranged. In some embodiments, the ITRs incorporated into viral genomes as described herein may be comprised of naturally occurring polynucleotide sequences or recombinantly derived polynucleotide sequences.

[0232] In some embodiments, the ITR may be from the same serotype as the capsid polypeptide, e.g., capsid variant, selected from any of the known serotypes, or a variant thereof. In some embodiments, the ITR may be of a different serotype than the capsid. In some embodiments, the viral genome comprises two ITR sequence regions, wherein the ITRs are of the same serotype as one another. In some embodiments, the viral genome comprises two ITR sequence regions, wherein the ITRs are of different serotypes. Non-limiting examples include zero, one or both of the ITRs having the same serotype as the capsid. In one embodiment both ITRs of the viral genome of the AAV particle are AAV2 ITRs.

Viral Genome Component: Promoters

[0233] In some embodiments, viral genome of an AAV particle described herein comprises at least one element to enhance the payload target specificity and expression (See e.g., Powell et al. Viral Expression Cassette Elements to Enhance Transgene Target Specificity and Expression in Gene Therapy, 2015; the contents of which are herein incorporated by reference in their entirety). Non-limiting examples of elements to enhance payload target specificity and expression include promoters, endogenous miRNAs, post-transcriptional regulatory elements (PREs), polyadenylation (PolyA) signal sequences and upstream enhancers (USEs), CMV enhancers and introns.

[0234] In some embodiments, an AAV particle comprising an AAV capsid variant described herein comprises a viral genome comprising a nucleic acid comprising a transgene encoding a payload, wherein the transgene is operably linked to a promoter. In some embodiments, the promoter is a species specific promoter, an inducible promoter, a tissue-specific promoter, or a cell cycle-specific promoter (e.g., a promoter as described in Parr et al., *Nat. Med.*3:1145-9 (1997); the contents of which are herein incorporated by reference in their entirety).

[0235] In some embodiments, the Promoter may be naturally occurring or non-naturally occurring. Non-limiting examples of promoters include those derived from viruses, plants, mammals, or humans. In some embodiments, the promoters may be those derived from human cells or systems. In some embodiments, the promoter may be truncated or mutated, e.g., a promoter variant.

[0236] In some embodiments, the promoter is a ubiquitous promoter, e.g., capable of expression in multiple tissues. In some embodiments the promoter is an human elongation factor 1α-subunit (EF1α) promoter, the cytomegalovirus (CMV) immediate-early enhancer and/or promoter, the chicken β-actin (CBA) promoter and its derivative CAG, β glucuronidase (GUSB) promoter, or ubiquitin C (UBC) promoter. In some embodiments, the promoter is a cell or tissue specific promoter, e.g., capable of expression in tissues or cells of the central or peripheral nervous systems, targeted regions within (e.g., frontal cortex), and/or sub-sets of cells therein (e.g., excitatory neurons). In some embodiments, the promoter is a cell-type specific promoters capable of expression of a payload in excitatory neurons (e.g., glutamatergic), inhibitory neurons (e.g., GABA-ergic), neurons of the sympathetic or parasympathetic nervous system, sensory neurons, neurons of the dorsal root ganglia, motor neurons, or supportive cells of the nervous systems such as microglia, glial cells, astrocytes, oligodendrocytes, and/or Schwann cells.

[0237] In some embodiments, the promoter is a liver specific promoter (e.g., hAAT, TBG), skeletal muscle specific promoter (e.g., desmin, MCK, C512), B cell promoter, monocyte promoter, leukocyte promoter, macrophage promoter, pancreatic acinar cell promoter, endothelial cell promoter, lung tissue promoter, and/or cardiac or cardiovascular promoter (e.g., αMHC, cTnT, and CMV-MLC2k).

[0238] In some embodiments, the promoter is a tissue-specific promoter for payload expression in a tissue or cell of the central nervous system. In some embodiments, the promoter is a synapsin (Syn) promoter, glutamate vesicular transporter (VGLUT) promoter, vesicular GABA transporter (VGAT) promoter, parvalbumin (PV) promoter, sodium channel Na_v 1.8 promoter, tyrosine hydroxylase (TH) promoter, choline acetyltransferase (ChaT) promoter, methyl-CpG binding protein 2 (MeCP2) promoter, Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII) promoter, metabotropic glutamate receptor 2 (mGluR2) promoter, neurofilament light (NFL) or heavy (NFH) promoter, neuron-specific enolase (NSE) promoter, β-globin minigene nβ2 promoter, preproenkephalin (PPE) promoter, enkephalin (Enk) promoter, and excitatory amino acid transporter 2 (EAAT2) promoter, or a fragment thereof. In some embodiments, the promoter is a cell-type specific promoter capable of expression in an astrocyte, e.g., a glial fibrillary acidic protein (GFAP) promoter and a EAAT2 promoter, or a fragment thereof. In some embodiments, the promoter is a cell-type specific promoter capable of expression in an oligodendrocyte, e.g., a myelin basic protein (MBP) promoter or a fragment thereof.

[0239] In some embodiments, the promoter is a GFAP promoter. In some embodiments, the promoter is a synapsin (syn or syn1) promoter, or a fragment thereof.

[0240] In some embodiments, the promoter comprises an insulin promoter or a fragment thereof.

[0241] In some embodiments, the promoter of the viral genome described herein (e.g., comprised within an AAV particle comprising an AAV capsid variant described herein) comprises an EF-1α promoter or variant thereof, e.g., as provided in Table 8. In some embodiments, the EF-1α promoter comprises the nucleotide sequence of any one of SEQ ID NOs: 987, 988, 990, 991, 995, 996, 998-1007 or any one of the sequences provided in Table 8, a nucleotide sequence comprising at least one, two, or three but no more than four modifications, e.g., substitutions, relative to the nucleotide sequence of SEQ ID NOs: 987, 988, 990, 991, 995, 996, 998-1007 or any one of the sequences provided in Table 8, or a nucleotide sequence with at least 70% (e.g., 80, 85%, 90%, 95%, 96%, 97%, 98%, or 99%) sequence identity to any one of SEQ ID NOs: 987, 988, 990, 991, 995, 996, 998-1007 or any one of the sequences provided in Table 8.

Table 8. Exemplary Promoter Variants

Description	Sequences	SEQ ID NO:
EF1a Promoter (intron underlined)	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGA GAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGG GTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGG GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTT TGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTT ACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGA TTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGC TTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGC CGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGT CTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGA TAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGGGGCC GCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCGGCGAGGCGGGGCC TGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGC TCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTG GCCCGGTCGGCACCAGTTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTG CAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACC CACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGT CGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTG GGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTT GCCCTTTTTGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAA GTTTTTTTCTTCCATTTCAGGTGTCGTGA	987
miniEF1a	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAG TGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATA AGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC GCGTAAG	988
Promoter Variant 1	GCATG	
Promoter Variant 2	GGTGGAGAAGAGCATG	990
Promoter Variant 3	GTCATCACTGAGGTGGAGAAGAGCATG	991
Promoter Variant 4	CGTGAG	
Promoter Variant 5	GT	
Promoter Variant 6	GCTCCGGT	

Promoter Variant 19	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAG TGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATA AGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC AG	995
Promoter Variant 20	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAG TGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATA AGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC GC	996
Promoter Variant 7	GTAAG	
Promoter Variant 8	GTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGG AGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAA AGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATA TAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAAC ACGCGTAAG	998
Promoter Variant 9	GCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTT GGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAAC TGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAAC CGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGC CAGAACACGCGTAAG	999
Promoter Variant 10	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGA GAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGG GTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGG GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTT TGCCGCCAGAACACGCGTAAG	1000
Promoter Variant 11	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGA GAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGG GTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGG GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTT TGCCGCCAGAACACAG	1001
Promoter Variant 12	GCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTC CCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGC GCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGG TGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAAC GGGTTTGCCGCCAGAACACGCGTAAG	1002
Promoter Variant 13	GCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTC CCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGC GCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGG TGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAAC GGGTTTGCCGCCAGAACACAG	1003
Promoter Variant 14	GGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACAT CGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTA GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTT TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTC TTTTTCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	1004
Promoter Variant 15	GGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACAT CGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTA GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTT TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTC TTTTTCGCAACGGGTTTGCCGCCAGAACACAG	1005
Promoter Variant 16	GTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGC AGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGA ACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACT GGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGC CGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	1006
Promoter Variant 18	GTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGC AGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGA ACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACT	1007

	GGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGC CGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAG	

Viral Genome Component: Untranslated Regions (UTRs)

[0242] In some embodiments, wild type untranslated regions (UTRs) of a gene are transcribed but not translated. Generally, the 5' UTR starts at the transcription start site and ends at the start codon and the 3' UTR starts immediately following the stop codon and continues until the termination signal for transcription.

[0243] Features typically found in abundantly expressed genes of specific target organs (e.g., CNS tissue, muscle, or DRG) may be engineered into UTRs to enhance stability and protein production. As a non-limiting example, a 5' UTR from mRNA normally expressed in the brain (e.g., huntingtin) may be used in the viral genomes of the AAV particles described herein to enhance expression in neuronal cells or other cells of the central nervous system.

[0244] While not wishing to be bound by theory, wild-type 5′ untranslated regions (UTRs) include features which play roles in translation initiation. Kozak sequences, which are commonly known to be involved in the process by which the ribosome initiates translation of many genes, are usually included in 5' UTRs. Kozak sequences have the consensus CCR(A/G)CCAUGG, where R is a purine (adenine or guanine) three bases upstream of the start codon (ATG), which is followed by another 'G'.

[0245] In one embodiment, the 5'UTR in the viral genome includes a Kozak sequence.

[0246] In one embodiment, the 5'UTR in the viral genome does not include a Kozak sequence.

[0247] While not wishing to be bound by theory, wild-type 3′ UTRs are known to have stretches of Adenosines and Uridines embedded therein. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into three classes (Chen et al, 1995, the contents of which are herein incorporated by reference in its entirety): Class I AREs, such as, but not limited to, c-Myc and MyoD, contain several dispersed copies of an AUUUA motif within U-rich regions. Class II AREs, such as, but not limited to, GM-CSF and TNF-a, possess two or more overlapping UUAUUUA(U/A)(U/A) nonamers. Class III ARES, such as, but not limited to, c-Jun and Myogenin, are less well defined. These U rich regions do not contain an AUUUA motif. Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3′ UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message *in vivo*.

[0248] Introduction, removal or modification of 3′ UTR AU rich elements (AREs) can be used to modulate the stability of a polynucleotide. When engineering specific polynucleotides, e.g., payload regions of viral genomes, one or more copies of an ARE can be introduced to make polynucleotides

less stable and thereby curtail translation and decrease production of the resultant protein. Likewise, AREs can be identified and removed or mutated to increase the intracellular stability and thus increase translation and production of the resultant protein.

[0249] In one embodiment, the 3' UTR of the viral genome may include an oligo(dT) sequence for templated addition of a poly-A tail.

[0250] In one embodiment, the viral genome may include at least one miRNA seed, binding site or full sequence. microRNAs (or miRNA or miR) are 19-25 nucleotide noncoding RNAs that bind to the sites of nucleic acid targets and down-regulate gene expression either by reducing nucleic acid molecule stability or by inhibiting translation. In some embodiments, a microRNA sequence comprises a seed region, e.g., a sequence in the region of positions 2-8 of the mature microRNA, which has Watson-Crick sequence fully or partially complementarity to the miRNA target sequence of the nucleic acid.

[0251] In one embodiment, the viral genome may be engineered to include, alter or remove at least one miRNA binding site, full sequence or seed region.

[0252] Any UTR from any gene known in the art may be incorporated into the viral genome of the AAV particle. These UTRs, or portions thereof, may be placed in the same orientation as in the gene from which they were selected or they may be altered in orientation or location. In one embodiment, the UTR used in the viral genome of the AAV particle may be inverted, shortened, lengthened, made with one or more other 5′ UTRs or 3′ UTRs known in the art. As used herein, the term "altered" as it relates to a UTR, means that the UTR has been changed in some way in relation to a reference sequence. For example, a 3′ or 5′ UTR may be altered relative to a wild type or native UTR by the change in orientation or location as taught above or may be altered by the inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides.

[0253] In one embodiment, the viral genome of the AAV particle comprises at least one artificial UTR which is not a variant of a wild type UTR.

[0254] In one embodiment, the viral genome of the AAV particle comprises UTRs which have been selected from a family of transcripts whose proteins share a common function, structure, feature or property.

Viral Genome Component: Polyadenylation Sequence

[0255] The viral genome of the AAV particle described herein (e.g., an AAV particle comprising an AAV capsid variant, described herein) may comprise a polyadenylation sequence. In some embodiments, the viral genome of the AAV particle (e.g., an AAV particle comprising an AAV capsid variant, described herein) comprises a polyadenylation sequence between the 3' end of the nucleotide sequence encoding the payload and the 5' end of the 3'ITR.

Viral Genome Component: Introns

[0256] In some embodiments, the viral genome of the AAV particle as described herein (e.g., an AAV particle comprising an AAV capsid variant), comprises an element to enhance the payload target specificity and expression (See e.g., Powell et al. *Viral Expression Cassette Elements to Enhance Transgene Target Specificity and Expression in Gene Therapy*, Discov. Med, 2015, 19(102): 49-57; the contents of which are herein incorporated by reference in their entirety), such as an intron. Non-limiting examples of introns include, MVM (67-97 bps), F.IX truncated intron 1 (300 bps), β-globin SD/immunoglobulin heavy chain splice acceptor (250 bps), adenovirus splice donor/immunoglobin splice acceptor (500 bps), SV40 late splice donor/splice acceptor (19S/16S) (180 bps) and hybrid adenovirus splice donor/IgG splice acceptor (230 bps).

Viral Genome Component: Stuffer sequences

[0257] In some embodiments, the viral genome of an AAV particle described herein (e.g., an AAV particle comprising an AAV capsid polypeptide, e.g., an AAV capsid variant), comprises an element to improve packaging efficiency and expression, such as a stuffer or filler sequence. Non-limiting examples of stuffer sequences include albumin and/or alpha-1 antitrypsin. Any known viral, mammalian, or plant sequence may be manipulated for use as a stuffer sequence.

[0258] In one embodiment, the stuffer or filler sequence may be from about 100-3500 nucleotides in length. The stuffer sequence may have a length of about 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900 or 3000 nucleotides.

Viral Genome Component: miRNA

[0259] In one embodiment, the viral genome comprises a sequence encoding a miRNA to reduce the expression of the payload in a tissue or cell, e.g., the DRG (dorsal root ganglion), or neurons of other ganglia, such as those of the sympathetic or parasympathetic nervous system. In some embodiments, a miRNA, e.g., a miR183, a miR182, and/or miR96, may be encoded in the viral genome to modulate, e.g., reduce the expression, of the viral genome in a DRG neuron. As another non-limiting example, a miR-122 miRNA may be encoded in the viral genome to modulate, e.g., reduce, the expression of the viral genome in the liver. In some embodiments, a miRNA, e.g., a miR-142-3p, may be encoded in the viral genome to modulate, e.g., reduce, the expression, of the viral genome in a cell or tissue of the hematopoietic lineage, including for example immune cells (e.g., antigen presenting cells or APC, including dendritic cells (DCs), macrophages, and B-lymphocytes). In some embodiments, a miRNA, e.g., a miR-1, may be encoded in the viral genome to modulate, e.g., reduce, the expression, of the viral genome in a cell or tissue of the heart.

Viral Genome Component: miR Binding Site

[0260] Tissue- or cell-specific expression of the AAV viral particles disclosed herein can be enhanced by introducing tissue- or cell-specific regulatory sequences, e.g., promoters, enhancers, microRNA binding sites, e.g., a detargeting site. Without wishing to be bound by theory, it is believed that an encoded miR binding site can modulate, e.g., prevent, suppress, or otherwise inhibit, the expression of a gene of interest on the viral genome disclosed herein, based on the expression of the corresponding endogenous microRNA (miRNA) or a corresponding controlled exogenous miRNA in a tissue or cell, e.g., a non-targeting cell or tissue. In some embodiments, a miR binding site modulates, e.g., reduces, expression of the payload encoded by a viral genome of an AAV particle described herein in a cell or tissue where the corresponding mRNA is expressed.

[0261] In some embodiments, the viral genome of an AAV particle described herein comprises a nucleotide sequence encoding a microRNA binding site, e.g., a detargeting site. In some embodiments, the viral genome of an AAV particle described herein comprises a nucleotide sequence encoding a miR binding site, a microRNA binding site series (miR BSs), or a reverse complement thereof.

[0262] In some embodiments, the nucleotide sequence encoding the miR binding site series or the miR binding site is located in the 3'-UTR region of the viral genome (e.g., 3' relative to the nucleotide sequence encoding a payload), e.g., before the polyA sequence, 5'-UTR region of the viral genome (e.g., 5' relative to the nucleotide sequence encoding a payload), or both.

[0263] In some embodiments, the encoded miR binding site series comprise at least 1-5 copies, e.g., at least 1-3, 2-4, 3-5, 1, 2, 3, 4, 5 or more copies of a miR binding site (miR BS). In some embodiments, all copies are identical, e.g., comprise the same miR binding site. In some embodiments, the miR binding sites within the encoded miR binding site series are continuous and not separated by a spacer. In some embodiments, the miR binding sites within an encoded miR binding site series are separated by a spacer, e.g., a non-coding sequence. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides, nucleotides in length. In some embodiments, the spacer coding sequence or reverse complement thereof comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions relative to the nucleotide sequence of GATAGTTA.

[0264] In some embodiments, the encoded miR binding site series comprise at least 1-5 copies, e.g., at least 1-3, 2-4, 3-5, 1, 2, 3, 4, 5 or more copies of a miR binding site (miR BS). In some embodiments, at least 1, 2, 3, 4, 5, or all of the copies are different, e.g., comprise a different miR binding site. In some embodiments, the miR binding sites within the encoded miR binding site series

are continuous and not separated by a spacer. In some embodiments, the miR binding sites within an encoded miR binding site series are separated by a spacer, e.g., a non-coding sequence. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides, in length. In some embodiments, the spacer comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, but no more than four modifications, e.g., substitutions, insertions, or deletions relative to the nucleotide sequence of GATAGTTA.

[0265] In some embodiments, the encoded miR binding site is substantially identical (e.g., at least 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% identical), to the miR in the host cell. In some embodiments, the encoded miR binding site comprises at least 1, 2, 3, 4, or 5 mismatches or no more than 6, 7, 8, 9, or 10 mismatches to a miR in the host cell. In some embodiments, the mismatched nucleotides are contiguous. In some embodiments, the mismatched nucleotides are non-contiguous. In some embodiments, the mismatched nucleotides occur outside the seed region-binding sequence of the miR binding site, such as at one or both ends of the miR binding site. In some embodiments, the miR binding site is 100% identical to the miR in the host cell.

[0266] In some embodiments, the nucleotide sequence encoding the miR binding site is substantially complementary (e.g., at least 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% complementary), to the miR in the host cell. In some embodiments, to complementary sequence of the nucleotide sequence encoding the miR binding site comprises at least 1, 2, 3, 4, or 5 mismatches or no more than 6, 7, 8, 9, or 10 mismatches to a miR in the host cell. In some embodiments, the mismatched nucleotides are contiguous. In some embodiments, the mismatched nucleotides are non-contiguous. In some embodiments, the mismatched nucleotides occur outside the seed region-binding sequence of the miR binding site, such as at one or both ends of the miR binding site. In some embodiments, the encoded miR binding site is 100% complementary to the miR in the host cell.

[0267] In some embodiments, an encoded miR binding site or sequence region is at least about 10 to about 125 nucleotides in length, e.g., at least about 10 to 50 nucleotides, 10 to 100 nucleotides, 50 to 100 nucleotides, 50 to 125 nucleotides, or 100 to 125 nucleotides in length. In some embodiments, an encoded miR binding site or sequence region is at least about 7 to about 28 nucleotides in length, e.g., at least about 8-28 nucleotides, 7-28 nucleotides, 8-18 nucleotides, 12-28 nucleotides, 20-26 nucleotides, 22 nucleotides, 24 nucleotides, or 26 nucleotides in length, and optionally comprises at least one consecutive region (e.g., 7 or 8 nucleotides) complementary (e.g., fully or partially complementary) to the seed sequence of a miRNA (e.g., a miR122, a miR142, a miR183, or a miR1).

[0268] In some embodiments, the encoded miR binding site is complementary (e.g., fully or partially complementary) to a miR expressed in liver or hepatocytes, such as miR122. In some embodiments, the encoded miR binding site or encoded miR binding site series comprises a miR122

binding site sequence. In some embodiments, the encoded miR122 binding site comprises the nucleotide sequence of ACAAACACCATTGTCACACTCCA (SEQ ID NO: 4673), or a nucleotide sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., insertions, deletions, or substitutions, relative to the nucleotide sequence of SEQ ID NO: 4673, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the viral genome comprises at least 2, 3, 4, or 5 copies of the encoded miR122 binding site, e.g., an encoded miR122 binding site series, optionally wherein the encoded miR122 binding site series comprises the nucleotide sequence of: ACAAACACCATTGTCACACTCCACACAAACACCATTGTCACACTCCACACAAACACCATTGTCACACTCCA (SEQ ID NO: 4674), or a nucleotide sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4674, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, at least two of the encoded miR122 binding sites are connected directly, e.g., without a spacer. In other embodiments, at least two of the encoded miR122 binding sites are separated by a spacer, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length, which is located between two or more consecutive encoded miR122 binding site sequences. In embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8, in length. In some embodiments, the spacer coding sequence or reverse complement thereof comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, an encoded miR binding site series comprises at least 3-5 copies (e.g., 4 copies) of a miR122 binding site, with or without a spacer, wherein the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions relative to the nucleotide sequence of GATAGTTA.

[0269] In some embodiments, the encoded miR binding site is complementary (e.g., fully or partially complementary) to a miR expressed in the heart. In embodiments, the encoded miR binding site or encoded miR binding site series comprises a miR-1 binding site. In some embodiments, the encoded miR-1 binding site comprises the nucleotide sequence of ATACATACTTCTTTACATTCCA (SEQ ID NO: 4679), a nucleotide sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three,

four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4679, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the viral genome comprises at least 2, 3, 4, or 5 copies of the encoded miR-1 binding site, e.g., an encoded miR-1 binding site series. In some embodiments, the at least 2, 3, 4, or 5 copies (e.g., 2 or 3 copies) of the encoded miR-1 binding site are continuous (e.g., not separated by a spacer) or separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA.

[0270] In some embodiments, the encoded miR binding site is complementary (e.g., fully or partially complementary) to a miR expressed in hematopoietic lineage, including immune cells (e.g., antigen presenting cells or APC, including dendritic cells (DCs), macrophages, and B-lymphocytes). In some embodiments, the encoded miR binding site complementary to a miR expressed in hematopoietic lineage comprises a nucleotide sequence disclosed, e.g., in US 2018/0066279, the contents of which are incorporated by reference herein in its entirety.

[0271] In embodiments, the encoded miR binding site or encoded miR binding site series comprises a miR-142-3p binding site sequence. In some embodiments, the encoded miR-142-3p binding site comprises the nucleotide sequence of TCCATAAAGTAGGAAACACTACA (SEQ ID NO: 4675), a nucleotide sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4675, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the viral genome comprises at least 2, 3, 4, or 5 copies of the encoded miR-142-3p binding site, e.g., an encoded miR-142-3p binding site series. In some embodiments, the at least 2, 3, 4, or 5 copies (e.g., 2 or 3 copies) of the encoded miR-142-3p binding site are continuous (e.g., not separated by a spacer) or separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at

least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA.

[0272] In some embodiments, the encoded miR binding site is complementary (e.g., fully complementary or partially complementary) to a miR expressed in a DRG (dorsal root ganglion) neuron, e.g., a miR183, a miR182, and/or miR96 binding site. In some embodiments, the encoded miR binding site is complementary to a miR expressed in expressed in a DRG neuron comprises a nucleotide sequence disclosed, e.g., in WO2020/132455, the contents of which are incorporated by reference herein in its entirety.

[0273] In some embodiments, the encoded miR binding site or encoded miR binding site series comprises a miR183 binding site sequence. In some embodiments, the encoded miR183 binding site comprises the nucleotide sequence of AGTGAATTCTACCAGTGCCATA (SEQ ID NO: 4676), or a nucleotide sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4676, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the sequence complementary to the seed sequence corresponds to the double underlined of the encoded miR-183 binding site sequence. In some embodiments, the viral genome comprises at least comprises at least 2, 3, 4, or 5 copies (e.g., at least 2 or 3 copies) of the encoded miR183 binding site, e.g., an encoded miR183 binding site. In some embodiments, the at least 2, 3, 4, or 5 copies (e.g., 2 or 3 copies) of the encoded miR183 binding site are continuous (e.g., not separated by a spacer) or separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii).

[0274] In some embodiments, the encoded miR binding site or the encoded miR binding site series comprises a miR182 binding site sequence. In some embodiments, the encoded miR182 binding site comprises, the nucleotide sequence of AGTGTGAGTTCTACCATTGCCAAA (SEQ ID NO: 4677), a sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g.,

substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4677, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the viral genome comprises at least 2, 3, 4, or 5 copies of the encoded miR182 binding site, e.g., an encoded miR182 binding site series. In some embodiments, the at least 2, 3, 4, or 5 copies (e.g., 2 or 3 copies) of the encoded miR182 binding site are continuous (e.g., not separated by a spacer) or separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii).

[0275] In certain embodiments, the encoded miR binding site or the encoded miR binding site series comprises a miR96 binding site sequence. In some embodiments, the encoded miR96 binding site comprises the nucleotide sequence of AGCAAAAATGTGCTAGTGCCAAA (SEQ ID NO: 4678), a sequence having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, at least 95%, at least 99%, or 100% sequence identity, or having at least one, two, three, four, five, six, or seven modifications, e.g., substitutions, insertions, or deletions, but no more than ten modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of SEQ ID NO: 4678, e.g., wherein the modification can result in a mismatch between the encoded miR binding site and the corresponding miRNA. In some embodiments, the viral genome comprises at least 2, 3, 4, or 5 copies of the encoded miR96 binding site, e.g., an encoded miR96 binding site series. In some embodiments, the at least 2, 3, 4, or 5 copies (e.g., 2 or 3 copies) of the encoded miR96 binding site are continuous (e.g., not separated by a spacer) or separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii).

[0276] In some embodiments, the encoded miR binding site series comprises a miR122 binding site, a miR-1, a miR142 binding site, a miR183 binding site, a miR182 binding site, a miR 96 binding site, or a combination thereof. In some embodiments, the encoded miR binding site series comprises at least 2, 3, 4, or 5 copies of a miR122 binding site, a miR142 binding site, a miR183 binding site, a

miR182 binding site, a miR 96 binding site, or a combination thereof. In some embodiments, at least two of the encoded miR binding sites are connected directly, e.g., without a spacer. In other embodiments, at least two of the encoded miR binding sites are separated by a spacer, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length, which is located between two or more consecutive encoded miR binding site sequences. In embodiments, the spacer is at least about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer coding sequence or reverse complement thereof comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA.

[0277] In some embodiments, an encoded miR binding site series comprises at least 2-5 copies (e.g., 2 or 3 copies) of a combination of at least two, three, four, five, or all of a miR-1, miR122 binding site, a miR142 binding site, a miR183 binding site, a miR182 binding site, a miR96 binding site, wherein each of the miR binding sites within the series are continuous (e.g., not separated by a spacer) or are separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA.

[0278] In some embodiments, an encoded miR binding site series comprises at least 2-5 copies (e.g., 2 or 3 copies) of a combination of a miR-122 binding site and a miR-1 binding site, wherein each of the miR binding sites within the series are continuous (e.g., not separated by a spacer) or are separated by a spacer. In some embodiments, the spacer is about 1 to 6 nucleotides or about 5 to 10 nucleotides, e.g., about 7-8 nucleotides or about 8 nucleotides, in length. In some embodiments, the spacer sequence comprises one or more of (i) GGAT; (ii) CACGTG; (iii) GCATGC, or a repeat of one or more of (i)-(iii). In some embodiments, the spacer comprises the nucleotide sequence of GATAGTTA, or a nucleotide sequence having at least one, two, or three modifications, e.g., substitutions, insertions, or deletions, but no more than four modifications, e.g., substitutions, insertions, or deletions, relative to the nucleotide sequence of GATAGTTA.

Genome Size

[0279] In one embodiment, the AAV particle described herein (e.g., an AAV particle comprising an AAV capsid variant), may comprise a single-stranded or double-stranded viral genome. The size of

the viral genome may be small, medium, large or the maximum size. As described above, the viral genome may comprise a promoter and a polyA tail.

[0280] In one embodiment, the viral genome may be a small single stranded viral genome. A small single stranded viral genome may be 2.1 to 3.5 kb in size such as, but not limited to, about 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, and 3.5 kb in size.

[0281] In one embodiment, the viral genome may be a small double stranded viral genome. A small double stranded viral genome may be 1.3 to 1.7 kb in size such as, but not limited to, about 1.3, 1.4, 1.5, 1.6, and 1.7 kb in size.

[0282] In one embodiment, the viral genome may be a medium single stranded viral genome. A medium single stranded viral genome may be 3.6 to 4.3 kb in size such as, but not limited to, about 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2 and 4.3 kb in size.

[0283] In one embodiment, the viral genome may be a medium double stranded viral genome. A medium double stranded viral genome may be 1.8 to 2.1 kb in size such as, but not limited to, about 1.8, 1.9, 2.0, and 2.1 kb in size.

[0284] In one embodiment, the viral genome may be a large single stranded viral genome. A large single stranded viral genome may be 4.4 to 6.0 kb in size such as, but not limited to, about 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 and 6.0 kb in size.

[0285] In one embodiment, the viral genome may be a large double stranded viral genome. A large double stranded viral genome may be 2.2 to 3.0 kb in size such as, but not limited to, about 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 and 3.0 kb in size.

Payloads and Active Agents

[0286] In some embodiments, a ligand described herein is fused to an active agent. In some embodiments the active agent is a therapeutic agent or a diagnostic agent. In some embodiments, the ligand is a component of an AAV particle, wherein the AAV particle comprises a viral genome encoding a payload. In some embodiments, the encoded payload comprises a therapeutic agent.

[0287] In some embodiments, the encoded payload or active agent comprises a therapeutic protein, an antibody molecule, an enzyme, one or more components of a genome editing system, an Fc polypeptide fused or coupled (e.g., covalently or non covalently) to a therapeutic agent, and/or an RNAi agent (e.g., a dsRNA, siRNA, shRNA, pre-miRNA, pri-miRNA, miRNA, stRNA, lncRNA, piRNA, or snoRNA). In some embodiments, the encoded payload or active agent modulates, e.g., increases or decreases, the presence, level, and/or activity of a gene, mRNA, protein, or a combination thereof, e.g., in a cell or a tissue.

Polypeptides

[0288] In some embodiments, the encoded payload or the active agent comprises a polypeptide, protein, or peptide, e.g., a polypeptide, protein, or peptide described herein. The nucleic acid encoding

the payload, may encode a product of any known gene and/or a recombinant version thereof. The active agent can be any known protein or a recombinant version thereof. In some embodiments, the encoded payload or the active agent is an apolipoprotein E (APOE) protein such as, but not limited to ApoE2, ApoE3 and/or ApoE4 protein. In one embodiment, the encoded payload or the active agent is ApoE2 (cys112, cys158) protein or a fragment or variant thereof. In one embodiment, the encoded payload or the active agent is an ApoE3 (cys112, arg158) protein or fragment or variant thereof. In one embodiment, the encoded payload or the active agent is an ApoE4 (arg112, arg158) protein or fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises an aromatic L-amin acid decarboxylase (AADC) protein. As another non-limiting example, the encoded payload or the active agent comprises an antibody, or a fragment thereof. As another non-limiting example, the encoded payload or the active agent comprises a human survival of motor neuron (SMN) 1 or SMN2 protein, or fragments or variants thereof. As another non-limiting example, the encoded payload or the active agent comprises a glucocerebrosidase (GBA1) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises a granulin precursor or progranulin (GRN) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises an aspartoacylase (ASPA) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises a tripeptidyl peptidase I (CLN2) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises a beta-galactosidase (GLB1) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises a N-sulphoglucosamine sulphohydrolase (SGSH) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises an N-acetyl-alpha-glucosaminidase (NAGLU) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises an iduronate 2-sulfatase (IDS) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises an intracellular cholesterol transporter (NPC1) protein, or a fragment or variant thereof. As another non-limiting example, the encoded payload or the active agent comprises a gigaxonin (GAN) protein, or a fragment or variant thereof.

[0289] In some embodiments, the encoded payload or the active agent comprises an Fc polypeptide. In some embodiments, the Fc polypeptide is fused or coupled to a therapeutic agent, e.g., a therapeutic protein or enzyme.

Antibody Molecules and Antibody Binding Fragments

[0290] In some embodiments, the encoded payload or the active agent is an antibody molecule. In some embodiments, the antibody molecule binds a CNS related target, e.g., an antigen associated with a neurological or neurodegenerative disorder. In some embodiments, the antibody molecule

binds a muscular or neuromuscular related target, e.g., an antigen associated with a muscular or neuromuscular disorder. In some embodiments, the antibody molecule binds a neuro-oncology related target, e.g., an antigen associated with a neuro-oncological disorder.

[0291] In some embodiments, the antibody molecule binds to β-amyloid, APOE, tau, SOD1, TDP-43, huntingtin, and/or synuclein. In some embodiments, the encoded payload comprises an antibody or antibody fragment that binds to a neuro-oncology related target, e.g., HER2, EGFR (e.g., EGFRvIII). In some embodiments, the antibody molecule binds to HER2/neu. In some embodiments, the antibody molecule binds to β-amyloid. In some embodiments, the antibody molecule binds to tau.

[0292] In some embodiments, the active agent comprises an antibody-drug conjugate. In some embodiments, the antibody molecule is conjugated to a cytotoxic or cytostatic agent, e.g., a chemotherapeutic agent or an anti-neoplastic drug. In some embodiments, the antibody is conjugated to a radioactive isotope, e.g., α-, β-, or γ-emitter, or β-and γ-emitter.

Gene Editing System

[0293] In some embodiments, the encoded payload or the encoded active agent comprises a gene editing system or one or more components thereof. In some embodiments, the gene editing system comprises nucleic acid sequences that encode proteins having enzymatic activity to (i) selectively induce double or single stranded breaks in a DNA or RNA sequence, or (ii) substitute, insert or delete a particular base or set of bases of a DNA or RNA sequence in the absence of a double or single stranded break in the DNA or RNA. In some embodiments, the gene editing system includes, but is not limited to a CRISPR-Cas system (including different Cas or Cas-related nucleases), a Zinc finger nuclease, a meganuclease, a TALEN or a base editors. In some embodiments, the gene editing system comprises a chromosomal integration of a transgene, e.g., introduced by a parvovirus vector in the absence of an exogenous nuclease or an enzymatic entity.

RNAi agents

[0294] In some embodiments, the encoded payload or the active agent comprises an RNAi agent, e.g., an RNAi agent described herein. In some embodiments, the encoded payload or the active agent comprises a dsRNA, a siRNA, a shRNA, a pre-miRNA, a pri-miRNA, a miRNA, a stRNA, a lncRNA, a piRNA, an antisense oligonucleotide (ASO), or a snoRNA. In some embodiments, the encoded payload or the active agent comprises an RNAi agent for inhibiting expression of a SOD1, MAPT, APOE, HTT, C9ORF72, TDP-43, APP, BACE, SNCA, ATXN1, ATXN3, ATXN7, SCN1A-SCN5A, or SCN8A-SCN11A gene, protein, and/or mRNA. In some embodiments, the RNAi agent described herein inhibits SOD1, MAPT, APOE, HTT, C9ORF72, TDP-43, APP, BACE, SNCA, ATXN1, ATXN3, ATXN7, SCN1A-SCN5A, or SCN8A-SCN11A.

[0295] In some embodiment, the encoded payload or the active agent comprises an RNAi agent which targets the mRNA of a gene to modulate, e.g., interfere with gene expression and/or protein

production. In some embodiments, the RNAi agent may target a gene at the location of a single-nucleotide polymorphism (SNP) or variant within the nucleotide sequence of the gene. In some embodiments, the RNAi agent is a siRNA. In some embodiments, the RNAi agent is an ASO.

[0296] The RNAi agent may be an siRNA duplex, wherein the siRNA duplex contains an antisense strand (guide strand) and a sense strand (passenger strand) hybridized together forming a duplex structure, wherein the antisense strand is complementary to the nucleic acid sequence of the targeted gene, and wherein the sense strand is homologous to the nucleic acid sequence of the targeted gene. In some aspects, the 5'end of the antisense strand has a 5' phosphate group and the 3'end of the sense strand contains a 3'hydroxyl group. In other aspects, there are none, one or 2 nucleotide overhangs at the 3'end of each strand.

[0297] Each strand of a siRNA duplex targeting a gene of interest may be about 19 to 25, 19 to 24 or 19 to 21 nucleotides in length, preferably about 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, or 25 nucleotides in length.

[0298] In one embodiment, a siRNA or dsRNA includes at least two sequences that are complementary to each other. The dsRNA includes a sense strand having a first sequence and an antisense strand having a second sequence. The antisense strand includes a nucleotide sequence that is substantially complementary to at least part of an mRNA encoding the target gene, and the region of complementarity is 30 nucleotides or less, and at least 15 nucleotides in length. Generally, the dsRNA is 19 to 25, 19 to 24 or 19 to 21 nucleotides in length. In some embodiments, the dsRNA is from about 15 to about 25 nucleotides in length, and in other embodiments the dsRNA is from about 25 to about 30 nucleotides in length. In some embodiments, the dsRNA is about 15 nucleotides in length, 16 nucleotides in length, 17 nucleotides in length, 18 nucleotides in length, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides in length, 26 nucleotides in length, 27 nucleotides in length, 28 nucleotides in length, 29 nucleotides in length, or 30 nucleotides in length.

[0299] In some embodiments, the siRNA or the ASO is conjugated to the ligand directly. In some embodiments, the siRNA or the ASO is conjugated to the ligand via a linker, e.g., a cross-linker. In some embodiments, the crosslinker comprises succinimidyl-4-(N-maleimidomethyl) and/or a saturated or unsaturated hydrocarbon chain (e.g., cyclohexane-1-carboxylate). In some embodiments, the crosslinker comprises succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate. In some embodiments, the ligand is conjugated to the RNAi agent via a linker comprising an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate. In some embodiments, the ligand is conjugated directly or indirectly via a linker, to the N-terminus of at least one strand of the RNAi agent. In some embodiments, the ligand is conjugated, e.g., directly or indirectly via a linker, to the C-terminus of at least one strand of the RNAi agent. In some embodiments, the ligand is conjugated, e.g., directly or

indirectly via a linker, to an internal nucleotide of at least one strand of the RNAi agent. In some embodiments, the ligand is conjugated to the sense strand. In other embodiments, the ligand is conjugated to the antisense strand. In some embodiments, the ligand is conjugated to the siRNA agent, e.g., as described in WO2021207189; WO2004065601; US8034376; WO2019217459; Brown et al. Expanding RNAi therapeutics to extrahepatic tissues with lipophilic conjugates. Nature Biotechnology. 2022; Eyford et al. A Nanomule Peptide Carrier Delivers siRNA Across the Intact Blood Brain Barrier to Attenuate Ischemic Stroke. Front Mol Biosci 2021 8:611367; which are hereby incorporated by reference in their entirety.

[0300] In some embodiments, the RNAi agent, e.g., the siRNA or ASO, further comprises a lipophilic moiety. In some embodiments, the lipophilic moiety is an aliphatic, alicyclic, or polyalicyclic compound. In some embodiments, the lipophilic moiety is selected from the group consisting of lipid, cholesterol, retinoic acid, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-bis-O(hexadecyl)glycerol, geranyloxyhexyanol, hexadecylglycerol, borneol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid, O3-(oleoyl)lithocholic acid, O3-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxazine. In some embodiments, the lipophilic moiety contains a saturated or unsaturated C4-C30 hydrocarbon chain, and an optional functional group selected from the group consisting of hydroxyl, amine, carboxylic acid, sulfonate, phosphate, thiol, azide, and alkyne. In some embodiments, the lipophilic moiety contains a saturated or unsaturated C6-C18 hydrocarbon chain, e.g., a saturated or unsaturated C16 hydrocarbon chain. In some embodiments, the lipophilic moiety is conjugated via a carrier that replaces one or more nucleotide(s) in the internal position(s) of the double stranded region. In some embodiments, the carrier is a cyclic group selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl; or is an acyclic moiety based on a serinol backbone or a diethanolamine backbone. In some embodiments, the lipophilic moiety is conjugated to the RNAi agent, e.g., the siRNA or ASO, via a linker containing an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate. In some embodiments, the lipophilic moiety is conjugated to a nucleobase, sugar moiety, or internucleosidic linkage. In some embodiments, the lipophilic moiety is conjugated via a bio-cleavable linker selected from the group consisting of DNA, RNA, disulfide, amide, functionalized monosaccharides or oligosaccharides of galactosamine, glucosamine, glucose, galactose, mannose, and combinations thereof. In some embodiments, the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to the N-terminus of at least one strand of the RNAi agent. In some embodiments, the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to the C-terminus of at least one strand of the RNAi agent. In some embodiments, the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to an internal nucleotide of at

least one strand of the RNAi agent. In some embodiments, the lipophilic moiety is conjugated, e.g., directly or indirectly via a linker to the sense strand. In some embodiments, lipophilic moiety is conjugated, e.g., directly or indirectly via a linker, to the antisense strand. In some embodiments, the lipophilic moiety and the ligand are present on the same strand, e.g., the sense strand. In some embodiments, lipophilic moiety and the ligand are present on different strands. In some embodiments, the lipophilic moiety is as described in WO2021207189; WO2004065601; US8034376; WO2019217459; Brown et al. Expanding RNAi therapeutics to extrahepatic tissues with lipophilic conjugates. Nature Biotechnology. 2022 (the contents of which are hereby incorporated in their entirety).

[0301] In some embodiments, the RNAi agent, e.g., the siRNA or the ASO, further comprises an N-acetylgalactosamine (GalNAc) conjugate. In some embodiments, the GalNAc conjugate is attached through a monovalent linker; or a bivalent, trivalent, or tetravalent branched linker. In some embodiments, the GalNAc conjugate is as described in WO2013155204, which is hereby incorporated by reference in its entirety.

[0302] In some embodiments, the RNAi agent, e.g., an RNAi agent described herein inhibits the expression of the gene, mRNA, and/or protein by at least 10%, at least 20%, at least 25%, at least 30%, at least 35% or at least 40% or more, such as when assayed by a method known in the art. In some embodiments, the RNAi agent inhibits expression of a gene, mRNA, and protein by 50-100%, e.g., by 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95% and 100%.

[0303] In some embodiments, the AAV particle described herein, comprising a viral genome encoding an RNAi agent targeting a gene of interest is administered to a subject in need for treating and/or ameliorating a disease, e.g., a neurological disorder of any disease associated with the central or peripheral nervous systems.

Design of siRNA

[0304] An AAV particle described herein (e.g., an AAV particle comprising an AAV capsid variant described herein) may comprise a viral genome encoding a siRNA molecule (e.g., siRNA duplex or encoded dsRNA) that target a gene of interest and suppress target gene expression, mRNA expression, and protein production. In some aspects, the siRNA molecules are designed and used to knock out target gene variants in cells, e.g., transcripts that are identified in neurological disease. In some aspects, the siRNA molecules are designed and used to knock down target gene variants in cells.

[0305] Some guidelines for designing siRNAs (for insertion into a viral genome of the AAV particles described herein) have been proposed in the art. These guidelines generally recommend generating a 19-nucleotide duplexed region, symmetric 2-3 nucleotide 3'overhangs, 5-phosphate and 3-hydroxyl groups targeting a region in the gene to be silenced. Other rules that may govern siRNA sequence preference include, but are not limited to, (i) A/U at the 5′ end of the antisense strand; (ii) G/C at the 5′ end of the sense strand; (iii) at least five A/U residues in the 5′ terminal one-third of the

antisense strand; and (iv) the absence of any GC stretch of more than 9 nucleotides in length. In accordance with such considerations, together with the specific sequence of a target gene, highly effective siRNA molecules essential for suppressing mammalian target gene expression may be readily designed. In some embodiments, an RNAi agent described herein, e.g., a siRNA or an ASO, is chemically modified to enhance one or more properties of the RNAi agent, e.g. stability.

[0306] In one embodiment, the sense and/or antisense strand is designed based on the method and rules outlined in European Patent Publication No. EP1752536, the contents of which are herein incorporated by reference in their entirety. As a non-limiting example, the 3'-terminal base of the sequence is adenine, thymine or uracil. As a non-limiting example, the 5'-terminal base of the sequence is guanine or cytosine. As a non-limiting example, the 3'-terminal sequence comprises seven bases rich in one or more bases of adenine, thymine and uracil.

[0307] In one embodiment, a siRNA molecule comprises a sense strand and a complementary antisense strand in which both strands are hybridized together to form a duplex structure. The antisense strand has sufficient complementarity to the target mRNA sequence to direct target-specific RNAi, e.g., the siRNA molecule has a sequence sufficient to trigger the destruction of the target mRNA by the RNAi machinery or process.

[0308] In some embodiments, the antisense strand and target mRNA sequences have 100% complementarity. The antisense strand may be complementary to any part of the target mRNA sequence. Neither the identity of the sense sequence nor the homology of the antisense sequence need be 100% complementary to the target.

[0309] In other embodiments, the antisense strand and target mRNA sequences comprise at least one mismatch. As a non-limiting example, the antisense strand and the target mRNA sequence have at least 50-90%, 50-95%, 50-99%, 60-70%, 60-80%, 60-90%, 60-95%, 60-99%, 70-80%, 70-90%, 70-95%, 70-99%, 80-90%, 80-95%, 80-99%, 90-95%, 90-99% or 95-99% complementary.

[0310] The siRNA molecule may have a length from about 10-50 or more nucleotides, e.g., each strand comprising 10-50 nucleotides (or nucleotide analogs). Preferably, the siRNA molecule has a length from about 15-30, e.g., 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in each strand, wherein one of the strands is sufficiently complementary to a target region. In one embodiment, the siRNA molecule has a length from about 19 to 25, 19 to 24 or 19 to 21 nucleotides.

[0311] In some embodiments, the siRNA molecule can be a synthetic RNA duplex comprising about 19 nucleotides to about 25 nucleotides, and two overhanging nucleotides at the 3'-end.

[0312] The siRNA molecule may comprise an antisense sequence and a sense sequence, or a fragment or variant thereof. As a non-limiting example, the antisense sequence and the sense sequence have at least 50-90%, 50-95%, 50-99%, 60-70%, 60-80%, 60-90%, 60-95%, 60-99%, 70-

80%, 70-90%, 70-95%, 70-99%, 80-90%, 80-95%, 80-99%, 90-95%, 90-99% or 95-99% complementary.

[0313] The sense and antisense sequences may be completely complementary across a substantial portion of their length. In other embodiments, the sense sequence and antisense sequence may be at least 70, 80, 90, 95 or 99% complementary across independently at least 50, 60, 70, 80, 85, 90, 95, or 99% of the length of the strands.

[0314] In some embodiments, the sense and antisense strands of a siRNA duplex are linked by a short spacer sequence leading to the expression of a stem-loop structure termed short hairpin RNA (shRNA). The hairpin is recognized and cleaved by Dicer, thus generating mature siRNA molecules.

[0315] In some embodiments, the siRNA molecules, as well as associated spacer and/or flanking regions once designed, can be encoded by the viral genome of the AAV particles described herein, for delivery to a cell.

siRNA modification

[0316] In some embodiments, the RNAi agents, e.g., siRNA molecules or ASOs may be chemically modified to modulate some features of RNA molecules, such as, but not limited to, increasing the stability of siRNAs *in vivo*. The chemically modified siRNA molecules can be used in human therapeutic applications, and are improved without compromising the RNAi activity of the siRNA molecules. As a non-limiting example, the siRNA molecules modified at both the 3′ and the 5′ end of both the sense strand and the antisense strand.

[0317] In some aspects, the RNAi agent, e.g., siRNA or ASO, may contain one or more modified nucleotides such as, but not limited to, sugar modified nucleotides, nucleobase modifications and/or backbone modifications. In some aspects, the siRNA molecule may contain combined modifications, for example, combined nucleobase and backbone modifications. In some embodiments, the RNAi agent, e.g., siRNA or ASO comprises at least one modified nucleotide. In some embodiments, not more than five of the sense strand nucleotides of the siRNA and not more than five of the nucleotides of the antisense strand of the siRNA are unmodified nucleotides. In some embodiments, all of the nucleotides of the sense strand of the siRNA and all of the nucleotides of the antisense strand of the siRNA are modified. In some embodiments, not more than five of the nucleotides of ASO are unmodified nucleotides. In some embodiments, all of the nucleotides of the ASO are modified.

[0318] In one embodiment, the modified nucleotide may be a sugar-modified nucleotide. Sugar modified nucleotides include, but are not limited to 2′-fluoro, 2′-amino and 2′-thio modified ribonucleotides, e.g. 2′-fluoro modified ribonucleotides. Modified nucleotides may be modified on the sugar moiety, as well as nucleotides having sugars or analogs thereof that are not ribosyl. For example, the sugar moieties may be, or be based on, mannoses, arabinoses, glucopyranoses, galactopyranoses, 4′-thioribose, and other sugars, heterocycles, or carbocycles. In one embodiment, the modified nucleotide may be a nucleobase-modified nucleotide.

[0319] In one embodiment, the modified nucleotide may be a backbone-modified nucleotide. In some embodiments, the RNAi agent may further comprise other modifications on the backbone. In some embodiments, the phosphodiester bonds/linker (PO linkage) may be modified as "phosphorothioate backbone (PS linkage). In some cases, the natural phosphodiester bonds may be replaced by amide bonds but the four atoms between two sugar units are kept. Such amide modifications can facilitate the solid phase synthesis of oligonucleotides and increase the thermodynamic stability of a duplex formed with siRNA complement. See e.g. Mesmaeker et al., *Pure & Appl. Chem.*, 1997, 3, 437-440; the content of which is incorporated herein by reference in its entirety.

[0320] Modified bases refer to nucleotide bases such as, for example, adenine, guanine, cytosine, thymine, uracil, xanthine, inosine, and queuosine that have been modified by the replacement or addition of one or more atoms or groups. Some examples of modifications on the nucleobase moieties include, but are not limited to, alkylated, halogenated, thiolated, aminated, amidated, or acetylated bases, individually or in combination. More specific examples include, for example, 5-propynyluridine, 5-propynylcytidine, 6-methyladenine, 6-methylguanine, N,N,-dimethyladenine, 2-propyladenine, 2-propylguanine, 2-aminoadenine, 1-methylinosine, 3-methyluridine, 5-methylcytidine, 5-methyluridine and other nucleotides having a modification at the 5 position, 5-(2-amino)propyl uridine, 5-halocytidine, 5-halouridine, 4-acetylcytidine, 1-methyladenosine, 2-methyladenosine, 3-methylcytidine, 6-methyluridine, 2-methylguanosine, 7-methylguanosine, 2,2-dimethylguanosine, 5-methylaminoethyluridine, 5-methyloxyuridine, deazanucleotides such as 7-deaza-adenosine, 6-azouridine, 6-azocytidine, 6-azothymidine, 5-methyl-2-thiouridine, other thio bases such as 2-thiouridine and 4-thiouridine and 2-thiocytidine, dihydrouridine, pseudouridine, queuosine, archaeosine, naphthyl and substituted naphthyl groups, any O- and N-alkylated purines and pyrimidines such as N6-methyladenosine, 5-methylcarbonylmethyluridine, uridine 5-oxyacetic acid, pyridine-4-one, pyridine-2-one, phenyl and modified phenyl groups such as aminophenol or 2,4,6-trimethoxy benzene, modified cytosines that act as G-clamp nucleotides, 8-substituted adenines and guanines, 5-substituted uracils and thymines, azapyrimidines, carboxyhydroxyalkyl nucleotides, carboxyalkylaminoalkyl nucleotides, and alkylcarbonylalkylated nucleotides.

[0321] In some embodiments, the 3' end of the sense strand of the RNAi agent, e.g., the siRNA agent, is protected via an end cap which is a cyclic group having an amine, said cyclic group being selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl.

[0322] In some embodiments, the siRNA or ASO comprises a modification, e.g., as described in Table 1 of WO2021207189, the contents of which are hereby incorporated by reference in its entirety. In some embodiments, the siRNA or ASO comprises a modification, e.g., as described in

WO2012/037254, US9587240, US7786290, or WO2009086558, which are incorporated by reference herein. In some embodiments, the siRNA or ASO comprises a modification to increase stability, e.g., a 2'-O-methoxyethyl sugar modification.

Molecular Scaffold

[0323] In some embodiments, the siRNA molecules may be encoded in a modulatory polynucleotide which also comprises a molecular scaffold.

[0324] In some embodiments, the modulatory polynucleotide which comprises the payload (e.g., siRNA, miRNA or other RNAi agent described herein) includes a molecular scaffold which comprises a 5' flanking sequence, a loop region, and/or a 3' flanking region. In some embodiments a 5' or 3' flanking region may be of any length and may a wild type microRNA sequence or a portion thereof, or may be completely artificial. A 3' flanking sequence may mirror the 5' flanking sequence in size and origin. Either flanking sequence may be absent. In one embodiment, both the 5' and 3' flanking sequences are absent. The 3' flanking sequence may optionally contain one or more CNNC motifs, where "N" represents any nucleotide. In some embodiments, the loop comprises at least one UGUG motif. In some embodiments, the UGUG motif is located at the 5' terminus of the loop. In some embodiments the 5' and 3' flanking sequences are the same sequence. In some embodiments they differ by 2%, 3%, 4%, 5%, 10%, 20% or more than 30% when aligned to each other.

[0325] In some embodiments, modulatory polynucleotide comprises a stem loop structure. In some embodiments, the modulatory polynucleotide comprises in 5' to 3' order: a 5' flanking sequence, a guide strand sequence, a loop region, a passenger strand sequence, and a 3' flanking sequence. In some embodiments, the modulatory polynucleotide comprises in 5' to 3' order: a 5' flanking sequence, a passenger strand sequence, a loop region, a guide strand sequence, and a 3' flanking sequence.

[0326] In one embodiment, the molecular scaffold comprises a dual-function targeting modulatory polynucleotide.

[0327] In one embodiment, the molecular scaffold may comprise one or more linkers known in the art. The linkers may separate regions or one molecular scaffold from another. As a non-limiting example, the molecular scaffold may be polycistronic.

[0328] In one embodiment, the modulatory polynucleotide is designed using at least one of the following properties: loop variant, seed mismatch/bulge/wobble variant, stem mismatch, loop variant and basal stem mismatch variant, seed mismatch and basal stem mismatch variant, stem mismatch and basal stem mismatch variant, seed wobble and basal stem wobble variant, or a stem sequence variant.

Other Active Agents

[0329] In some embodiments the active agent comprises a diagnostic agent. In some embodiments, the diagnostic agent is or comprises an imaging agent (e.g., a protein or small molecule

compound coupled to a detectable moiety). In some embodiments, the imaging agent comprises a PET or MRI ligand, or an antibody molecule coupled to a detectable moiety. In some embodiments, the detectable moiety is or comprises a radiolabel, a fluorophore, a chromophore, or an affinity tag. In some embodiments, the radiolabel is or comprises tc99m, iodine-123, a spin label, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen- 15, oxygen-17, gadolinium, manganese, or iron.

[0330] In some embodiments, the active agent is a small molecule. In some embodiments, the active agent is a ribonucleic acid complex (e.g., a Cas9/gRNA complex), a plasmid, a closed-end DNA, a circ-RNA, or an mRNA.

Therapeutic Applications

[0331] The present disclosure provides a method for treating a disease, disorder and/or condition in a subject, including a human subject, comprising administering to the subject a composition described herein, e.g., a composition comprising a ligand that binds to a GPI anchor protein fused or coupled (e.g., covalently or non-covalently) to an active agent (e.g., a therapeutic agent or a diagnostic agent).

[0332] In some embodiments, a composition described herein is administered to a subject prophylactically, to prevent on-set of disease. In another embodiment, the composition is administered to treat (e.g., lessen the effects of) a disease or symptoms thereof. In yet another embodiment, the composition is administered to cure (eliminate) a disease. In another embodiment, the composition is administered to prevent or slow progression of disease. In yet another embodiment, the composition is used to reverse the deleterious effects of a disease. Disease status and/or progression may be determined or monitored by standard methods known in the art.

[0333] In some embodiments, a composition described herein is useful for treatment, prophylaxis, palliation or amelioration of a genetic disorder, e.g., an autosomal dominant genetic disorder, an autosomal recessive disorder, X-linked dominant genetic disorder, an X-linked recessive genetic disorder, or a Y-linked genetic disorder. In some embodiments, the genetic disorder is a monogenetic disorder or a polygenic disorder. In some embodiments, treatment of a genetic disorder, e.g., a monogenic disorder, comprises the use of a composition described herein for a gene replacement therapy.

[0334] In some embodiments, provided herein is method for treating a neurological disorder and/or neurodegenerative disorder in a subject, comprising administering to the subject an effective amount of a composition described herein. In some embodiments, treatment of a neurological disorder and/or neurodegenerative disorder comprises prevention of said neurological disorder and/or neurological disorder.

[0335] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of neurological diseases and/or disorders. In some

embodiments, the composition is useful for the treatment, prophylaxis, palliation or amelioration of tauopathy.

[0336] In some embodiments, a composition described herein is for the treatment, prophylaxis, palliation or amelioration of Alzheimer's Disease. In some embodiments, treatment of Alzheimer's Disease comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ApoE2 protein, ApoE4 protein, an ApoE3 protein, BDNF protein, CYP46A1 protein, Klotho protein, fractalkine (FKN) protein, neprilysin protein (NEP), CD74 protein, caveolin-1, or a combination or variant thereof. In some embodiments, treatment of Alzheimer's Disease comprises the use of the composition for a reduction in the expression of a tau gene and/or protein, a synuclein gene and/or protein, or a combination or variant thereof. In some embodiments, the encoded payload or active agent comprises an antibody molecule that binds to tau or synuclein, an RNAi agent for inhibiting tau or synuclein, a gene editing system (e.g., a CRISPR-Cas system) for altering tau or synuclein expression, or a combination thereof.

[0337] In some embodiments, a composition described herein is useful the treatment, prophylaxis, palliation or amelioration of Friedreich's ataxia, or any disease stemming from a loss or partial loss of frataxin protein.

[0338] In some embodiments, a composition described herein is for the treatment, prophylaxis, palliation or amelioration of frontal temporal dementia. In some embodiments, treatment of frontal temporal dementia comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a progranulin protein or variant thereof.

[0339] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of Parkinson's Disease. In some embodiments, treatment of Parkinson's disease comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an AADC protein, GAD protein, GDNF protein, TH-GCH1 protein, GBA protein, AIMP2-DX2 protein, or a combination or variant thereof. In some embodiments, treatment of Parkinson's disease comprises the composition for a gene knock-down therapy or a gene editing therapy (e.g., knock-out, repression, or correction). In some embodiments, the encoded payload or active agent comprises a modulator, e.g., an RNAi agent or a CRISPR-Cas system, for altering expression of an alpha-synuclein gene, mRNA, and/or protein, or variant thereof. In some embodiments, the composition is useful for the treatment, prophylaxis, palliation or amelioration of an AADC deficiency. In some embodiments, treatment of AADC deficiency comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an AADC protein or variant thereof.

[0340] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of Amyotrophic lateral sclerosis. In some embodiments, treatment of ALS comprises the use of the composition for a gene replacement therapy. In some

embodiments, the encoded payload or the active agent comprises a TDP-43 protein, UPF1 protein, C9orf72 protein, CCNF protein, HSF1 protein, Factor H protein, NGF protein, ADAR2 protein, GDNF protein, VEGF protein, HGF protein, NRTN protein, AIMP2-DX2 protein, or a combination or variant thereof. In some embodiments, treatment of ALS comprises the use of the composition for a gene knock-down therapy or a gene editing therapy (e.g., knock-out, repression, or correction). In some embodiments, the encoded payload or the active agent comprises a modulator, e.g., an RNAi agent or a CRISPR-Cas system, for altering expression of a SOD1 or C9ORF72 gene, mRNA, and/or protein, or a combination or variant thereof.

[0341] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of Huntington's Disease. In some embodiments, treatment of ALS comprises the use of the composition for a gene knock-down (e.g., knock-out) therapy or a gene editing therapy (e.g., knock-out, repression, or correction). In some embodiments, the encoded payload or active agent comprises a modulator, e.g., an RNAi agent or a CRISPR-Cas system, for altering expression of an HTT gene, mRNA, and/or protein, or a variant thereof.

[0342] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of spinal muscular atrophy. In some embodiments, treatment of spinal muscular atrophy comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an SMN1 protein, an SMN2 protein, or a combination or variant thereof.

[0343] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of multiple system atrophy. In some embodiments, treatment of multiple system atrophy comprises the use of the composition for a gene replacement therapy.

[0344] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of Gaucher disease (GD) (e.g., Type 1 GD, Type 2 GD, or Type 3 GD). In some embodiments, the composition is useful for the treatment, prophylaxis, palliation or amelioration of Parkinson's disease associated with a GBA mutation. In some embodiments, the composition is useful for the treatment, prophylaxis, palliation or amelioration of dementia with Lewy Bodies (DLB).

[0345] In some embodiments, the a composition described herein is useful for treatment, prophylaxis, palliation or amelioration of a leukodystrophy, e.g., Alexander disease, autosomal dominant leukodystrophy with autonomic diseases (ADLD), Canavan disease, cerebrotendinous xanthomatosis (CTX), metachromatic leukodystrophy (MLD), Pelizaeus-Merzbacher disease, or Refsum disease. In some embodiments, treatment of MLD comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ARSA protein or variant thereof. In some embodiments, treatment of ALD comprises the use of the

composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ABCD-1 protein or variant thereof.

[0346] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of megalencephalic leukoencephalopathy (MLC). In some embodiments, treatment of MLC comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload encoded or active agent comprises an MLC1 protein or variant thereof.

[0347] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Krabbe disease. In some embodiments, treatment of Krabbe disease comprises the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a GALC protein or variant thereof.

[0348] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Mucopolysaccharidosis, e.g., a Type I (MPS I), Type II (MPS II), Type IIIA (MPS IIIA), Type IIIB (MPS IIIB), or Type IIIC (MPS IIIC). In some embodiments, treatment of Mucopolysaccharidosis comprises the use of the composition for a gene replacement therapy or a gene editing therapy (e.g., enhancement or correction). In some embodiments, the encoded payload or active agent comprises an IDUA protein, IDS protein, SGSH protein, NAGLU protein, HGSNAT protein, or a combination or variant thereof.

[0349] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Batten/NCL. In some embodiments, treatment of Batten/NCL comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a CLN1 protein, CLN2 protein, CLN3 protein, CLN5 protein, CLN6 protein, CLN7 protein, CLN8 protein, or a combination or variant thereof.

[0350] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of Rett Syndrome. In some embodiments, treatment of Rett Syndrome comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload comprises a capsid variant described herein comprises an MeCP2 protein or variant thereof.

[0351] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Angelman Syndrome. In some embodiments, treatment of Angelman Syndrome comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a UBE3A protein or variant thereof.

[0352] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Fragile X Syndrome. In some embodiments, treatment of Fragile X Syndrome comprises the use of the composition for a gene replacement therapy. In some

embodiments, the encoded payload or active agent comprises a Reelin protein, a DgkK protein, a FMR1 protein, or a combination or variant thereof.

[0353] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Canavan Disease. In some embodiments, treatment of Canavan Disease comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ASPA protein or variant thereof.

[0354] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Gangliosidosis, e.g., a GM1 Gangliosidosis or a GM2 Gangliosidosis (e.g., Tay Sachs Sandhoff). In some embodiments, treatment of a Gangliosidosis, e.g., a GM1 Gangliosidosis or a GM2 Gangliosidosis (e.g., Tay Sachs Sandhoff), comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprising a capsid variant described herein comprises a GLB1 protein, a HEXA protein, a HEXB protein, a GM2A protein, or a combination or variant thereof.

[0355] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of GM3 Synthase Deficiency. In some embodiments, treatment of GM3 Synthase Deficiency comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ST3GAL5 protein or variant thereof.

[0356] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Niemann-Pick disorder, e.g., a Niemann-Pick A or a Niemann-Pick C1 (NPC-1). In some embodiments, treatment of a Niemann-Pick disorder, e.g., a Niemann-Pick A or a Niemann-Pick C1 (NPC-1) comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an ASM protein, an NPC1 protein, or variant thereof.

[0357] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Schwannoma (e.g., Neuroma). In some embodiments, treatment of Schwannoma (e.g., Neuroma) comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a Caspase-1 protein or variant thereof.

[0358] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Tuberous Sclerosis, e.g., Tuberous Sclerosis Type 1 or Tuberous Sclerosis Type 2. In some embodiments, treatment of Tuberous Sclerosis, e.g., Tuberous Sclerosis Type 1 or Tuberous Sclerosis Type 2 comprises the use the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a TSC1 protein, a TSC2 protein, or variant thereof.

[0359] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a CDKL5 Deficiency. In some embodiments, treatment of a CDKL5 Deficiency comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a CDKL5 protein or variant thereof.

[0360] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Charcot-Marie-Tooth disorder, e.g., a Charcot-Marie-Tooth Type 1X (CMT1X) disorder, a Charcot-Marie-Tooth Type 2A (CMT2A) disorder, or a Charcot-Marie-Tooth Type 4J (CMT4J) disorder. In some embodiments, treatment of a Charcot-Marie-Tooth disorder, e.g., a Charcot-Marie-Tooth Type 1X (CMT1X) disorder, a Charcot-Marie-Tooth Type 2A (CMT2A) disorder, or a Charcot-Marie-Tooth Type 4J (CMT4J) disorder, comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a GJB1 protein, a MFN2 protein, a FIG4 protein, or variant thereof.

[0361] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of an Aspartylglucosaminuria (AGU). In some embodiments, treatment of an AGU comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an AGA protein or variant thereof.

[0362] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Leigh Syndrome. In some embodiments, treatment of a Leigh Syndrome comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a SURF1 protein or variant thereof.

[0363] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of epilepsy. In some embodiments, treatment of epilepsy comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an NPY/Y2 protein, a Galanin protein, a Dynorphin protein, an AIMP2-DX2 protein, an SLC6A1 protein, an SLC13A5 protein, a KCNQ2 protein, or variant thereof.

[0364] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Dravet Syndrome. In some embodiments, treatment of Dravet Syndrome comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises an SCN1a protein, or variant thereof.

[0365] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of a Duchenne muscular dystrophy (DMD). In some embodiments, treatment of DMD comprises the use of the composition for a gene replacement therapy or enhancement (e.g., correction of exon-skipping), or a gene editing therapy (e.g., enhancement or correction). In some embodiments, the encoded payload or active agent comprises a

Dystrophin gene and/or protein, an Utrophin gene and/or protein, or a GALGT2 gene and/or protein, or a Follistatin gene and/or protein, or a combination or variant thereof.

[0366] In some embodiments, a compositions described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Pompe Disease. In some embodiments, treatment of Pompe Disease comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a GAA protein, or variant thereof.

[0367] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation, or amelioration of Limb-Girdle Muscular Dystrophy (LGMD2A). In some embodiments, treatment of LGMD2A comprises the use of the composition for a gene replacement therapy. In some embodiments, the encoded payload or active agent comprises a CAPN-3 protein, DYSF protein, a SGCG protein, a SGCA protein, a SGCB protein, a FKRP protein, a ANO5 protein, or a combination or variant thereof.

[0368] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of chronic or neuropathic pain.

[0369] In some embodiments, a composition described herein is useful for treatment, prophylaxis, palliation or amelioration of a disease associated with the central nervous system.

[0370] In some embodiments, a composition described herein is useful for treatment, prophylaxis, palliation or amelioration of a disease associated with the peripheral nervous system.

[0371] In some embodiments, provided herein is a method for treating a neuro-oncological disorder in a subject, comprising administering to the subject an effective amount of a composition described herein. In some embodiments, treatment of a neuro-oncological disorder comprises prevention of said neuro-oncological disorder. In some embodiments, a neuro-oncological disorder comprises a cancer of a primary CNS origin (e.g., a CNS cell, a tissue, or a region), or a metastatic cancer in a CNS cell, tissue, or region. Examples of primary CNS cancers could be gliomas (which may include glioblastoma (also known as glioblastoma multiforme), astrocytomas, oligodendrogliomas, and ependymomas, and mixed gliomas), meningiomas, medulloblastomas, neuromas, and primary CNS lymphoma (in the brain, spinal cord, or meninges), among others. Examples of metastatic cancers include those originating in another tissue or organ, e.g., breast, lung, lymphoma, leukemia, melanoma (skin cancer), colon, kidney, prostate, or other types that metastasize to brain.

[0372] In some embodiments, a composition described herein is useful for the treatment, prophylaxis, palliation or amelioration of a disease associated with expression of HER2, e.g., a disease associated with overexpression of HER2. In some embodiments, the composition is useful for the treatment, prophylaxis, palliation or amelioration of a HER2-positive cancer. In some embodiments, the HER2-positive cancer is a HER2-positive solid tumor. Additionally, or alternatively, the HER2-positive cancer may be a locally advanced or metastatic HER2-positive

cancer. In some instances, the HER2-positive cancer is a HER2-positive breast cancer or a HER2-positive gastric cancer. In some embodiments, the HER2-positive cancer is selected from the group consisting of a HER2- positive gastroesophageal junction cancer, a HER2-positive colorectal cancer, a HER2-positive lung cancer (e.g., a HER2-positive non-small cell lung carcinoma), a HER2-positive pancreatic cancer, a HER2-positive colorectal cancer, a HER2-positive bladder cancer, a HER2-positive salivary duct cancer, a HER2-positive ovarian cancer (e.g., a HER2-positive epithelial ovarian cancer), or a HER2-positive endometrial cancer. In some instances, the HER2-positive cancer is prostate cancer. In some embodiments, the HER2-positive cancer has metastasized to the central nervous system (CNS). In some instances, the metastasized HER2-cancer has formed CNS neoplasms.

[0373] In some embodiments, a composition described herein is administered to a subject having at least one of the diseases or symptoms described herein. In some embodiments, the composition is administered to a subject having or diagnosed with having a disease or disorder described herein.

[0374] In some embodiments, provided herein is a method for treating a muscular disorder and/or neuromuscular disorder in a subject, comprising administering to the subject an effective amount of a composition described herein. In some embodiments, treatment of a muscular disorder and/or neuromuscular disorder comprises prevention of said muscular disorder and/or neuromuscular disorder.

[0375] In some embodiments, a composition described herein is administered to a subject having at least one of the diseases or symptoms described herein. In some embodiments, the composition is administered to a subject having or diagnosed with having a disease or disorder described herein.

[0376] Any neurological disease or disorder, neurodegenerative disorder, muscular disorder, neuromuscular disorder, and/or neuro-oncological disorder may be treated with compositions described herein, or pharmaceutical compositions thereof.

Pharmaceutical Composition and Formulations

[0377] According to the present disclosure, an AAV particle comprising an AAV capsid variant described herein may be prepared as a pharmaceutical composition. In some embodiments a composition described herein, e.g., a composition comprising a ligand that binds to a GPI anchor protein fused or coupled (e.g., covalently or non-covalently) to an active agent (e.g., a therapeutic agent or a diagnostic agent) can be prepared as pharmaceutical composition. In some embodiments, the pharmaceutical composition comprises at least one active ingredients. In some embodiments, the pharmaceutical composition comprises a pharmaceutically acceptable excipient.

[0378] In some embodiments, an AAV particle or composition described herein can be formulated using an excipient to: (1) increase stability; (2) increase cell transfection or transduction; (3) permit the sustained or delayed expression of the payload; (4) alter the biodistribution (e.g., target the viral particle to specific tissues or cell types); (5) increase the translation of encoded protein; (6)

alter the release profile of encoded protein; and/or (7) allow for regulatable expression of the payload. Formulations of the present disclosure can include, without limitation, saline, liposomes, lipid nanoparticles, polymers, peptides, proteins, cells transfected with viral vectors (e.g., for transfer or transplantation into a subject) and combinations thereof.

[0379] In some embodiments, the relative amount of the active ingredient (e.g. an AAV particle or a composition described herein), a pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the present disclosure may vary, depending upon the identity, size, and/or condition of the subject being treated and further depending upon the route by which the composition is to be administered. For example, the composition may comprise between 0.1% and 99% (w/w) of the active ingredient. By way of example, the composition may comprise between 0.1% and 100%, e.g., between .5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

[0380] The present disclosure also provides in some embodiments, a pharmaceutical composition suitable for administration to a subject, e.g., a human. In some embodiments, the pharmaceutical composition is administered to a subject, e.g., a human.

Administration

[0381] In some embodiments, a composition described herein may be administered to a subject by a delivery route, e.g., a localized delivery route or a systemic delivery route.

[0382] In some embodiments, a composition described herein may be administered via such a route that it is able to cross the blood-brain barrier, vascular barrier, or other epithelial barrier. In some embodiments, a composition described herein may be administered in any suitable form, either as a liquid solution or suspension, as a solid form suitable for liquid solution or suspension in a liquid solution. In some embodiments, a composition described herein may be formulated with any appropriate and pharmaceutically acceptable excipient.

[0383] In some embodiments, a composition described herein is administered intramuscularly, intravenously, intracerebrally, intrathecally, intratumorally, intracerebroventricularly, via intraparenchymal administration, or via intra-cisterna magna injection (ICM). In some embodiments, the composition is administered intravenously. In some embodiments, the composition is administered via intra-cisterna magna injection (ICM). In some embodiments, the composition is administered intratumorally. In some embodiments, the composition is administered intraarterially.

[0384] In some embodiments, a composition described herein may be delivered to a subject via a single route administration. In some embodiments, the composition may be delivered to a subject via a multi-site route of administration. In some embodiments, a subject may be administered at 2, 3, 4, 5, or more than 5 sites.

[0385] In some embodiments, a composition described herein is administered via a bolus infusion. In some embodiments, the composition is administered via sustained delivery over a period

of minutes, hours, or days. In some embodiments, the infusion rate may be changed depending on the subject, distribution, formulation, and/or another delivery parameter. In some embodiments, the composition is administered using a controlled release. In some embodiments, the composition is administered using a sustained release, e.g., a release profile that conforms to a release rate over a specific period of time.

[0386] In some embodiments, a composition described herein may be delivered by more than one route of administration. As non-limiting examples of combination administrations, the composition may be delivered by intrathecal and intracerebroventricular, or by intravenous and intraparenchymal administration.

Intravenous administration

[0387] In some embodiments, a composition described herein may be administered to a subject by systemic administration. In some embodiments, the systemic administration is intravenous administration. In another embodiment, the systemic administration is intraarterial administration. In some embodiments, the composition is administered to a subject by intravenous administration. In some embodiments, the intravenous administration may be achieved by subcutaneous delivery. In some embodiments, the composition is administered to the subject via focused ultrasound (FUS), e.g., coupled with the intravenous administration of microbubbles (FUS-MB) or MRI-guided FUS coupled with intravenous administration, e.g., as described in Terstappen et al. (Nat Rev Drug Discovery, doi.org/10.1038/s41573-021-00139-y (2021)), the contents of which are incorporated herein by reference in its entirety. In some embodiments, the composition is administered to the subject intravenously. In some embodiments, the subject is a human.

Administration to the CNS

[0388] In some embodiments, a composition described herein may be delivered by direct injection into the brain. As a non-limiting example, the brain delivery may be by intrahippocampal administration. In some embodiments, the composition is administered to a subject by intraparenchymal administration. In some embodiments, the intraparenchymal administration is to tissue of the central nervous system. In some embodiments, the composition is administered to a subject by intracranial delivery (See, e.g., US Pat. No. 8119611; the content of which is incorporated herein by reference in its entirety). In some embodiments, the composition is delivered by injection into the CSF pathway. Non-limiting examples of delivery to the CSF pathway include intrathecal and intracerebroventricular administration. In some embodiments, the composition is administered via intracisternal magna (ICM) injection.

[0389] In some embodiments, a composition described herein is delivered to the brain by systemic delivery. As a non-limiting example, the systemic delivery may be by intravascular

administration. As a non-limiting example, the systemic or intravascular administration may be intravenous.

[0390] In some embodiments, a composition described herein is delivered by an intraocular delivery route. A non-limiting example of an intraocular administration includes an intravitreal injection.

Intramuscular administration

[0391] In some embodiments, a composition described herein is delivered by intramuscular administration. Without wishing to be bound by theory, it is believed in some embodiments, that the multi-nucleated nature of muscle cells provides an advantage to gene transduction subsequent to delivery. In some embodiments, cells of the muscle are capable of expressing recombinant proteins with the appropriate post-translational modifications. Without wishing to be bound by theory, it is believed in some embodiments, the enrichment of muscle tissue with vascular structures allows for transfer to the blood stream and whole-body delivery. Examples of intramuscular administration include systemic (e.g., intravenous), subcutaneous or directly into the muscle. In some embodiments, more than one injection is administered. In some embodiments, an AAV particle of the present disclosure may be delivered by an intramuscular delivery route. (See, e.g., U. S. Pat. No. 6506379; the content of which is incorporated herein by reference in its entirety). Non-limiting examples of intramuscular administration include an intravenous injection or a subcutaneous injection.

[0392] In some embodiments, a composition described herein is administered to a subject and transduces the muscle of a subject. As a non-limiting example, the composition is administered by intramuscular administration. In some embodiments, the composition is administered to a subject by subcutaneous administration. In some embodiments, the intramuscular administration is via systemic delivery. In some embodiments, the intramuscular administration is via intravenous delivery. In some embodiments, the intramuscular administration is via direct injection to the muscle.

[0393] In some embodiments, the muscle is transduced by administration, e.g., intramuscular administration. In some embodiments, an intramuscular delivery comprises administration at one site. In some embodiments, an intramuscular delivery comprises administration at more than one site. In some embodiments, an intramuscular delivery comprises administration at two, three, four, or more sites. In some embodiments, intramuscular delivery is combined with at least one other method of administration.

[0394] In some embodiments, a composition described herein is administered to a subject by peripheral injections. Non-limiting examples of peripheral injections include intraperitoneal, intramuscular, intravenous, conjunctival, or joint injection. It was disclosed in the art that the peripheral administration of AAV vectors can be transported to the central nervous system, for example, to the motor neurons (e.g., U. S. Patent Publication Nos. US20100240739 and US20100130594; the content of each of which is incorporated herein by reference in their entirety).

[0395] In some embodiments, a composition described herein is administered to a subject by intraparenchymal administration. In some embodiments, the intraparenchymal administration is to muscle tissue. In some embodiments, an AAV particle or composition described herein is delivered as described in Bright et al 2015 (Neurobiol Aging. 36(2):693-709), the contents of which are herein incorporated by reference in their entirety. In some embodiments, the composition is administered to the gastrocnemius muscle of a subject. In some embodiments, the composition is administered to the bicep femorii of the subject. In some embodiments, the composition is administered to the tibialis anterior muscles. In some embodiments, the composition is administered to the soleus muscle.

Depot administration

[0396] In some embodiments, a composition as described herein is formulated in depots for extended release. Generally, specific organs or tissues are targeted for administration.

[0397] In some embodiments, a composition described herein is spatially retained within or proximal to target tissues. Provided are methods of providing a composition described herein to target tissues of mammalian subjects by contacting target tissues (which comprise one or more target cells) with the composition, under conditions such that they are substantially retained in target tissues, e.g., such that at least 10, 20, 30, 40, 50, 60, 70, 80, 85, 90, 95, 96, 97, 98, 99, 99.9, 99.99 or greater than 99.99% of the composition is retained in the target tissues. In some embodiments, retention is determined by measuring the amount of the composition that enter a target cell or a plurality of target cells. For example, at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.9%, 99.99%, or greater than 99.99% of a pharmaceutical composition and/or an AAV particle, administered to a subject are present intracellularly at a period of time following administration. For example, intramuscular injection to a subject may be performed using aqueous compositions comprising a composition described herein and a transfection reagent, and retention is determined by measuring the amount of the composition, present in the muscle cell or plurality of muscle cells.

[0398] In some embodiments, disclosed herein are methods of providing a composition described herein to a tissue of a subject, by contacting the tissue (comprising a cell, e.g., a plurality of cells) with the composition under conditions such that they are substantially retained in the tissue. In some embodiments, a composition described herein comprises a sufficient amount of an active ingredient such that the effect of interest is produced in at least one cell. In some embodiments, a composition described herein generally comprises one or more cell penetration agents. In some embodiments, the disclosure provides a naked formulation (such as without cell penetration agents or other agents), with or without pharmaceutically acceptable carriers.

Methods of Treatment

[0399] Provided in the present disclosure are methods for introducing (e.g., delivering) a composition described herein into cells. In some embodiments, the method comprises introducing

into said cells an AAV particle or vector described herein in an amount sufficient to modulate, e.g., increase, the production of a target gene, mRNA, and/or protein. In some embodiments, the method comprises introducing into said cells a composition or vector described herein in an amount sufficient to modulate, e.g., decrease, expression of a target gene, mRNA, and/or protein. In some aspects, the cells may be neurons such as but not limited to, motor, hippocampal, entorhinal, thalamic, cortical, sensory, sympathetic, or parasympathetic neurons, and glial cells such as astrocytes, microglia, and/or oligodendrocytes.

[0400] Disclosed in the present disclosure are methods for treating a neurological disease/disorder or a neurodegenerative disorder, a muscular or neuromuscular disorder, or a neuro-oncological disorder associated with aberrant, e.g., insufficient or increased, function/presence of a protein, e.g., a target protein in a subject in need of treatment.

[0401] In some embodiments, the method comprises administering to the subject a therapeutically effective amount of a composition described herein.

[0402] In some embodiments, the composition comprising the AAV particles of the present disclosure (e.g., an AAV particle comprising an AAV capsid variant described herein) is administered to the central nervous system of the subject via systemic administration. In some embodiments, the systemic administration is intravenous (IV) injection. In some embodiments, the AAV particle described herein or a pharmaceutical composition comprising an AAV particle described herein is administered by focused ultrasound (FUS), e.g., coupled with the intravenous administration of microbubbles (FUS-MB) or MRI-guided FUS coupled with intravenous administration.

[0403] In some embodiments, a composition described herein is administered to the central nervous system of the subject via intraventricular administration. In some embodiments, the composition comprising the AAV particle of the present disclosure (e.g., an AAV particle comprising an AAV capsid variant) is administered via intra-cisterna magna injection (ICM).

[0404] In some embodiments, a composition described herein is administered to the central nervous system of the subject via intraventricular injection and intravenous injection.

[0405] In some embodiments, a composition described herein is administered to the central nervous system of the subject via ICM injection and intravenous injection at a specific dose per subject. As a non-limiting example, the AAV particles are administered via ICM injection at a dose of $1x10^4$ VG per subject. As a non-limiting example, the AAV particles are administered via IV injection at a dose of $2x10^{13}$ VG per subject.

[0406] In some embodiments, a composition described herein is administered to the central nervous system of the subject. In other embodiments, the composition comprising the AAV particles of the present disclosure is administered to a CNS tissue of a subject (e.g., putamen, hippocampus, thalamus, or cortex of the subject).

[0407] In some embodiments, a composition described herein is administered to the central nervous system of the subject via intraparenchymal injection. Non-limiting examples of intraparenchymal injections include intraputamenal, intracortical, intrathalamic, intrastriatal, intrahippocampal or into the entorhinal cortex.

[0408] In some embodiments, a composition described herein is administered to the central nervous system of the subject via intraparenchymal injection and intravenous injection.

[0409] In some embodiments, a composition described herein is administered to the central nervous system of the subject via intraventricular injection, intraparenchymal injection and intravenous injection.

[0410] In some embodiments, a composition described herein is administered to a muscle of the subject via intravenous injection.

[0411] In some embodiments, a composition described herein is delivered into specific types of cells, including, but not limited to, thalamic, hippocampal, entorhinal, cortical, motor, sensory, excitatory, inhibitory, sympathetic, or parasympathetic neurons; glial cells including oligodendrocytes, astrocytes and microglia; and/or other cells surrounding neurons such as T cells. In some embodiments, a composition described herein is delivered to a cell or region of the midbrain. In some embodiments, a composition described herein is delivered to a cell or region of the brains stem. In some embodiments, a composition described herein is delivered to neurons in the putamen, hippocampus, thalamus and/or cortex.

[0412] In some embodiments, administration of a composition described herein to a subject may increase target gene, mRNA, and/or protein levels in a subject, relative to a control, e.g., the gene, mRNA, and/or mRNA levels in the subject prior to the composition. The target gene, mRNA, and/or protein levels may be increased by about 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95% and 100%, or at least 20-30%, 20-40%, 20-50%, 20-60%, 20-70%, 20-80%, 20-90%, 20-95%, 20-100%, 30-40%, 30-50%, 30-60%, 30-70%, 30-80%, 30-90%, 30-95%, 30-100%, 40-50%, 40-60%, 40-70%, 40-80%, 40-90%, 40-95%, 40-100%, 50-60%, 50-70%, 50-80%, 50-90%, 50-95%, 50-100%, 60-70%, 60-80%, 60-90%, 60-95%, 60-100%, 70-80%, 70-90%, 70-95%, 70-100%, 80-90%, 80-95%, 80-100%, 90-95%, 90-100% or 95-100% in a subject such as, but not limited to, the CNS, a region of the CNS, or a specific cell of the CNS, or a muscle, a region of a muscle, or a cell of a muscle, of a subject. In some embodiments, cell of the CNS comprises an astrocyte, microglia, cortical neuron, hippocampal neuron, DRG and/or sympathetic neuron, sensory neuron, oligodendrocyte, motor neuron, or combination thereof. As a non-limiting example, the composition may increase the gene, mRNA, and/or protein levels of a target protein by fold increases over baseline. In some embodiments, the composition leads to 5-6 times higher levels of a target gene, mRNA, or protein.

[0413] In some embodiments, administration of a composition described herein e.g., a comprising a siRNA molecule, to a subject may decrease target gene, mRNA, and/or protein levels in a subject,

relative to a control, e.g., the gene, mRNA, and/or mRNA levels in the subject prior to receiving the composition. The target gene, mRNA, and/or protein levels may be decreased by about 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95% and 100%, or at least 20-30%, 20-40%, 20-50%, 20-60%, 20-70%, 20-80%, 20-90%, 20-95%, 20-100%, 30-40%, 30-50%, 30-60%, 30-70%, 30-80%, 30-90%, 30-95%, 30-100%, 40-50%, 40-60%, 40-70%, 40-80%, 40-90%, 40-95%, 40-100%, 50-60%, 50-70%, 50-80%, 50-90%, 50-95%, 50-100%, 60-70%, 60-80%, 60-90%, 60-95%, 60-100%, 70-80%, 70-90%, 70-95%, 70-100%, 80-90%, 80-95%, 80-100%, 90-95%, 90-100% or 95-100% in a subject such as, but not limited to, the CNS, a region of the CNS, or a specific cell of the CNS, or a muscle, a region of a muscle, or a cell of a muscle, of a subject. In some embodiments, cell of the CNS comprises an astrocyte, microglia, cortical neuron, hippocampal neuron, DRG and/or sympathetic neuron, sensory neuron, oligodendrocyte, motor neuron, or combination thereof.

[0414] In some embodiments, a composition described herein may be used to increase target protein and reduce symptoms of neurological disease in a subject. In some embodiments, the composition may be used to decrease target protein and reduce symptoms of neurological disease in a subject.

[0415] In some embodiments, a composition described herein may be used to reduce the decline of functional capacity and activities of daily living as measured by a standard evaluation system such as, but not limited to, the total functional capacity (TFC) scale.

[0416] In some embodiments, a composition described herein may be used to improve performance on any assessment used to measure symptoms of neurological disease. Such assessments include, but are not limited to ADAS-cog (Alzheimer Disease Assessment Scale – cognitive), MMSE (Mini-Mental State Examination), GDS (Geriatric Depression Scale), FAQ (Functional Activities Questionnaire), ADL (Activities of Daily Living), GPCOG (General Practitioner Assessment of Cognition), Mini-Cog, AMTS (Abbreviated Mental Test Score), Clock-drawing test, 6-CIT (6-item Cognitive Impairment Test), TYM (Test Your Memory), MoCa (Montreal Cognitive Assessment), ACE-R (Addenbrookes Cognitive Assessment), MIS (Memory Impairment Screen), BADLS (Bristol Activities of Daily Living Scale), Barthel Index, Functional Independence Measure, Instrumental Activities of Daily Living, IQCODE (Informant Questionnaire on Cognitive Decline in the Elderly), Neuropsychiatric Inventory, The Cohen-Mansfield Agitation Inventory, BEHAVE-AD, EuroQol, Short Form-36 and/or MBR Caregiver Strain Instrument, or any of the other tests as described in Sheehan B (Ther Adv Neurol Disord. 5(6):349-358 (2012)), the contents of which are herein incorporated by reference in their entirety.

[0417] In some embodiments, the present composition is administered as a solo therapeutic or as combination therapeutic for the treatment of a neurological disease/disorder or a neurodegenerative disorder, a muscular disorder or neuromuscular disorder, and/or a neuro-oncological disorder.

[0418] A composition described herein may be used in combination with one or more other therapeutic agents. In some embodiments, compositions can be administered concurrently with, prior to, or subsequent to, additional therapeutic or medical procedures. In general, each agent will be administered at a dose and/or on a time schedule determined for that agent.

[0419] Therapeutic agents that may be used in combination with the composition can be small molecule compounds which are antioxidants, anti-inflammatory agents, anti-apoptosis agents, calcium regulators, anti-glutamatergic agents, structural protein inhibitors, compounds involved in muscle function, and compounds involved in metal ion regulation. As a non-limiting example, the combination therapy may be in combination with one or more neuroprotective agents such as small molecule compounds, growth factors and hormones which have been tested for their neuroprotective effect on motor neuron degeneration.

[0420] Compounds tested for treating neurological disease which may be used in combination with the AAV particles described herein include, but are not limited to, cholinesterase inhibitors (donepezil, rivastigmine, galantamine), NMDA receptor antagonists such as memantine, anti-psychotics, anti-depressants, anti-convulsants (e.g., sodium valproate and levetiracetam for myoclonus), secretase inhibitors, amyloid aggregation inhibitors, copper or zinc modulators, BACE inhibitors, inhibitors of tau aggregation, such as Methylene blue, phenothiazines, anthraquinones, n-phenylamines or rhodamines, microtubule stabilizers such as NAP, taxol or paclitaxel, kinase or phosphatase inhibitors such as those targeting GSK3β (lithium) or PP2A, immunization with Aβ peptides or tau phospho-epitopes, anti-tau or anti-amyloid antibodies, dopamine-depleting agents (e.g., tetrabenazine for chorea), benzodiazepines (e.g., clonazepam for myoclonus, chorea, dystonia, rigidity, and/or spasticity), , amino acid precursors of dopamine (e.g., levodopa for rigidity), skeletal muscle relaxants (e.g., baclofen, tizanidine for rigidity and/or spasticity), inhibitors for acetylcholine release at the neuromuscular junction to cause muscle paralysis (e.g., botulinum toxin for bruxism and/or dystonia), atypical neuroleptics (e.g., olanzapine and quetiapine for psychosis and/or irritability, risperidone, sulpiride and haloperidol for psychosis, chorea and/or irritability, clozapine for treatment-resistant psychosis, aripiprazole for psychosis with prominent negative symptoms), selective serotonin reuptake inhibitors (SSRIs) (e.g., citalopram, fluoxetine, paroxetine, sertraline, mirtazapine, venlafaxine for depression, anxiety, obsessive compulsive behavior and/or irritability), hypnotics (e.g., xopiclone and/or zolpidem for altered sleep-wake cycle), anticonvulsants (e.g., sodium valproate and carbamazepine for mania or hypomania) and mood stabilizers (e.g., lithium for mania or hypomania).

[0421] Neurotrophic factors may be used in combination therapy with a composition described herein for treating neurological disease. Generally, a neurotrophic factor is defined as a substance that promotes survival, growth, differentiation, proliferation and/or maturation of a neuron, or stimulates increased activity of a neuron. In some embodiments, the present methods further comprise delivery

of one or more trophic factors into the subject in need of treatment. Trophic factors may include, but are not limited to, IGF-I, GDNF, BDNF, CTNF, VEGF, Colivelin, Xaliproden, Thyrotrophin-releasing hormone and ADNF, and variants thereof.

[0422] In one aspect, a composition described herein may be co-administered with AAV particles expressing neurotrophic factors such as AAV-IGF-I (See e.g., Vincent et al., *Neuromolecular medicine*, 2004, 6, 79-85; the contents of which are incorporated herein by reference in their entirety) and AAV-GDNF (See e.g., Wang et al., *J Neurosci.*, 2002, 22, 6920-6928; the contents of which are incorporated herein by reference in their entirety).

[0423] In some embodiments, administration of a composition described herein to a subject will modulate, e.g., increase or decrease, the expression of a target protein in a subject and the modulation, e.g., increase or decrease of the presence, level, activity, and/or expression of the target protein will reduce the effects and/or symptoms of a neurological disease/disorder or a neurodegenerative disorder, a muscular disorder or neuromuscular disorder, and/or a neuro-oncological disorder in a subject.

DEFINITIONS

[0424] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains.

[0425] Articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The disclosure includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The disclosure includes embodiments in which more than one, or the entire group members are present in, employed in, or otherwise relevant to a given product or process.

[0426] It is also noted that the term "comprising" is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term "comprising" is used herein, the term "consisting of" and "consisting essentially thereof" is thus also encompassed and disclosed.

[0427] Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the disclosure, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[0428] *Adeno-associated virus:* As used herein, the term "adeno-associated virus" or "AAV" refers to members of the dependovirus genus or a variant, e.g., a functional variant, thereof. In some

embodiments, the AAV is wildtype, or naturally occurring. In some embodiments, the AAV is recombinant.

[0429] *AAV Particle*: As used herein, an "AAV particle" refers to a particle or a virion comprising an AAV capsid, e.g., an AAV capsid variant, and a polynucleotide, e.g., a viral genome or a vector genome. In some embodiments, the viral genome of the AAV particle comprises at least one payload region and at least one ITR. In some embodiments, an AAV particle of the disclosure is an AAV particle comprising an AAV variant. In some embodiments, the AAV particle is capable of delivering a nucleic acid, e.g., a payload region, encoding a payload to cells, typically, mammalian, e.g., human, cells. In some embodiments, an AAV particle of the present disclosure may be produced recombinantly. In some embodiments, an AAV particle may be derived from any serotype, described herein or known in the art, including combinations of serotypes (e.g., "pseudotyped" AAV) or from various genomes (e.g., single stranded or self-complementary). In some embodiments, the AAV particle may be replication defective and/or targeted. It is to be understood that reference to the AAV particle of the disclosure also includes pharmaceutical compositions thereof, even if not explicitly recited.

[0430] *Administering:* As used herein, the term "administering" refers to providing a pharmaceutical agent or composition to a subject.

[0431] *Amelioration*: As used herein, the term "amelioration" or "ameliorating" refers to a lessening of severity of at least one indicator of a condition or disease. For example, in the context of neurodegeneration disorder, amelioration includes the reduction of neuron loss.

[0432] *Amplicon*: As used herein, "amplicon" may refer to any piece of RNA or DNA formed as the product of amplification events, e.g. PCR. In some embodiments, full-length capsid amplicons may be used as templates for next generation sequencing (NGS) library generation. Full-length capsid amplicons may be used for cloning into a DNA library for any number of additional rounds of AAV selection as described herein.

[0433] *Animal:* As used herein, the term "animal" refers to any member of the animal kingdom. In some embodiments, "animal" refers to humans at any stage of development. In some embodiments, "animal" refers to non-human animals at any stage of development. In certain embodiments, the non-human animal is a mammal (*e.g.*, a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, and worms. In some embodiments, the animal is a transgenic animal, genetically engineered animal, or a clone.

[0434] *Antisense strand:* As used herein, the term "the antisense strand" or "the first strand" or "the guide strand" of a siRNA molecule refers to a strand that is substantially complementary to a section of about 10-50 nucleotides, e.g., about 15-30, 16-25, 18-23 or 19-22 nucleotides of the mRNA of a gene targeted for silencing. The antisense strand or first strand has sequence sufficiently

complementary to the desired target mRNA sequence to direct target-specific silencing, e.g., complementarity sufficient to trigger the destruction of the desired target mRNA by the RNAi machinery or process.

[0435] *Approximately:* As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

[0436] *Biopanning:* As used herein, the term "biopanning" refers to an AAV capsid library selection process comprising administration of an AAV particle with enhanced tissue- and/or cell type-specific transduction to a cell and/or subject; extraction of nucleotides encoded by said AAV particle from said transduced tissue- and/or cell type-specific; and, use of the extracted nucleotides for cloning into a nucleotide library for the generation of AAV particles for subsequent rounds of the same.

[0437] *Capsid*: As used herein, the term "capsid" refers to the exterior, e.g., a protein shell, of a virus particle, e.g., an AAV particle, that is substantially (e.g., >50%, >60%, >70%, >80%, >90%, >95%, >99%, or 100%) protein. In some embodiments, the capsid is an AAV capsid comprising an AAV capsid protein described herein, e.g., a VP1, VP2, and/or VP3 polypeptide. The AAV capsid protein can be a wild-type AAV capsid protein or a variant, e.g., a structural and/or functional variant from a wild-type or a reference capsid protein, referred to herein as an "AAV capsid variant." In some embodiments, the AAV capsid variant described herein has the ability to enclose, e.g., encapsulate, a viral genome and/or is capable of entry into a cell, e.g., a mammalian cell. In some embodiments, the AAV capsid variant described herein may have modified tropism compared to that of a wild-type AAV capsid, e.g., the corresponding wild-type capsid.

[0438] *Complementary and substantially complementary:* As used herein, the term "complementary" refers to the ability of polynucleotides to form base pairs with one another. Base pairs are typically formed by hydrogen bonds between nucleotide units in antiparallel polynucleotide strands. Complementary polynucleotide strands can form base pairs in the Watson-Crick manner (e.g., A to T, A to U, C to G), or in any other manner that allows for the formation of duplexes. As persons skilled in the art are aware, when using RNA as opposed to DNA, uracil rather than thymine is the base that is considered to be complementary to adenine. However, when a U is denoted in the context of the present disclosure, the ability to substitute a T is implied, unless otherwise stated. Perfect complementarity or 100% complementarity refers to the situation in which each nucleotide unit of one polynucleotide strand can form a hydrogen bond with a nucleotide unit of a second polynucleotide

strand. Less than perfect complementarity refers to the situation in which some, but not all, nucleotide units of two strands can form hydrogen bond with each other. For example, for two 20-mers, if only two base pairs on each strand can form a hydrogen bond with each other, the polynucleotide strands exhibit 10% complementarity. In the same example, if 18 base pairs on each strand can form hydrogen bonds with each other, the polynucleotide strands exhibit 90% complementarity. The term "complementary" as used herein can encompass fully complementary, partially complementary, or substantially complementary. As used herein, the term "substantially complementary" means that the siRNA has a sequence (e.g., in the antisense strand) which is sufficient to bind the desired target mRNA, and to trigger the RNA silencing of the target mRNA. "Fully complementary", "perfect complementarity", or "100% complementarity" refers to the situation in which each nucleotide unit of one polynucleotide or oligonucleotide strand can base-pair with a nucleotide unit of a second polynucleotide or oligonucleotide strand.

[0439] *Control Elements:* As used herein, "control elements", "regulatory control elements" or "regulatory sequences" refers to promoter regions, polyadenylation signals, transcription termination sequences, upstream regulatory domains, origins of replication, internal ribosome entry sites ("IRES"), enhancers, and the like, which provide for the replication, transcription and translation of a coding sequence in a recipient cell. Not all of these control elements need always be present as long as the selected coding sequence is capable of being replicated, transcribed and/or translated in an appropriate host cell.

[0440] *Delivery:* As used herein, "delivery" refers to the act or manner of delivering an AAV particle, a compound, substance, entity, moiety, cargo or payload.

[0441] *Element:* As used herein, the term "element" refers to a distinct portion of an entity. In some embodiments, an element may be a polynucleotide sequence with a specific purpose, incorporated into a longer polynucleotide sequence.

[0442] *Encapsulate:* As used herein, the term "encapsulate" means to enclose, surround or encase. As an example, a capsid protein, e.g., an AAV capsid variant, often encapsulates a viral genome. In some embodiments, encapsulate within a capsid, e.g., an AAV capsid variant, encompasses 100% coverage by a capsid, as well as less than 100% coverage, e.g., 95%, 90%, 85%, 80%, 70%, 60% or less. For example, gaps or discontinuities may be present in the capsid so long as the viral genome is retained in the capsid, e.g., prior to entry into a cell.

[0443] *Effective Amount:* As used herein, the term "effective amount" of an agent is that amount sufficient to effect beneficial or desired results, for example, clinical results, and, as such, an "effective amount" depends upon the context in which it is being applied. For example, in the context of administering an agent that treats cancer, an effective amount of an agent is, for example, an amount sufficient to achieve treatment, as defined herein, of cancer, as compared to the response obtained without administration of the agent.

[0444] *Expression*: As used herein, "expression" of a nucleic acid sequence refers to one or more of the following events: (1) production of an RNA template from a DNA sequence (*e.g.*, by transcription); (2) processing of an RNA transcript (*e.g.*, by splicing, editing, 5′ cap formation, and/or 3′ end processing); (3) translation of an RNA into a polypeptide or protein; and (4) post-translational modification of a polypeptide or protein.

[0445] *Formulation*: As used herein, a "formulation" includes at least one AAV particle (active ingredient) and an excipient, and/or an inactive ingredient.

[0446] *Fragment:* A "fragment," as used herein, refers to a portion. For example, an antibody fragment may comprise a CDR, or a heavy chain variable region, or a scFv, etc.

[0447] *Homology*: As used herein, the term "homology" refers to the overall relatedness between polymeric molecules, *e.g.* between polynucleotide molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. In some embodiments, polymeric molecules are considered to be "homologous" to one another if their sequences are at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identical or similar. The term "homologous" necessarily refers to a comparison between at least two sequences (polynucleotide or polypeptide sequences). In accordance with the disclosure, two polynucleotide sequences are considered to be homologous if the polypeptides they encode are at least about 50%, 60%, 70%, 80%, 90%, 95%, or even 99% for at least one stretch of at least about 20 amino acids. In some embodiments, homologous polynucleotide sequences are characterized by the ability to encode a stretch of at least 4–5 uniquely specified amino acids. For polynucleotide sequences less than 60 nucleotides in length, homology is determined by the ability to encode a stretch of at least 4–5 uniquely specified amino acids. In accordance with the disclosure, two protein sequences are considered to be homologous if the proteins are at least about 50%, 60%, 70%, 80%, or 90% identical for at least one stretch of at least about 20 amino acids.

[0448] *Identity*: As used herein, the term "identity" refers to the overall relatedness between polymeric molecules, *e.g.*, between polynucleotide molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of the percent identity of two polynucleotide sequences, for example, can be performed by aligning the two sequences for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second nucleic acid sequences for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In certain embodiments, the length of a sequence aligned for comparison purposes is at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or 100% of the length of the reference sequence. The nucleotides at corresponding nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical

positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which needs to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. For example, the percent identity between two nucleotide sequences can be determined using methods such as those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; the contents of each of which are incorporated herein by reference in their entirety. For example, the percent identity between two nucleotide sequences can be determined using the algorithm of Meyers and Miller (CABIOS, 1989, 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 percent identity between two nucleotide sequences can, alternatively, be determined using the GAP program in the GCG software package using an NWSgapdna.CMP matrix. Methods commonly employed to determine percent identity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM J Applied Math., 48:1073 (1988); incorporated herein by reference. Techniques for determining identity are codified in publicly available computer programs. Exemplary computer software to determine homology between two sequences include, but are not limited to, GCG program package, Devereux, J., *et al.*, *Nucleic Acids Research*, 12(1), 387 (1984)), BLASTP, BLASTN, and FASTA Altschul, S. F. *et al.*, *J. Molec. Biol.*, 215, 403 (1990)).

[0449] *Inhibit expression of a gene:* As used herein, the phrase "inhibit expression of a gene" means to cause a reduction in the amount of an expression product of the gene. The expression product can be an RNA transcribed from the gene (*e.g.*, an mRNA) or a polypeptide translated from an mRNA transcribed from the gene. Typically, a reduction in the level of an mRNA results in a reduction in the level of a polypeptide translated therefrom. The level of expression may be determined using standard techniques for measuring mRNA or protein.

[0450] *Inverted terminal repeat:* As used herein, the term "inverted terminal repeat" or "ITR" refers to a cis-regulatory element for the packaging of polynucleotide sequences into viral capsids.

[0451] *Isolated*: As used herein, the term "isolated" refers to a substance or entity that is altered or removed from the natural state, e.g., altered or removed from at least some of component with which it is associated in 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. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of the environment in which it is found in nature. In some embodiments, an isolated nucleic acid is recombinant, e.g., incorporated into a vector.

[0452] *Library:* As used herein, the term "library" refers to a diverse collection of linear polypeptides, polynucleotides, viral particles, or viral vectors. As examples, a library may be a DNA library or an AAV capsid library.

[0453] *Ligand*: As used herein, the term "ligand" refers to molecule that binds to a target, e.g., a receptor. In some embodiments, the receptor is a GPI-anchored protein, e.g., as described herein. In some embodiments, the receptor is alkaline phosphatase (ALPL), e.g., human ALPL, NHP ALPL, or murine ALPL. In some embodiments, the ligand is or comprises a peptide, a protein, an antibody molecule, a nucleic acid molecule (e.g., an aptamer), or a small molecule, optionally in isolation or as part of a fusion or a conjugate, e.g., with an active agent.

[0454] *Molecular scaffold:* As used herein a "molecular scaffold" is a framework or starting molecule that forms the sequence or structural basis against which to design or make a subsequent molecule.

[0455] *Neurological disease:* As used herein, a "neurological disease" is any disease associated with the central or peripheral nervous system and components thereof (e.g., neurons).

[0456] *Orthogonal evolution:* As used herein, the term "orthogonal evolution" refers to a method wherein AAV particles are administered for a first round of AAV selection as described herein across a set of any number of cell- and/or subject-types that may be from different species and/or strains, and wherein any number of additional, i.e., subsequent, AAV selection rounds are performed either across a set of any number of cell- and/or subject-types that may be from different species and/or strains, or across a set of any number of cell- and/or subject-types that may be from the same species and/or strain.

[0457] *Open reading frame:* As used herein, "open reading frame" or "ORF" refers to a sequence which does not contain a stop codon in a given reading frame.

[0458] *Particle*: As used herein, a "particle" is a virus comprised of at least two components, a protein capsid and a polynucleotide sequence enclosed within the capsid.

[0459] *Payload region:* As used herein, a "payload region" is any nucleic acid sequence (e.g., within the viral genome) which encodes one or more "payloads" of the disclosure. As non-limiting examples, a payload region may be a nucleic acid sequence within the viral genome of an AAV particle, which encodes a payload, wherein the payload is an RNAi agent or a polypeptide. Payloads of the present disclosure may be, but are not limited to, peptides, polypeptides, proteins, antibodies, RNAi agents, etc.

[0460] *Polypeptide:* As used herein, "polypeptide" means a polymer of amino acid residues (natural or unnatural) linked together most often by peptide bonds. The term, as used herein, refers to proteins, polypeptides, and peptides of any size, structure, or function. In some instances, the polypeptide encoded is smaller than about 50 amino acids and the polypeptide is then termed a peptide. If the polypeptide is a peptide, it will be at least about 2, 3, 4, or at least 5 amino acid residues long. Thus, polypeptides include gene products, naturally occurring polypeptides, synthetic polypeptides, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing. A polypeptide may be a single molecule or may be a multi-molecular complex such as a dimer, trimer or tetramer. They may also comprise single chain or multichain polypeptides and may be associated or linked. The term polypeptide may also apply to amino acid polymers in which one or more amino acid residues are an artificial chemical analogue of a corresponding naturally occurring amino acid.

[0461] *Polypeptide variant:* The term "polypeptide variant" refers to molecules which differ in their amino acid sequence from a native or reference sequence. The amino acid sequence variants may possess substitutions, deletions, and/or insertions at certain positions within the amino acid sequence, as compared to a native or reference sequence. In some embodiments, a variant comprises a sequence having at least about 50%, at least about 80%, or at least about 90%, identical (homologous) to a native or a reference sequence.

[0462] *Peptide:* As used herein, "peptide" is less than or equal to 50 amino acids long, e.g., about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 amino acids long.

[0463] *Pharmaceutically acceptable*: The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

[0464] *Preventing*: As used herein, the term "preventing" or "prevention" refers to partially or completely delaying onset of an infection, disease, disorder and/or condition; partially or completely delaying onset of one or more symptoms, features, or clinical manifestations of a particular infection, disease, disorder, and/or condition; partially or completely delaying onset of one or more symptoms, features, or manifestations of a particular infection, disease, disorder, and/or condition; partially or completely delaying progression from an infection, a particular disease, disorder and/or condition; and/or decreasing the risk of developing pathology associated with the infection, the disease, disorder, and/or condition.

[0465] *Prophylactic*: As used herein, "prophylactic" refers to a therapeutic or course of action used to prevent the spread of disease.

[0466] *Prophylaxis:* As used herein, a "prophylaxis" refers to a measure taken to maintain health and prevent the spread of disease.

[0467] *Region:* As used herein, the term "region" refers to a zone or general area. In some embodiments, when referring to a protein or protein module, a region may comprise a linear sequence of amino acids along the protein or protein module or may comprise a three-dimensional area, an epitope and/or a cluster of epitopes. In some embodiments, regions comprise terminal regions. As used herein, the term "terminal region" refers to regions located at the ends or termini of a given agent. When referring to proteins, terminal regions may comprise N- and/or C-termini.

[0468] In some embodiments, when referring to a polynucleotide, a region may comprise a linear sequence of nucleic acids along the polynucleotide or may comprise a three-dimensional area, secondary structure, or tertiary structure. In some embodiments, regions comprise terminal regions. As used herein, the term "terminal region" refers to regions located at the ends or termini of a given agent. When referring to polynucleotides, terminal regions may comprise 5' and/or 3' termini.

[0469] *RNA or RNA molecule*: As used herein, the term "RNA" or "RNA molecule" or "ribonucleic acid molecule" refers to a polymer of ribonucleotides; the term "DNA" or "DNA molecule" or "deoxyribonucleic acid molecule" refers to a polymer of deoxyribonucleotides. DNA and RNA can be synthesized naturally, e.g., by DNA replication and transcription of DNA, respectively; or be chemically synthesized. DNA and RNA can be single-stranded (i.e., ssRNA or ssDNA, respectively) or multi-stranded (e.g., double stranded, i.e., dsRNA and dsDNA, respectively). The term "mRNA" or "messenger RNA", as used herein, refers to a single stranded RNA that encodes the amino acid sequence of one or more polypeptide chains.

[0470] *RNA interfering or RNAi:* As used herein, the term "RNA interfering" or "RNAi" refers to a sequence specific regulatory mechanism mediated by RNA molecules which results in the inhibition or interfering or "silencing" of the expression of a corresponding protein-coding gene. RNAi has been observed in many types of organisms, including plants, animals and fungi. RNAi occurs in cells naturally to remove foreign RNAs (e.g., viral RNAs). Natural RNAi proceeds via fragments cleaved from free dsRNA which direct the degradative mechanism to other similar RNA sequences. RNAi is controlled by the RNA-induced silencing complex (RISC) and is initiated by short/small dsRNA molecules in cell cytoplasm, where they interact with the catalytic RISC component argonaute. The dsRNA molecules can be introduced into cells exogenously. Exogenous dsRNA initiates RNAi by activating the ribonuclease protein Dicer, which binds and cleaves dsRNAs to produce double-stranded fragments of 21-25 base pairs with a few unpaired overhang bases on each end. These short double stranded fragments are called small interfering RNAs (siRNAs).

[0471] *RNAi agent:* As used herein, the term "RNAi agent" refers to an RNA molecule, or its derivative, that can induce inhibition, interfering, or "silencing" of the expression of a target gene and/or its protein product. An RNAi agent may knock-out (virtually eliminate or eliminate)

expression, or knock-down (lessen or decrease) expression. The RNAi agent may be, but is not limited to, dsRNA, siRNA, shRNA, pre-miRNA, pri-miRNA, miRNA, stRNA, lncRNA, piRNA, or snoRNA.

[0472] *miR binding site:* As used herein, a "miR binding site" comprises a nucleic acid sequence (whether RNA or DNA, e.g., differ by "U" of RNA or "T" in DNA) that is capable of binding, or binds, in whole or in part to a microRNA (miR) through complete or partial hybridization . Typically, such binding occurs between the miR and the miR binding site in the reverse complement orientation. In some embodiments, the miR binding site is transcribed from the AAV vector genome encoding the miR binding site.

[0473] In some embodiments, a miR binding site may be encoded or transcribed in series. Such a "miR binding site series" or "miR BSs" may include two or more miR binding sites having the same or different nucleic acid sequence.

[0474] *Spacer*: As used here, a "spacer" is generally any selected nucleic acid sequence of, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length, which is located between two or more consecutive miR binding site sequences. Spacers may also be more than 10 nucleotides in length, e.g., 20, 30, 40, or 50 or more than 50 nucleotides.

[0475] *Sample:* As used herein, the term "sample" or "biological sample" refers to a subset of its tissues, cells, nucleic acids, or component parts (e.g. body fluids, including but not limited to blood, serum, mucus, lymphatic fluid, synovial fluid, cerebrospinal fluid, saliva, amniotic fluid, amniotic cord blood, urine, vaginal fluid and semen).

[0476] *Self-complementary viral particle*: As used herein, a "self-complementary viral particle" is a particle comprised of at least two components, a protein capsid and a self-complementary viral genome enclosed within the capsid.

[0477] *Sense Strand:* As used herein, the term "the sense strand" or "the second strand" or "the passenger strand" of a siRNA molecule refers to a strand that is complementary to the antisense strand or first strand. The antisense and sense strands of a siRNA molecule are hybridized to form a duplex structure. As used herein, a "siRNA duplex" includes a siRNA strand having sufficient complementarity to a section of about 10-50 nucleotides of the mRNA of the gene targeted for silencing and a siRNA strand having sufficient complementarity to form a duplex with the other siRNA strand.

[0478] *Similarity*: As used herein, the term "similarity" refers to the overall relatedness between polymeric molecules, *e.g.* between polynucleotide molecules (*e.g.* DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Calculation of percent similarity of polymeric molecules to one another can be performed in the same manner as a calculation of percent identity, except that calculation of percent similarity takes into account conservative substitutions as is understood in the art.

[0479] *Short interfering RNA or siRNA:* As used herein, the terms "short interfering RNA," "small interfering RNA" or "siRNA" refer to an RNA molecule (or RNA analog) comprising between about 5-60 nucleotides (or nucleotide analogs) which is capable of directing or mediating RNAi. Preferably, a siRNA molecule comprises between about 15-30 nucleotides or nucleotide analogs, such as between about 16-25 nucleotides (or nucleotide analogs), between about 18-23 nucleotides (or nucleotide analogs), between about 19-22 nucleotides (or nucleotide analogs) (e.g., 19, 20, 21 or 22 nucleotides or nucleotide analogs), between about 19-25 nucleotides (or nucleotide analogs), and between about 19-24 nucleotides (or nucleotide analogs). The term "short" siRNA refers to a siRNA comprising 5-23 nucleotides, preferably 21 nucleotides (or nucleotide analogs), for example, 19, 20, 21 or 22 nucleotides. The term "long" siRNA refers to a siRNA comprising 24-60 nucleotides, preferably about 24-25 nucleotides, for example, 23, 24, 25 or 26 nucleotides. Short siRNAs may, in some instances, include fewer than 19 nucleotides, e.g., 16, 17 or 18 nucleotides, or as few as 5 nucleotides, provided that the shorter siRNA retains the ability to mediate RNAi. Likewise, long siRNAs may, in some instances, include more than 26 nucleotides, e.g., 27, 28, 29, 30, 35, 40, 45, 50, 55, or even 60 nucleotides, provided that the longer siRNA retains the ability to mediate RNAi or translational repression absent further processing, e.g., enzymatic processing, to a short siRNA. siRNAs can be single stranded RNA molecules (ss-siRNAs) or double stranded RNA molecules (ds-siRNAs) comprising a sense strand and an antisense strand which hybridized to form a duplex structure called an siRNA duplex.

[0480] *Subject:* As used herein, the term "subject" or "patient" refers to any organism to which a composition in accordance with the disclosure may be administered, *e.g.*, for experimental, diagnostic, prophylactic, and/or therapeutic purposes. Typical subjects include animals (*e.g.*, mammals such as mice, rats, rabbits, non-human primates, and humans) and/or plants.

[0481] *Substantially*: As used herein, the term "substantially" refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term "substantially" is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

[0482] *Target Cells:* As used herein, "target cells" or "target tissue" refers to any one or more cells of interest. The cells may be found *in vitro*, *in vivo*, *in situ* or in the tissue or organ of an organism. The organism may be an animal, preferably a mammal, more preferably a human and most preferably a patient.

[0483] *Therapeutic Agent:* The term "therapeutic agent" refers to any agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

[0484] *Therapeutically effective amount:* As used herein, the term "therapeutically effective amount" means an amount of an agent to be delivered (*e.g.*, nucleic acid, drug, therapeutic agent, diagnostic agent, prophylactic agent, *etc.*) that is sufficient, when administered to a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition. In some embodiments, a therapeutically effective amount is provided in a single dose.

[0485] *Therapeutically effective outcome*: As used herein, the term "therapeutically effective outcome" means an outcome that is sufficient in a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition.

[0486] *Treating*: As used herein, the term "treating" refers to partially or completely alleviating, ameliorating, improving, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of a particular infection, disease, disorder, and/or condition. For example, "treating" cancer may refer to inhibiting survival, growth, and/or spread of a tumor. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition and/or to a subject who exhibits only early signs of a disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

[0487] *Conservative amino acid substitution:* As used herein, a "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

[0488] *Variant:* As used herein, the term "variant" refers to a polypeptide or polynucleotide that has an amino acid or a nucleotide sequence that is substantially identical, e.g., having at least 70%, 75%, 80%, 85%, 90%, 95% or 99% sequence identity to a reference sequence. In some embodiments, the variant is a functional variant.

[0489] *Functional Variant*: As used herein, the term "functional variant" refers to a polypeptide variant or a polynucleotide variant that has at least one activity of the reference sequence.

[0490] *Insertional Variant*: "Insertional variants" when referring to polypeptides are those with one or more amino acids inserted, e.g., immediately adjacent or subsequent, to a position in an amino

acid sequence. "Immediately adjacent" or "immediately subsequent" to an amino acid means connected to either the alpha-carboxy or alpha-amino functional group of the amino acid.

[0491] *Deletional Variant*: "Deletional variants" when referring to polypeptides, are those with one or more amino acids in deleted from a reference protein.

[0492] *Vector:* As used herein, the term "vector" refers to any molecule or moiety which transports, transduces or otherwise acts as a carrier of a heterologous molecule. In some embodiments, vectors may be plasmids. In some embodiments, vectors may be viruses. An AAV particle is an example of a vector. Vectors of the present disclosure may be produced recombinantly and may be based on and/or may comprise adeno-associated virus (AAV) parent or reference sequences. The heterologous molecule may be a polynucleotide and/or a polypeptide.

[0493] *Viral Genome:* As used herein, the terms "viral genome" or "vector genome" refer to the nucleic acid sequence(s) encapsulated in an AAV particle. A viral genome comprises a nucleic acid sequence with at least one payload region encoding a payload and at least one ITR.

Equivalents and Scope

[0494] The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety. While this invention has been disclosed with reference to certain embodiments, it is apparent that further embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

[0495] The present disclosure is further illustrated by the following non-limiting examples.

EXAMPLES

Example 1. High-throughput screen of TRACER AAV library in NHP and Mice

[0496] A TRACER based method as described in WO2020072683, WO 2021/202651, and WO2021230987, the contents of which are herein incorporated by reference in their entirety, was used to generate the AAV capsid variants described herein. An orthogonal evolution approach was combined with a high throughput screening by NGS. Briefly, the library of AAV capsid variants was generated using a sliding window approach, where 6 amino acid sequences were inserted into 8 different positions across loop IV of AAV9, including immediately subsequent to positions 453, 454, 455, 456, 457, 458, 459, and 460, relative to a reference sequence numbered according to SEQ ID NO: 138. The initial library was passed twice through non-human primates (NHP, 2-4 years of age). After the second passage (e.g., 28 days post injection into two NHPs), RNA was extracted from six brain regions. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate fold enrichment relative to an AAV9 wild-type control. Following these two passages, approximately 21195 variants were identified with an average fold

change greater than wild-type. Of the 21195 variants, 1558 demonstrated a fold-change of greater than 6 compared to wild-type and were detected across all brain regions investigated. Of these 1558, approximately 1470 variants were selected for constructing a synthetic library and a third passage through two NHPs. Within the 1470 variants selected for further characterization and investigation, there was a relatively even distribution for each insertion position of the sliding window used to generate the initial library.

[0497] After creation of the synthetic library with the sub-selected variants, the synthetic library was screened (passage 3) in two NHPs (2-4 years of age) and two strains of mice, BALB/c (n=3, 6-8 weeks of age) and C57Bl/6 mice (n=3, 6-8 weeks of age), in a first cross-species evolution screen. The animals were injected intravenously with the synthetic library. After a period *in vivo*, (e.g., 28-days) RNA was extracted from nervous tissue, e.g., brain, spinal cord, and DRG of the NHPs and the brains of mice. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed, and the peptides comprised within the variants were identified and the capsid enrichment ratio for each variant compared to the wild-type AAV9 control was calculated (fold enrichment relative to wild-type AAV9) (**Table 9**). Values above 1 indicate an increase in expression relative to AAV9. All animals were dosed intravenously at 2-3 VG/kg across the screen.

[0498] As shown in Table 9, approximately 700 variants demonstrated an increase in expression relative to AAV9, and several variants demonstrated a greater than 10-fold enrichment relative to AAV9 in the brain of NHPs. Further, the variants demonstrating the greatest fold enrichment in the brain also demonstrated the greatest fold enrichment in the spinal cord relative to AAV9 in NHPs. These variants also demonstrated de-targeting in the DRG (data not shown). For instance, the variant comprising GSGSPHSKAQNQQT (SEQ ID NO: 200) demonstrated a 76.6 fold enrichment in the brain, a 29.4 fold enrichment in the spinal cord, and 0.4 fold enrichment in the DRG of NHPs relative to AAV9; and GHDSPHKSGQNQQT (SEQ ID NO: 201) demonstrated a 62.6 fold enrichment in the brain, a 15.6 fold enrichment in the spinal cord, and 0.0 fold enrichment in the DRG of NHPs relative to AAV9. Also, across the peptides comprised within the AAV capsid variants with the greatest fold-enrichment in the NHP brain relative wild-type AAV9, it was observed that each of these peptides comprised an SPH motif in the same position (e.g., immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138), regardless of the insertion position within the variant capsid, as well as a positive amino acid (e.g., K or R) in one of the next three residues subsequent to the SPH motif.

[0499] Those variants with the greatest fold enrichment in the brains of NHPs also had the greatest fold enrichment in the brains of both mouse species. Also, when comparing the fold enrichment relative to wild-type for each variant between the two species of mice investigated (C57Bl/6 and BALB/c mice), they were highly correlated (R^2= 0.8591).

Table 9. NGS fold-enrichment of AAV capsid variants in NHPs and mice

Peptide Sequence	SEQ ID NO:	Fold enrichment over AAV9 in the brain of NHPs	Fold enrichment over AAV9 in the spinal cord of NHPs	Fold enrichment over AAV9 in the brain of Mice (C57Bl/6)	Fold enrichment over AAV9 in the brain of Mice (BALB/c)
GSGSPHSKAQNQQT	200	73.615	29.402	25.293	41.304
GHDSPHKSGQNQQT	201	62.612	15.641	63.993	49.760
GSGSPHARMQNQQT	202	56.138	22.690	7.795	4.164
GSGSPHVKSQNQQT	203	37.551	13.649	8.069	15.861
GQDSPHKSGQNQQT	204	24.569	3.548	57.344	42.615
GSGSPHASRQNQQT	205	18.265	7.804	28.028	36.577
GSGSPHASRQNKQT	206	17.520	35.029	13.096	18.114
GSGSPHVKIQNQQT	207	16.854	9.068	2.173	2.227
GSGSPHSKAKNQQT	208	14.458	0.049	21.494	23.556
GSGSPHKKNQNQQT	209	12.991	0.379	25.958	7.415
GSGSPHVRMQNQQT	210	11.574	6.764	9.121	10.076
GSGSPHASRQKQQT	211	11.417	0.005	7.413	12.400
GHSSPHRSGQNQQT	212	10.357	1.887	23.197	25.442
GMRTYHLSGQNQQT	213	9.241	1.939	2.033	1.586
GSGSPHTRGQNQQT	214	7.092	3.815	10.801	6.240
GSGIIPVSSQNQQT	215	6.352	0.000	0.642	0.253
GSEYGHKSGQNQQT	216	6.308	2.750	5.198	5.332
GRGQNVSSVHRQQT	217	5.404	0.000	1.206	0.691
GSSHRFYGGQNQQT	218	4.732	0.000	0.787	0.110
GYFVAAWSGQNQQT	219	4.488	0.000	0.071	0.175
GSVLHSHAAQNQQT	220	4.150	6.448	0.675	0.423
GSGDLVVSTQNQQT	221	3.874	1.177	0.411	0.273
GSYGMAASGQNQQT	222	3.817	10.052	1.274	0.829
GLNHFGASGQNQQT	223	3.802	3.188	0.774	0.579
GSTGSHSAGQNQHT	224	3.717	0.285	1.190	0.850
GLAGHTVSGQNQQT	225	3.632	0.229	0.972	0.202
GIILGASSGQNQQT	226	3.630	4.868	1.378	0.865
GSGVSTYNIQNQQT	227	3.609	2.912	0.769	0.520
GSLVSVQTGQNQQT	228	3.534	6.043	0.903	0.469
GQSSPHRSGQNQQT	229	3.496	2.142	12.352	19.366
GREYGHKSGQNQQT	230	3.453	0.000	1.422	0.959
GHTLTLSSGQNQQT	231	3.405	5.648	0.648	0.606
GSITLIPSGQNQQT	232	3.361	3.917	0.326	0.435
GSNGFTALGQNQQT	233	3.361	2.663	0.830	0.332
GSGHSSHSVQNQQT	234	3.339	3.318	0.942	0.424
GSGIPQRSGKNQQT	235	3.331	0.000	1.418	1.685
GSGDTLHMLQNQQT	236	3.317	1.174	0.393	0.482
GERHTVLSGQNQQT	237	3.289	3.008	1.027	0.607
GSGMPQSHIQNQQT	238	3.289	11.609	0.514	0.334
GSGQLSGIGGNQQT	239	3.266	0.287	0.993	0.626
GSGQNRKPASFAQT	240	3.204	0.000	0.892	1.061
GSGSVSQLGQNQQT	241	3.184	2.307	0.596	0.375
GSDFLGTHGQNQQT	242	3.171	0.348	1.038	0.750
GQIVQNPSGQNQQT	243	3.133	0.406	1.446	0.635
GSGTQIPSQQNQQT	244	3.112	1.224	0.470	0.151
GSGQNQQSAREGLT	245	3.111	5.632	1.221	1.104
GSGLGMSTGQNQQT	246	3.110	5.499	0.458	0.660

GSGLPVLSGQNQQT	247	3.100	4.149	0.631	0.210
GSGHSIRTDQNQQT	248	3.074	15.600	0.229	0.148
GSGQSVQTVVNQQT	249	3.057	5.441	0.582	0.240
GSGQNRAQSRFQQT	250	3.043	0.000	0.619	1.788
GGGDLGRSSQNQQT	251	3.036	4.830	0.916	0.539
GGGTKMDSGQNQQT	252	3.034	0.000	0.733	0.297
GSGSPHPSRQNQQT	253	3.017	1.993	1.869	0.975
GSGQFTNAGMNQQT	254	2.969	0.936	0.565	0.418
GGRNGHTVGQNQQT	255	2.965	3.732	1.105	1.003
GSGFGPQTGQNQQT	256	2.964	2.861	1.280	0.849
GRTDSHTSGQNQQT	257	2.913	1.510	1.299	0.704
GYEVLGSSGQNQQT	258	2.891	0.000	2.459	0.319
GSVHLSVTGQNQQT	259	2.882	1.157	0.741	0.282
GFMSYKGSGQNQQT	260	2.865	0.209	1.808	0.569
GNIAGSVSGQNQQT	261	2.849	1.187	0.446	0.257
GSGSHRDVSQNQQT	262	2.843	4.022	0.626	0.550
GGLGSMSSGQNQQT	263	2.812	1.405	1.802	0.822
GSGHLPQSAQNQQT	264	2.803	7.828	0.826	0.496
GGVLVGGSGQNQQT	265	2.778	0.178	1.527	0.688
GTHPYTSSGQNQQT	266	2.775	1.684	0.758	0.471
GSGQNQQLKENRST	267	2.765	0.062	1.149	1.118
GSGQNQQTSPHNHT	268	2.761	3.132	1.524	0.845
GSGTLYPQSQNQQT	269	2.761	5.558	0.324	0.160
GSGQNQQSNWITKT	270	2.711	0.000	0.540	0.634
GSGYTSLFLQNQQT	271	2.710	0.010	0.490	1.044
GSGVMTHVLQNQQT	272	2.692	0.347	0.370	0.533
GSVSDVRAGQNQQT	273	2.661	1.647	0.267	0.747
GSGQSHMATLNQQT	274	2.657	0.724	1.173	0.504
GSGLSVHLAQNQQT	275	2.657	1.234	0.806	0.508
GSGLSHATQQNQQT	276	2.640	7.819	1.111	0.638
GSGLSVQSGQNQQT	277	2.637	2.929	1.695	1.005
GSGHMTYREKNQQT	278	2.633	5.267	1.257	0.540
GSKGVPTPGQNQQT	279	2.625	1.292	1.452	0.459
GSGLLPLSSQNQQT	280	2.612	1.130	0.501	0.293
GNGLYAVSGQNQQT	281	2.611	9.148	0.322	0.213
GFNGSPSSGQNQQT	282	2.609	12.197	2.338	0.924
GSGQIRHSDQNQQT	283	2.600	12.884	1.170	0.320
GGQVAPSSGQNQQT	284	2.581	2.427	1.433	0.709
GSFSMHTHGQNQQT	285	2.535	0.118	1.027	0.693
GSGQNQQVIQGSNT	286	2.521	8.778	0.935	0.810
GRVLHSHAGQNQQT	287	2.513	0.826	1.294	0.908
GSGQNQQTSLQDQT	288	2.505	0.500	0.315	0.968
GSGLGRAPVQNQQT	289	2.503	2.214	0.841	0.383
GNGFSSASGQNQQT	290	2.493	0.772	0.240	0.182
GSGQMASRESNQQT	291	2.492	0.300	0.341	0.288
GPGLPNHSGQNQQT	292	2.486	1.992	1.197	0.659
GNIQWQGSGQNQQT	293	2.468	6.266	1.182	0.837
GMSAHMSSGQNQQT	294	2.456	5.255	1.310	0.947
GHSFVNRSGQNQQT	295	2.447	11.148	1.305	0.756
GRAVMDHSGQNQQT	296	2.408	3.209	0.728	0.283
GALTVMQSGQNQQT	297	2.381	0.430	0.246	0.199
GSGQRSPVLPNQQT	298	2.369	6.230	0.434	0.526
GSGQNGHLSLKQQT	299	2.362	1.896	0.718	0.270

GSLPRGTSDQNQQT	300	2.362	0.000	0.453	0.495
GVAGSLVSGQNQQT	301	2.358	7.670	1.321	1.160
GRGGIPQSGQNQQT	302	2.352	8.683	1.639	1.181
GSGQYASSIPNQQT	303	2.346	3.321	1.022	0.489
GTDFGRQSSQNQQT	304	2.346	3.196	1.021	0.797
GIFMQTPSGQNQQT	305	2.344	6.198	0.938	0.252
GSGQNQQTRLVDLT	306	2.342	9.348	1.268	0.490
GTREMPLSGQNQQT	307	2.339	2.830	1.436	0.538
GSRLVHVHGQNQQT	308	2.334	1.174	1.277	0.934
GSGRLVPNGPNQQT	309	2.314	3.925	0.639	0.411
GSGYLRESPQNQQT	310	2.311	0.878	0.331	0.677
GARIQNASGQKQQT	311	2.300	2.103	1.220	1.039
GLSNPMPSGQNQQT	312	2.280	6.033	1.190	0.829
GSTVQDTRGQNQQT	313	2.270	4.979	0.576	0.473
GPFGMPSSGQNQQT	314	2.260	2.700	0.727	0.560
GSGQNHGVLSNQQT	315	2.254	1.603	1.113	0.701
GSGYSMSQAQNQQT	316	2.250	4.479	0.519	0.329
GSGMLTHTLQNQQT	317	2.246	2.272	0.496	0.199
GRGSPHASRQNQQT	318	2.241	0.000	5.050	5.856
GLSWPSTSGQNQQT	319	2.238	0.000	0.910	0.610
GNSMERTSGQNQQT	320	2.221	4.177	1.047	0.935
GSGMSPSTLQNQQT	321	2.216	3.053	0.318	0.153
GSGHGQVLSQNQQT	322	2.213	12.133	1.880	0.661
GRGQIYSTGGNQQT	323	2.210	11.629	1.329	0.743
GVVAAHNSGQNQQT	324	2.202	1.301	1.196	1.336
GDSSLRHSGQNQQT	325	2.194	0.000	0.662	0.412
GSLVSQGAGQNQQT	326	2.188	4.414	1.436	1.246
GSLLQAHSGQNQQT	327	2.182	1.008	0.575	0.748
GSGHIYVGIQNQQT	328	2.178	6.428	0.989	0.337
GHHTTVQSGQNQQT	329	2.177	6.245	0.851	0.755
GSRQSKRNELNQQT	330	2.177	0.000	1.325	0.232
GSGQNQQHVSSPRT	331	2.176	1.279	1.847	0.938
GSSKELLWGQNQQT	332	2.163	0.000	0.506	0.883
GSLSTPSSGQNQQT	333	2.159	1.279	1.094	0.669
GSIGYAGQGQNQQT	334	2.157	4.951	1.604	0.712
GSGQNQRVSNSQQT	335	2.146	0.492	1.086	0.985
GSGYASHVQQNQQT	336	2.146	3.038	1.157	0.758
GSGEYSRSGQNQQT	337	2.145	0.745	0.617	0.205
GSVSTHSSGQNQQT	338	2.145	3.446	1.198	0.918
GSGQNQHSLGNYQT	339	2.143	1.896	1.077	0.606
GSGGLDTRGQNQQT	340	2.139	6.216	0.236	0.197
GNILHATSGQNQQT	341	2.136	0.125	1.159	0.424
GSGQSYTMTQNQQT	342	2.136	6.755	0.297	0.231
GSGQNQHSAPNSQT	343	2.134	4.143	1.187	0.731
GSGQNQQTMDHNRT	344	2.130	4.944	0.642	0.440
GSNGGVGTGQNQQT	345	2.130	0.788	1.191	1.087
GAGSIIPSGQNQQT	346	2.129	7.164	0.595	0.249
GSGQTHGGQHNQQT	347	2.125	12.251	1.448	1.098
GSNLSFQSGQNQQT	348	2.122	5.853	1.087	0.719
GATLQVHSGQNQQT	349	2.122	2.219	0.623	0.545
GSGFNQRSEQNQQT	350	2.121	4.491	1.770	0.758
GSGSLRDFDQNQQT	351	2.120	6.846	0.586	0.272
GSGDSITGKQNQQT	352	2.112	1.295	0.793	0.306

GSGQDRNIVQNQQT	353	2.112	0.229	0.454	0.632
GSGLSHSHQQNQQT	354	2.109	5.852	1.256	0.592
GSGQNQQTGMSSVK	355	2.109	4.544	1.451	0.679
GSVTHGISGQNQQT	356	2.105	4.542	1.135	0.789
GVVAHQPSGQNQQT	357	2.103	0.152	0.910	2.267
GSGPILGQLQNQQT	358	2.097	2.058	0.470	0.123
GSGHVPNSGLNQQT	359	2.091	0.653	1.636	1.154
GDAGVRSSGQNQQT	360	2.068	3.918	1.033	1.193
GSGSQLMSLQNQQT	361	2.065	3.559	0.563	0.172
GSGLDYSQRQNQQT	362	2.056	0.837	0.484	0.217
GSGQSSGRLINKQT	363	2.055	28.135	0.543	0.277
GSSVSPSSGQNQQT	364	2.054	0.579	1.064	0.787
GSGQVVGLSGNQQT	365	2.052	7.212	0.785	0.881
GSNMGVPLGQNQQT	366	2.049	2.448	0.334	0.420
GSFYPSSTGQNQQT	367	2.047	2.374	0.420	0.277
GSGQNQQTRLTDLT	368	2.046	8.470	0.910	0.776
GPTNGRSSGQNQQT	369	2.034	8.903	0.936	1.308
GSGLLHGKLQNQQT	370	2.032	2.521	1.068	0.917
GANMGHVSGQNQQT	371	2.020	0.810	1.302	1.138
GSGQNQQSGRGDLT	372	2.019	6.919	0.524	1.152
GSHGHYASGQKQQT	373	2.016	0.000	0.895	0.685
GSGDLRISPQNQQT	374	2.012	16.207	0.620	0.237
GSGMPVILGQNQQT	375	2.005	0.150	0.840	0.287
GRGVITSSGQNHQT	376	2.004	0.864	1.656	0.669
GSGHSVSGPQNQQT	377	1.993	6.259	1.370	0.619
GSRNGHTVGRNQQT	378	1.993	0.000	1.162	0.367
GAGVHMVSGQNQQT	379	1.987	6.488	1.055	0.791
GSGQNHRPSVLQQT	380	1.983	5.582	0.433	0.582
GSGSPRDSIQNQQT	381	1.981	4.914	0.171	0.446
GSGQGIHSSVNQQT	382	1.981	4.873	0.632	0.634
GSGQQLSITPNQQT	383	1.979	10.280	0.845	0.201
GGYHSQTSGQNQQT	384	1.978	2.642	1.740	1.525
GSLHHDNHGQNQQT	385	1.976	0.980	0.968	0.463
GIMARDSSGQNQQT	386	1.972	3.486	1.320	0.904
GVVHITNSGQNQQT	387	1.969	0.504	0.794	0.846
GSGQNQHSAPFNQT	388	1.969	0.499	0.759	0.870
GSGQTSGLKQNQQT	389	1.968	3.927	0.394	0.334
GSGQNQQTSLSNTA	390	1.959	1.186	1.567	1.182
GSGQNQAVHNKSQT	391	1.956	3.791	1.465	1.083
GVHTHLPSGQNQQT	392	1.952	1.364	1.414	0.796
GHLTMHNSGQNHQT	393	1.938	1.798	1.030	0.586
GSGSSSRPYQNQQT	394	1.934	3.823	0.962	0.496
GILLATPSGQNQQT	395	1.931	8.205	1.341	0.288
GSGQNAGSFPNQQT	396	1.928	12.575	1.091	0.286
GSRDGHTVGQNQQT	397	1.928	7.089	0.495	0.661
GSLLISTSGQNQQT	398	1.919	5.763	1.488	0.808
GSGAMPSHGQNQQT	399	1.915	0.000	1.142	0.912
GALVSPISGQNQQT	400	1.912	1.051	0.640	0.347
GSLSSHGVGQNQQT	401	1.911	7.498	1.218	0.804
GSGQNQQASLAMRT	402	1.910	3.577	2.066	1.638
GPGLGSHSGQNQQT	403	1.906	14.563	0.880	1.195
GHDSQHKSGQNQQT	404	1.904	6.988	1.154	0.869
GSGLTLSATQNQQT	405	1.901	0.193	0.708	0.340

GSGQVVAHVGNQQT	406	1.901	0.833	0.800	0.321
GSGLRTMTTQNQQT	407	1.900	8.939	0.838	0.594
GSGQVGRLLQNQQT	408	1.899	1.762	0.773	0.748
GSGQLSHQSVNQQT	409	1.898	4.032	0.720	0.695
GSGDRYQTLQNQQT	410	1.897	1.075	0.645	0.318
GSGQNQQLKSSAQT	411	1.891	1.197	0.908	0.716
GSGQNQYSIPVAQT	412	1.891	1.194	0.511	0.297
GSGERLHLTQNQQT	413	1.885	1.456	0.387	0.245
GSGHNQQVRTAPNT	414	1.885	1.022	1.006	0.580
GGLSHVMSGQNQQT	415	1.875	0.403	0.885	0.378
GSGQSHRDVLNQQT	416	1.872	15.082	0.138	0.280
GSGQNLAGRMDQQT	417	1.864	0.085	0.362	0.295
GSGQNQQTNRGNPM	418	1.860	3.402	1.349	1.098
GSGQSYQRDHNQQT	419	1.859	8.013	0.779	0.323
GSLLSAGMGQNQHT	420	1.856	6.168	0.589	0.342
GSGQNQQTAIYRNI	421	1.854	2.207	0.818	1.437
GSGQNQQTSGTTNC	422	1.854	8.161	1.040	0.806
GMTSHSVSGQNQQT	423	1.850	2.732	0.220	0.154
GSSQSTGYQPNQQT	424	1.847	3.388	0.522	0.577
GSLKPTTLGQNQQT	425	1.840	0.476	0.175	0.220
GRMFSLGSGQNQQT	426	1.836	8.429	1.630	1.174
GSGQNQQTALGVKC	427	1.835	1.343	1.378	1.014
GAMVSHSSGQNQQT	428	1.833	8.999	0.739	0.868
GSGQNQQRNSDSVT	429	1.829	0.000	1.238	0.842
GSGQSMTLHLNQQT	430	1.827	0.991	0.721	0.248
GSGQVHQAEVNQQT	431	1.825	0.152	0.436	0.287
GSGQNQSQNHLQQT	432	1.825	0.600	1.063	0.772
GSLLTTASGQNQQT	433	1.822	0.780	0.938	0.635
GSGLIRTAAQNQQT	434	1.822	8.339	0.808	0.998
GSGQNQQTVSRQST	435	1.820	0.472	1.330	0.796
GSGQYANHGINQQT	436	1.820	5.717	0.906	0.701
GSRSTGPSGQNQQT	437	1.819	2.479	0.440	0.466
GRGVQQKLQQNQQT	438	1.817	0.000	1.974	0.823
GSGQNQQVHLSTGT	439	1.811	0.266	1.011	0.455
GSGQNQQLSAKSST	440	1.809	1.567	1.224	1.115
GSGYKAARPQNQQT	441	1.803	0.000	1.418	0.337
GSAGISPSGQNQQT	442	1.797	1.812	0.784	0.622
GSGQNRAHAFLQQT	443	1.795	0.000	1.200	1.271
GSGLSGITMQNQQT	444	1.792	14.796	0.862	0.496
GPGSAHSSGQNQQT	445	1.785	4.392	1.099	0.872
GSSHTQALGQNQQT	446	1.784	0.143	0.882	0.874
GSGVHGVSSQNQQT	447	1.781	4.519	1.504	0.951
GSSGRDMGGQNQQT	448	1.778	2.177	1.052	0.595
GERAFPTSGQNQQT	449	1.775	6.515	0.972	0.362
GGRIVSLSGQNQQT	450	1.766	4.936	1.161	0.847
GSGQNSYSHTSQQT	451	1.765	2.262	1.130	0.658
GLGYPGSSGQNQQT	452	1.763	7.090	0.929	0.577
GSGPQSHTGQNQQT	453	1.757	9.490	0.958	0.447
GSGQNQQLSRDAST	454	1.754	3.716	1.877	0.611
GSGQILHSVPNQQT	455	1.752	1.316	0.398	0.240
GSGFHTDSRQNQQT	456	1.748	4.384	7.344	0.575
GSGQSHSLATNQQT	457	1.745	2.711	1.021	0.343
GSGQNQQTLSKPWT	458	1.743	0.253	0.845	0.733

GSGHAAISQQNQQT	459	1.742	2.373	1.211	0.520
GSGQNQQQIGGNST	460	1.741	6.169	0.877	0.576
GGGPMAGSGQNQQT	461	1.735	2.815	1.049	0.372
GMRMEYQSGQNQQT	462	1.729	3.695	0.644	0.632
GSGQNQQGTLLHQT	463	1.728	2.065	1.347	1.303
GSGQNQRSSGGVQT	464	1.723	2.056	1.805	1.165
GSGQNQRGALATQT	465	1.722	1.117	0.899	0.891
GSGTVHAATQNQQT	466	1.721	1.676	0.563	0.476
GSRMTQQFGQNQQT	467	1.720	25.798	1.233	0.976
GSSSPGASGQNQQT	468	1.717	1.244	1.378	0.660
GHPSPHVSGQNQQT	469	1.713	0.416	0.551	0.488
GSGSHHASRQNQQT	470	1.712	0.451	3.073	0.584
GAVGHSYSGQNQQT	471	1.706	0.808	0.306	0.536
GSRSQYDIGQNQQT	472	1.706	0.112	0.528	0.193
GSGQGPQERGNQQT	473	1.702	1.269	0.846	0.313
GSIAHVGTGQNQQT	474	1.696	1.264	0.837	1.045
GSGQNQQKQNHGNT	475	1.695	5.349	1.538	1.340
GSGQNQQALGSQRT	476	1.695	1.934	1.419	0.562
GSGAITHMPQNQQT	477	1.695	1.681	0.647	0.411
GSGQRNPLLLNQQT	478	1.693	0.144	0.662	0.740
GSSGIPVSHQNQQT	479	1.690	3.384	0.820	0.333
GVHSLTPSGQNQQT	480	1.687	4.104	0.475	0.215
GVIVLHGSGQNQQT	481	1.682	14.166	1.074	1.098
GGTRVVDSGQNQQT	482	1.676	9.735	0.676	0.370
GSGGVTYQSQNQQT	483	1.673	7.283	0.649	0.181
GSGQNQAGHGPGQT	484	1.670	2.861	1.040	0.887
GSGQLVTSGPNQQT	485	1.669	5.271	0.964	0.433
GSGIAAQRTQNQQT	486	1.665	2.691	1.062	0.754
GSTPAGVGGQNQQT	487	1.663	2.733	0.593	0.477
GSGQNQQTSTGVHS	488	1.660	8.271	1.039	1.075
GSGQIRQLVDNQQT	489	1.657	5.529	0.314	0.272
GSLIGMQSGQNQQT	490	1.656	6.783	0.797	0.392
GSGQIKGKMDNQQT	491	1.654	2.601	1.065	1.012
GSGSDMSSWQNQQT	492	1.651	0.175	0.281	0.303
GRGQNQQHTGLATT	493	1.650	6.174	1.134	0.691
GSGQNQQTLYSSNT	494	1.642	1.044	0.664	0.368
GSGQTQVLKSNQQT	495	1.640	3.031	1.599	0.975
GSRTLSNVGQNQQT	496	1.640	3.219	0.617	0.542
GSGVQHSLPQNQQT	497	1.639	0.764	0.440	0.387
GNYLHQASGQNQQT	498	1.635	1.454	0.816	0.181
GSGGTSVHQQNQQT	499	1.629	0.000	0.585	0.195
GMDHSRPSGQNQQT	500	1.627	3.976	0.918	0.648
GSGQNQQSMGTFTT	501	1.625	0.000	1.792	0.399
GSGQNQQTPLRPPT	502	1.624	0.352	0.874	0.472
GSGQNQHHSVSQQT	503	1.623	3.700	0.605	0.334
GSGQLRSLSTNQQT	504	1.622	6.855	1.310	0.382
GSGSPRQLSQNQQT	505	1.621	0.873	0.520	0.273
GSGQNQQTTASSHT	506	1.618	7.404	0.745	0.678
GRGQVVSTHQNQQT	507	1.607	3.318	0.931	0.561
GSAQVSMVGQNQQT	508	1.601	1.332	0.500	0.285
GSSTLVTIGKNQQT	509	1.592	4.316	0.917	0.819
GFAHQASSGQNQQT	510	1.587	1.852	1.638	1.080
GSGQPVLSISNQQT	511	1.586	2.695	0.390	0.282

GSGQSHRSELNQQT	512	1.585	11.974	0.668	0.256
GSSVGSPIGQNQQT	513	1.584	3.574	1.059	0.706
GSGMPIRNVQNQQT	514	1.584	0.138	0.684	0.631
GSSTRVDSGQNQQT	515	1.584	2.774	0.704	0.660
GSGQNQQTAMRSTT	516	1.581	2.588	0.656	0.665
GSGQNQQHSSSHLT	517	1.581	2.782	1.091	0.859
GSRNGHAVGQNQQT	518	1.574	2.688	0.434	0.939
GLGAYQSSGQNQQT	519	1.574	0.696	1.407	0.688
GPGLSGHSGQNQQT	520	1.571	1.603	1.154	1.297
GSTGIVSSGQNQQT	521	1.570	0.927	2.141	1.046
GSRTTQVIGQNQQT	522	1.570	1.838	0.773	0.564
GSGLLHRAQQNQQT	523	1.569	0.724	1.583	0.646
GSGQNAQQAAAQQT	524	1.568	4.239	0.937	0.604
GSGQNQQSALRTQT	525	1.568	1.913	1.581	1.421
GSGFLSDTRQNQQT	526	1.566	45.953	0.473	0.575
GSGLLYHDQQNQQT	527	1.565	2.760	0.405	0.107
GSGQNQHYSLHKQT	528	1.563	3.399	1.485	1.273
GSGHSPLPQQNQQT	529	1.562	0.556	0.387	0.247
GNGHSMRPNQNQQT	530	1.560	2.341	0.693	0.376
GSGLKWSTLQNQQT	531	1.556	0.000	1.134	2.442
GSGQMGRQAVNQQT	532	1.554	1.529	0.535	0.411
GSGQNQQTSGVLTL	533	1.553	0.000	1.104	0.782
GSGQNQQALHNPHT	534	1.553	0.664	0.638	0.213
GSGQNQQVIPNSKT	535	1.548	1.036	0.844	0.376
GSPLQDRVGQNQQT	536	1.548	0.753	0.469	0.391
GSGQNQYSSTNPQT	537	1.542	2.251	0.544	0.535
GAMTVTISGQNQQT	538	1.542	6.249	0.443	0.257
GSGQNQQLQTLIRT	539	1.538	1.425	0.813	0.514
GSGLRQTSQQNQQT	540	1.537	2.067	0.978	0.705
GSGQNQQTGLRQQT	541	1.533	2.120	1.217	1.103
GSGQTRQMKDNQQT	542	1.530	11.079	0.841	0.214
GSGQNHGLQSGQQT	543	1.530	4.960	0.938	0.779
GSGQSHRQPENQQT	544	1.529	2.153	0.209	0.159
GSGQDRHIVQNQQT	545	1.527	11.068	0.285	0.162
GSGQNQQLPHSNLT	546	1.521	1.838	0.442	0.283
GSGQLSVPYDNQQT	547	1.521	0.000	0.622	0.111
GSGRNPQTQPLQQT	548	1.519	0.040	0.733	0.573
GSGQPYSTGLNQQT	549	1.519	1.403	0.612	0.376
GSGQNQQTHGGLRD	550	1.519	6.487	1.913	1.298
GAYGMVSSGQNQQT	551	1.518	3.469	0.732	0.773
GSGIQSSYSQNQQT	552	1.517	15.978	1.032	0.684
GPRLSDQSGQNQQT	553	1.511	0.364	0.640	0.579
GSGQNQQTHPSPCT	554	1.510	1.003	1.120	0.546
GSGQSFQMHTNQQT	555	1.504	9.770	0.503	0.325
GSGQNQQTGNPKHT	556	1.504	5.973	1.391	1.139
GFSSAVHSGQNQQT	557	1.502	1.234	0.218	0.210
GSGQNQQTSMSNAT	558	1.501	6.766	1.605	0.745
GSGQDMKQHHNQQT	559	1.501	1.638	0.358	0.239
GLRLSTPSGQNQQT	560	1.498	4.334	0.804	0.522
GSGQNQQTSVYMNT	561	1.498	0.613	0.640	0.983
GSGQNQYSQSSMQT	562	1.494	4.278	0.375	0.309
GSGQNQQSMADHTT	563	1.494	1.728	0.428	0.215
GWERSFVSGQNQQT	564	1.492	0.943	0.490	0.538

GLLAGKSSGQNQQT	565	1.491	2.981	0.999	0.946
GKSFVPQSGQNQQT	566	1.489	2.502	1.798	0.430
GSGQMQSAGSNQQT	567	1.482	0.116	1.034	1.128
GSDQNQRLTSSMQT	568	1.479	0.164	0.875	0.670
GESRAVLSGQNQQT	569	1.476	0.938	0.789	0.368
GSVFGVPSGQNQQT	570	1.474	1.248	0.685	0.213
GSGLPDRNLQNQQT	571	1.471	7.306	1.136	0.611
GSGTHNSAIQNQQT	572	1.469	0.570	0.762	0.574
GSGMIIASMQNQQT	573	1.469	6.722	1.135	0.415
GGITWTDSGQNQQT	574	1.462	4.535	1.472	0.468
GSGQNQQASGRQQT	575	1.458	3.179	0.943	0.991
GSGQNQQPHLKSLT	576	1.457	5.016	1.096	0.740
GPPQHMTSGQNQQT	577	1.457	1.547	0.509	0.677
GSGQNQQASLPSRT	578	1.456	0.389	0.930	0.673
GSGQIVSTQTNQQT	579	1.456	1.103	0.453	0.512
GSGKGHSAGQNQQT	580	1.453	0.936	1.035	1.173
GSGQNTRLQLGQQT	581	1.452	1.747	0.181	0.234
GSVGSRPVGQNQQT	582	1.442	11.363	1.182	0.716
GSSHTLALGQNQQT	583	1.441	7.071	0.851	0.406
GMYEYSQSGQNQQT	584	1.438	0.000	1.410	0.448
GNGQNQQHSILHGT	585	1.435	0.000	0.777	0.415
GSGYNQPHLQNQQT	586	1.435	4.512	0.711	0.395
GPLVNASSGQNQQT	587	1.434	5.239	0.831	0.343
GSGQNQQVLTTART	588	1.434	4.142	1.071	0.948
GSGQNQHSVHNDQT	589	1.428	0.000	0.521	0.515
GAGLIMHSGQNQQT	590	1.425	1.408	0.565	0.511
GMGRHSASGQNQQT	591	1.417	6.500	0.470	0.389
GSHSQSGHGQNQQT	592	1.413	1.240	0.696	0.318
GSSTTIVSGQNHQT	593	1.411	0.000	0.993	0.672
GRHLVTASGQNQQT	594	1.411	2.885	0.648	0.404
GSGQNQQHANLNQT	595	1.410	0.094	0.416	0.544
GSGSTHSKAQNQQT	596	1.410	0.515	0.921	0.801
GSGQNKQMLSGNTT	597	1.410	2.219	1.074	0.404
GSGQVHNPTQNQQT	598	1.410	2.488	1.021	0.542
GSGQNQQIPHVHQT	599	1.409	0.768	0.576	0.218
GSLHAGLSGQNQQT	600	1.408	1.739	1.286	0.936
GPAQHGTSGQNQQT	601	1.407	0.866	1.030	0.615
GEKAVTSSGQNQQT	602	1.402	0.998	0.558	0.327
GSGQNQQTMANGQR	603	1.394	0.216	1.169	1.230
GSGSPHSKDQNQQT	604	1.394	0.000	2.041	4.680
GSFSMGYGGQNQQT	605	1.393	18.476	1.908	1.030
GSGTHLVSLQNQQT	606	1.392	0.000	0.715	1.167
GSGQMQPHVQNQQT	607	1.389	9.381	0.387	0.153
GSGQNQQVAGLNNT	608	1.386	3.218	0.449	0.492
GSSQNQQHDMRLRT	609	1.386	2.645	0.669	0.552
GPASLPISGQNQQT	610	1.386	9.008	0.312	0.155
GSGQNQQPPLATRT	611	1.386	2.295	0.593	0.287
GSSRVPVSGQNQQT	612	1.385	13.191	0.870	0.485
GSGQNQQTNLGHTT	613	1.383	1.523	1.343	1.281
GSGQNQQLVSRVQT	614	1.381	1.195	0.656	0.466
GPNSYPVSGQKQQT	615	1.381	4.040	0.736	0.834
GHAHYQASGQNQQT	616	1.377	7.299	0.803	0.745
GSGQALLSTGNQQT	617	1.377	0.847	0.536	0.370

GSGQLPRQMTNQQT	618	1.376	3.550	0.400	0.562
GSGFPKSTEQNQQT	619	1.376	2.058	0.610	0.194
GSRETSLSGQNQQT	620	1.373	5.193	1.364	0.203
GSGQNQQGTGVSHT	621	1.371	4.295	1.417	0.749
GSRTVPVYGQNQQT	622	1.371	0.363	1.226	0.969
GSNAQSAHGQNQQT	623	1.371	0.888	0.976	0.245
GAFHLAASGQNQQT	624	1.369	18.165	0.994	0.775
GSGQYRSSSDNQQT	625	1.369	6.209	0.681	0.409
GSGQVYISTPNQQT	626	1.367	0.000	0.859	0.282
GSGVSTQLLQNQQT	627	1.367	2.467	0.928	0.509
GSGQLGLSVTNQQT	628	1.364	6.906	1.395	0.376
GSGSNMRLSQNQQT	629	1.363	0.588	0.962	0.730
GSGQNLHSGLPQQT	630	1.363	1.594	1.054	0.592
GSSHTLALGQNKQT	631	1.362	2.160	0.838	0.643
GSGQNQHSLPAHRT	632	1.361	0.700	0.911	0.742
GSGQNQGTVYPNQT	633	1.358	7.648	0.835	0.815
GSGQNQQPSLRQST	634	1.356	2.905	1.315	0.554
GSGQNARLKDNQQT	635	1.354	2.395	0.580	0.938
GHAGSTGSGQNQQT	636	1.352	2.829	1.332	1.233
GSGQALSSSGNQQT	637	1.351	6.860	0.894	0.931
GSGASESHRQNQQT	638	1.350	0.850	0.325	0.313
GVGVITSSGQNQQT	639	1.348	0.918	1.296	0.777
GSLYGQSLGQNQQT	640	1.348	11.248	0.894	0.843
GSGQMSDVHGNQQT	641	1.346	7.172	0.408	0.548
GSGQNQQHSSKATT	642	1.345	12.248	1.350	1.401
GSGQNQQTSVSQQT	643	1.342	1.614	1.030	0.913
GSGQKMWKLDNQQT	644	1.341	0.000	0.990	1.418
GSGQNVSMQVNQQT	645	1.341	0.000	0.357	0.251
GSGQNQRATLSNQT	646	1.339	1.084	0.947	0.723
GSGQASSKSANQQT	647	1.339	1.138	0.500	0.175
GSGKNQTPIPKGQT	648	1.339	5.077	1.306	1.154
GSGQNQQTRQEGST	649	1.339	0.000	0.645	0.718
GASSLATSGQNQQT	650	1.337	0.703	0.423	0.217
GSGQRGSLTENQQT	651	1.337	2.482	0.300	0.567
GSEQTRQRGQNQQT	652	1.333	2.172	0.574	0.815
GSGQNQQTLTASKE	653	1.333	1.152	0.981	1.172
GSGTSGKTGKNQQT	654	1.333	4.033	0.358	0.676
GQLVTFTSGQNQQT	655	1.331	11.282	0.819	0.294
GSGQNQQSANKILT	656	1.331	3.789	0.894	1.236
GSGQNQQHHSSHTT	657	1.328	2.158	0.957	0.452
GSGQNQKGMQPNQT	658	1.326	3.139	0.775	1.059
GSGQLVSGLYNQQT	659	1.325	0.000	0.842	0.733
GSSVGVPSGQNQQT	660	1.322	4.867	0.336	1.157
GSGQNQQWDSRRQT	661	1.321	0.531	1.059	0.825
GSEQTRQSGQNQQT	662	1.321	0.514	0.734	0.900
GSGIGSHIPQNQQT	663	1.319	0.173	0.822	0.597
GSGQNQRLHGVDQT	664	1.318	4.655	0.459	0.341
GEVSRVLSGQNQQT	665	1.318	0.437	1.150	0.440
GSGQNQQKVSPLLT	666	1.314	1.602	0.755	0.806
GSGLALERSQNQQT	667	1.311	0.486	0.618	0.096
GPDRIGSSGQNQQT	668	1.308	0.426	0.654	0.342
GSGQNQDHQNKQQT	669	1.308	1.470	0.510	0.761
GSGQNQQTALYNNT	670	1.307	0.862	0.660	0.726

GSGAVHLTAQNQQT	671	1.306	1.668	0.541	0.466
GSLVSTQSGQNQQT	672	1.305	1.293	1.282	0.650
GSGVSARMVQNQQT	673	1.299	0.624	0.870	0.697
GSGQTRMPLANQQT	674	1.296	0.790	0.447	0.273
GSGISSRNMQNQQT	675	1.291	6.328	1.671	0.560
GSGEKVHSGQNQQT	676	1.289	0.062	0.862	0.671
GSGQNQQKLSSMST	677	1.286	1.586	1.160	1.052
GSGQNQQTGQHMRV	678	1.286	4.161	1.839	1.635
GSGMIHTTAQNQQT	679	1.285	0.105	0.678	0.276
GSGQNWPALKGQQT	680	1.284	2.031	1.101	1.222
GASHMSISGQNQQT	681	1.284	0.462	0.404	0.374
GSDQNQQLGYSKQT	682	1.283	0.000	0.853	0.660
GIPSIRESGQNQQT	683	1.282	0.166	0.484	0.254
GSGIPSVKFQNQQT	684	1.281	0.061	0.364	0.561
GSGQNQQTSVSQNV	685	1.281	0.750	0.788	0.715
GSGQNQQIGESRMT	686	1.279	0.103	0.890	0.453
GSGSSSMSFQNQQT	687	1.279	0.540	0.466	0.095
GSGQKQERAVSKQT	688	1.277	0.000	1.174	0.732
GCTTRLNSGQNQQT	689	1.276	0.000	0.184	0.618
GSGQNQQIISTKIT	690	1.275	0.000	0.951	0.710
GSGQNQQKSLNGNT	691	1.275	8.573	0.586	0.851
GSGIPAPRLQNQQT	692	1.273	4.162	0.583	0.396
GSGQIRESMGNQQT	693	1.270	1.676	0.833	0.523
GSGQNSGVHFNQQT	694	1.268	0.587	0.871	0.377
GSGQNIGHSLPQQT	695	1.264	6.183	0.740	0.478
GSGERSISVQNQQT	696	1.264	1.619	0.598	0.173
GSGLKPNVLQNQQT	697	1.263	0.975	0.701	0.268
GSGQVAYAQGNQQT	698	1.259	1.309	0.734	0.313
GSGQSSYGSGNQQT	699	1.257	1.686	1.161	0.456
GSGQNQAMTHGDQT	700	1.257	1.878	0.357	0.259
GSGQNQALVSMGQT	701	1.255	1.876	0.987	0.560
GSGQNPSFMRGQQT	702	1.252	1.454	1.293	1.094
GSGQNQQSHLRTNT	703	1.251	4.583	1.022	0.718
GYTRLETSGQNQQT	704	1.250	1.323	0.841	0.297
GSGQSYDMRGNQQT	705	1.248	0.567	0.588	0.368
GSRTTQDIGQNQQT	706	1.247	0.000	0.685	0.280
GSGHPYKAAQNQQT	707	1.246	0.000	0.872	0.507
GRLSNAHGGQNQQT	708	1.245	0.839	1.036	0.725
GSGQNQRAVLNDQT	709	1.242	3.023	0.556	0.259
GGSHTYGGGQNQQT	710	1.241	13.065	0.982	0.730
GSSVNSMIGQNQQT	711	1.239	0.000	0.976	0.580
GNSSMMGSGQNQQT	712	1.239	3.856	0.656	0.364
GNRDRPSSGQNQQT	713	1.239	3.947	0.298	0.178
GSGNMHASRQNQQT	714	1.238	3.878	0.782	0.687
GFIFPKVSGQNQQT	715	1.237	0.000	1.764	0.692
GSGQNQQLKNSTST	716	1.235	1.703	1.063	0.538
GSGQNQQSQYMPRT	717	1.234	0.401	0.549	0.520
GSGQRMADIGNQQT	718	1.233	2.539	0.352	0.427
GSGQNQSHYPSQQT	719	1.228	4.315	0.644	0.402
GSDGKMHRGQNQQT	720	1.227	0.000	1.826	0.776
GSGSVGFIGQNQQT	721	1.227	8.261	0.689	0.445
GLHGMTLSGQNQQT	722	1.226	3.552	0.470	0.338

GSDQSKRGDSNQQT	723	1.225	0.639	0.479	0.267
GSLFLATGGQNQQT	724	1.220	0.000	0.775	0.485
GSGQNQQPSAFSKT	725	1.220	4.906	1.309	0.754
GSGQLPQSGLNQQT	726	1.218	1.504	0.641	0.318
GSGSKQNALQNQQT	727	1.216	2.010	0.941	0.594
GSGQRRELSQNQQT	728	1.215	1.791	0.622	0.396
GSGQREPKASNQQT	729	1.214	2.793	0.399	0.520
GSGQNQQHPSTQQT	730	1.205	1.552	1.017	0.680
GSQSTLGLGQNQQT	731	1.204	3.246	0.594	0.400
GSGQNQQMPGLSST	732	1.204	1.887	0.234	0.181
GSGQNQQTVGGKNL	733	1.203	0.128	0.777	1.051
GSSREFHSGQNQQT	734	1.203	1.591	0.688	0.474
GSGQNQQTVPSNLV	735	1.201	0.791	0.434	0.281
GSGQNAYSSQAQQT	736	1.201	12.096	0.629	0.216
GSGQNKDHSTRRQT	737	1.197	0.000	0.384	0.477
GQLGSVGSGQDQQT	738	1.196	0.000	1.020	0.437
GSGQHAAPGHNQQT	739	1.195	5.999	0.600	0.199
GSGQNQQTSQSPPT	740	1.194	1.208	0.851	0.478
GSGNYRDHEQNQQT	741	1.193	7.389	0.287	0.222
GSGQHSNQHVNQQT	742	1.192	1.453	0.955	0.558
GSGQTARNGINQQT	743	1.192	2.030	1.002	0.472
GSGQNQQHYGSQGT	744	1.189	0.453	1.345	0.379
GSGSPQASRQNQQT	745	1.189	6.782	0.923	0.542
GSGFSHSMGKNQQT	746	1.188	9.809	1.381	0.611
GSGQSHSLETNQQT	747	1.188	1.319	0.520	0.363
GTEQTRQSGQNQQT	748	1.188	0.132	0.756	0.756
GSGRHLASVQNQQT	749	1.187	1.024	0.654	0.606
GLGSKNHSGQNQQT	750	1.187	5.046	0.825	0.224
GSGQNQQTSHFPSA	751	1.185	0.325	0.969	0.907
GSGQLSGTPQNQQT	752	1.185	1.382	1.025	0.643
GSGQNQQAPHKKET	753	1.180	0.598	0.994	0.689
GSGQNQQTLRGSLE	754	1.179	1.812	0.853	0.354
GSIAMTSHGQNQQT	755	1.178	1.435	0.551	0.438
GSPGVSPSGQNQQT	756	1.178	3.006	0.853	1.160
GSGQNQQTGSSSRV	757	1.176	0.580	0.995	1.128
GSGQHLPLLGNQQT	758	1.175	1.739	0.519	0.347
GSDHSHRGGQNQQT	759	1.174	0.504	0.818	0.331
GSGIVTKLGQNQQT	760	1.174	10.571	0.599	0.242
GSGQDVTKTGNQQT	761	1.173	4.523	0.531	0.035
GSGQNQQSHGRIGT	762	1.173	5.117	0.607	0.455
GSGQNQQINHRSPT	763	1.173	0.748	0.259	0.220
GSGDDSRVGQNQQT	764	1.172	0.191	0.466	0.156
GSGQSTLKRINQQT	765	1.168	13.442	0.534	1.184
GSGSQHSKAQNQQT	766	1.168	0.312	0.638	0.916
GSGQNQQHASSNNT	767	1.166	7.155	0.789	0.896
GSRTYQVSGQNQQT	768	1.164	1.853	0.638	0.641
GSGQNQGLLSSPQT	769	1.164	0.000	0.707	0.417
GSGGGLQHNQNQQT	770	1.163	4.098	1.137	0.778
GSGQNQQTTAATRM	771	1.163	3.925	0.947	1.005
GSGQNQRASILVQT	772	1.162	3.632	0.531	0.569
GSGQNLGLLGAQQT	773	1.161	1.458	0.524	0.226
GSLDLGRSGQNQQT	774	1.160	3.283	1.002	0.505
GNSQVKVSGQNQQT	775	1.158	4.930	1.422	0.728

GSSGSHQYGQNQQT	776	1.155	0.000	1.129	0.794
GSGQNQQQRDGTLT	777	1.152	0.387	0.760	0.730
GRGQHVSVANNQQT	778	1.152	1.896	1.032	0.589
GDSSSRISGQNQQT	779	1.151	3.787	0.916	0.348
GSGQNQQHSLSSQT	780	1.150	3.844	0.700	0.730
GSLMDVHRGQNQQT	781	1.150	0.387	1.009	0.238
GSIQYQSSGQNQQT	782	1.147	2.601	1.074	1.191
GLGSKNPSGQNQQT	783	1.147	1.629	1.184	0.424
GSGQLVLTLQNQQT	784	1.143	0.000	0.336	0.336
GSGQNQQTSQPLPG	785	1.141	0.080	0.748	0.530
GSGQNQQNLGKLNT	786	1.141	0.000	0.919	0.687
GTTAHQPSGQNQQT	787	1.138	0.211	0.726	0.275
GSGQNRAQIGTQQT	788	1.138	0.469	0.776	0.654
GSGQYVHVSSNQQT	789	1.137	1.803	0.739	0.366
GSGQNQQTAHAFNI	790	1.132	3.404	0.699	0.729
GSGQNQRTMVATQT	791	1.130	1.122	0.649	0.554
GSGQNPIRGAMQQT	792	1.126	1.327	1.296	0.427
GSGYVITGSQNQQT	793	1.125	6.271	0.971	0.248
GRGPKQSNIQNQQT	794	1.125	0.737	0.771	2.490
GSGQNQQTMLGKPC	795	1.125	0.047	1.090	0.992
GSGQNQQVGSTVRT	796	1.124	2.040	0.918	0.614
GNVTTQKSGQNQQT	797	1.122	2.546	1.215	0.922
GSGNPVSHLQNQQT	798	1.121	1.037	0.583	0.310
GSLSHMESGQNQQT	799	1.120	0.829	0.489	0.265
GRAPTNLSGQNQQT	800	1.118	0.687	0.757	0.169
GSGQNQQTVMTARA	801	1.117	1.535	0.995	0.843
GSGMPASRLQNQQT	802	1.117	1.689	0.790	0.372
GVVRNHQSGQNQQT	803	1.116	5.801	0.899	0.868
GSGQNQHSVQVRQT	804	1.116	1.909	0.782	0.916
GSGQNTGHLTMQQT	805	1.114	0.078	1.026	0.595
GSGQNQQYAGKILT	806	1.112	0.300	1.078	0.431
GSGNPHVRNQNQQT	807	1.112	0.873	0.732	0.755
GSGQNGGSSNRQQT	808	1.109	2.594	1.255	0.844
GSGQRLSQGVNHQT	809	1.108	3.394	0.931	1.141
GSGQNAHAKEGQQT	810	1.108	0.000	0.875	1.179
GSSPAPNSGQNQQT	811	1.106	2.229	0.719	0.368
GLAHKTSSGQNQQT	812	1.106	0.915	0.427	0.690
GSGQNQQTPGAHKT	813	1.105	3.827	0.957	0.277
GSGQNQQSLSGSFT	814	1.105	0.735	0.745	0.883
GSGQNQQSTGTSRT	815	1.103	4.054	1.209	0.935
GSGQNQQTVQSNLV	816	1.103	2.350	0.577	0.698
GSGQNQQLGSRQCT	817	1.102	0.183	0.987	0.407
GSGQNQYLRLELQT	818	1.101	0.000	0.416	0.839
GSGQNQQTSPRLQT	819	1.100	0.795	1.156	1.091
GSGQNQQTTSSNMT	820	1.099	0.569	0.638	0.698
GTASTYNSGQNQQT	821	1.099	2.560	0.250	0.625
GSGQNQQTMPQHKI	822	1.097	2.394	0.479	0.197
GSGQSHLHTGNQQT	823	1.096	2.584	0.721	0.295
GVKGVGHSGQNQQT	824	1.096	2.485	0.994	0.783
GSGKVTKQSQNQQT	825	1.095	0.000	0.928	1.035
GSGQNQQTALEKSL	826	1.092	0.000	0.625	0.702
GSGYKDTYGQNQQT	827	1.091	0.854	0.717	0.448
GSGQNQQSGTFLST	828	1.090	5.673	1.021	0.742

GSGQNTGQHMMQQT	829	1.090	1.058	1.147	0.917
GSGKNQQRPGLDQT	830	1.089	1.557	0.583	0.385
GSGQSREISLNQQT	831	1.088	6.954	0.594	0.282
GTPTSPSSGQNQQT	832	1.086	4.558	0.833	0.662
GKPAGGLSGQNQQT	833	1.085	2.805	0.708	0.739
GSGQNHRSADMQQT	834	1.084	12.001	0.417	0.212
GSGQNQQTLPSLSL	835	1.084	1.758	0.527	0.175
GSPYMGATGQNQQT	836	1.083	5.364	0.918	0.254
GSGHAKAVGQNQQT	837	1.081	4.357	0.703	0.824
GHMKGVTSGQNQQT	838	1.081	2.814	0.807	0.413
GSGQNQKILTLDQT	839	1.080	0.371	0.291	0.314
GSGQNQQTKVGHSA	840	1.079	1.256	0.669	1.019
GIARTTISGQNQQT	841	1.078	1.783	0.819	0.330
GSGQNQQTSVGFRT	842	1.077	3.737	0.648	0.534
GSGQNQQTMIANIR	843	1.076	0.000	0.379	0.458
GDMTRSSSGQNQQT	844	1.075	0.802	1.145	1.038
GSGHMSDLRQNQQT	845	1.073	4.291	0.555	0.328
GRGAVMASGQNQQT	846	1.072	0.923	0.783	0.605
GSGQNQQLSGKSVT	847	1.070	1.524	1.276	0.930
GSHTLVVSGQNQQT	848	1.069	1.535	0.671	0.748
GSGPWSAGLQNQQT	849	1.067	0.947	0.700	0.539
GSGQHSPHALNQQT	850	1.064	1.412	0.885	0.573
GSGQNQQPNSGSMT	851	1.064	0.925	0.588	0.339
GSGLAHLGGQNQQT	852	1.064	2.191	0.749	0.794
GSSVRYEPKQNQQT	853	1.063	1.564	0.450	0.501
GSGQNQQARPLELT	854	1.061	0.059	0.389	0.252
GSGQPRSTGINQQT	855	1.061	0.693	0.650	0.542
GSGQNQANWVKVQT	856	1.059	0.126	0.683	0.532
GSGHLFQSGQNQQT	857	1.057	0.615	0.751	0.386
GSGQNRGISISQQT	858	1.057	2.166	0.686	0.566
GSGTHYDNRQNQQT	859	1.054	0.072	0.612	0.486
GSGQNQQTSTTPLP	860	1.052	2.823	0.828	0.741
GSGQVHASQVNQKT	861	1.049	0.503	0.855	0.767
GSSGHRESGQNQQT	862	1.048	4.398	0.641	0.691
GLSAEKSSGQNQQT	863	1.047	7.203	0.629	0.303
GSGQEHRSLANQQT	864	1.046	0.000	0.507	0.344
GSGQTVVRIANQQT	865	1.046	4.156	0.661	0.390
GSGQNVSSVHRQQT	866	1.045	0.712	0.383	0.271
GSGASRMSIQNQQT	867	1.045	0.111	0.801	0.417
GVAFIGSSGQNQQT	868	1.043	0.000	0.744	0.648
GSGQNQQTVPTRQT	869	1.040	1.207	0.629	0.138
GSGQAAKSSQNQQT	870	1.036	0.681	0.778	0.737
GSGQNQQVAIRTST	871	1.035	2.447	0.963	0.370
GSVHMQNAGQNQQT	872	1.034	3.608	1.004	0.625
GSGMRQAGVQNQQT	873	1.032	0.811	0.736	0.775
GSGQNQQVGGKTVT	874	1.032	6.195	1.094	0.821
GVHDMRVSGQNQQT	875	1.032	8.083	1.171	0.818
GSGQHVSVANNQQT	876	1.029	5.734	0.974	0.577
GSAAMSVRGQNQQT	877	1.029	2.386	0.202	0.287
GVSRGGPSGQNQQT	878	1.028	1.611	0.750	0.591
GSGQMVHTIGNQQT	879	1.026	1.328	0.406	0.430
GRGGSMAETQNQQT	880	1.024	2.853	0.799	0.669
GSGHTNPTRQNQQT	881	1.021	0.688	0.726	0.807

GSGEAARYEQNQQT	882	1.020	0.000	0.107	0.125
GSGQNERHLVLQQT	883	1.019	5.354	0.416	0.150
GSGQNQQSKQQVLT	884	1.019	1.494	1.428	1.256
GSGQARAHRGNQQT	885	1.017	0.000	0.254	0.386
GSGQNQQPLDTSRT	886	1.015	0.775	0.491	0.376
GSGQNQQLANMVTT	887	1.014	1.739	1.253	0.987
GSGQMKDLHRNQQT	888	1.014	1.068	0.587	0.506
GSGQNQHLSSFVQT	889	1.013	0.110	1.090	0.364
GSGQNQQPSSRVTT	890	1.012	2.179	0.784	0.504
GSGQNQQLAITLGT	891	1.011	0.000	0.877	0.143
GSGQNQQTVGNPAT	892	1.008	3.014	0.856	0.395
GSGQNQGRAHPMQT	893	1.007	2.364	0.684	0.453
GSGQLIASVVNQQT	894	1.005	0.086	0.197	0.359
GSSVRSLVGQNQQT	895	1.004	3.840	0.412	0.608
GGAGSAHSGQNQQT	896	1.003	6.108	0.474	1.092
GSDQNQQTMSSTRT	897	1.003	2.428	1.306	0.835
GSGQNQQMAGAFRT	898	1.003	1.784	1.307	0.762
GSLGNLQRGQNQQT	899	1.003	0.895	0.947	0.385
GSGPSISHGQNQQT	900	1.000	0.000	0.614	0.665
GSGQNQQT	6406	1.000	1.000	1.000	1.000
GSGQNQQSSFNVQT	901	0.998	0.000	1.307	0.675
GSGQNQQTGQATHN	902	0.996	2.199	0.877	0.527

[0500] A second cross-species evolution screen was performed using an AAV capsid variant library with a modification in loop IV introduced as described above and passaging it once through NHPs (passage 1) and then subsequently injected it into two different strains of mice (passage 2), C57Bl/6 and BALB/c. The fold-enrichment for each variant in the brain of each mouse species was calculated by systematic NGS enrichment analysis following RNA recovery and RT-PCR amplification. The fold enrichment values in the second passage in mice were compared to those fold enrichment values from the second pass that was performed in NHPs as described above. As shown in **Table 10**, when comparing the second pass fold enrichment values in the mice versus NHPs, 12 variants were identified that had a fold-enrichment value greater than 10 in all three animal groups. Further, 10 of these 12 variants comprised the SPH motif and a positive residue in one of the next three subsequent residues (**Table 10**).

Table 10. NGS fold-enrichment of AAV capsid variants from a second passage (P2) in NHPs or mice (C57Bl/6 or BALB/c) following a first passage in NHPs

Peptide Sequence	SEQ ID NO:	Fold enrichment over AAV9 in NHP P2	Fold enrichment over AAV9 in BALB/c P2	Fold enrichment over AAV9 in C57Bl/6 P2
VSGSPHSKAQNQQT	903	99.76	92.99	34.29
CSGSPHSKAQNQQT	904	85.1	66.74	22.19
GSGSPHSKAQNQQT	200	56.33	44.58	14.48
GSGSPHRKAQNQQT	905	46.39	42.47	14.11
GRGSPHSKAQNQQT	906	43.68	59.65	28.13
GHDSPHKSGQNQQT	201	33.96	59.14	27.15
GSGSPHSKAKNQQT	208	31.27	41.51	14

GSGSPHSKAQNKQT	907	29.52	44.1	13.69
GSGSPHSKAQTQQT	908	24.27	41.75	18
GQDSPHKSGQNQQT	204	22.7	32.37	16.02
GSGSTHASRQNQQT	909	11.04	23.71	10.67
GHDSQHKSGQNQQT	404	10.36	21.3	13.55

[0501] Following the second passage in mice, a synthetic library was generated using those variants that demonstrated a fold-change in enrichment relative to wild-type AAV9 that was above 10 in the brain of either strain of mice, as measured by systematic NGS enrichment analysis following RNA recovery and RT-PCR amplification. There were approximately 500 variants in this synthetic library. This synthetic library was then injected back into both strains of mice (C57Bl/6 and BALB/c; passage 3). RNA was recovered from the mouse brains, RT-PCR amplification was performed, and fold-enrichment relative to wild-type AAV9 was calculated by NGS analysis, which is provided in **Table 11**. As shown in **Table 11**, the variants with the greatest fold-enrichment in the brain in each strain, were highly correlated across strains (R^2=0.8458).

Table 11. NGS fold-enrichment of AAV capsid variants in the brain from a third passage (P3) in mice (C57Bl/6 or BALB/c) following a first and second passage in mice

Peptide Sequence	SEQ ID NO:	Fold enrichment over AAV9 in BALB/c	Fold enrichment over AAV9 in C57Bl/6	Average
GSGSPHKYGQNQQT	910	150.445	103.488	126.966
GSGSPHKFGQNQQT	911	73.364	60.304	66.834
GHDSPHKSGQNQQT	201	82.460	51.125	66.792
GSGSPHSKAQNQQT	200	60.312	65.853	63.083
VSGSPHKFGQNQQT	912	60.186	59.142	59.664
GSGSPHSKAQNHQT	913	63.486	51.647	57.566
VSGSPHSKAQNQQT	903	73.555	37.429	55.492
GQDSPHKSGQNQQT	204	63.898	43.752	53.825
GSGSPHSKAQHQQT	914	45.309	45.600	45.454
GSGSPHKTYQNQQT	915	50.283	35.460	42.871
GSGSPHSKAQTQQT	908	43.120	39.098	41.109
VSGSPHASRQNQQT	916	46.572	32.480	39.526
GSGSPHSKAQNKQT	907	39.848	35.596	37.722
GSGSPHKFGKNQQT	917	31.948	34.899	33.423
GSGSPHASRQNQHT	918	28.145	30.928	29.537
GSHSPHKSGQNQQT	919	22.948	35.412	29.180
GSGQNQQRRMSPST	920	4.576	53.520	29.048
GSGSPHASRQNQQT	205	28.866	29.139	29.003
GSGSPHSKPQNQQT	921	26.958	28.599	27.779
GSGSPHKFGQKQQT	922	39.597	14.927	27.262
VSGSPHGARQNQQT	923	30.985	22.634	26.810
GSGSPHSKAQKQQT	924	25.052	27.459	26.256
GHSSPHRSGQNQQT	212	16.982	35.081	26.032
GSGSPHSKAKNQQT	208	21.069	25.711	23.390
GSHSPHKRGQNQQT	925	24.054	20.262	22.158
GRGSPHSKAQNQQT	906	20.939	22.720	21.830

GQSSPHRSGQNQQT	229	9.916	26.608	18.262
GSGQNRQRLKGLET	926	3.937	31.022	17.480
GSGSPHKLGQNQQT	927	18.905	14.732	16.818
GSGSPHKTSKNQQT	928	14.654	17.606	16.130
GSGSPHKIGQNQQT	929	16.999	14.794	15.897
GSGSPHKKNQNQQT	209	25.633	5.605	15.619
GSGSPHASRQNKQT	206	10.738	20.347	15.542
GSGSPHTRGQNQQT	214	16.899	13.869	15.384
GSGQDSPHVRNQQT	930	15.340	14.646	14.993
GSGSPHKTSQNQQT	931	20.428	8.818	14.623
GSGSPHASRKNQQT	932	13.799	12.749	13.274
GSGSPHASRQKQQT	211	13.624	11.188	12.406
GSHSPHKSGQKQQT	933	6.700	17.736	12.218
GSGSPHKTSQKQQT	934	12.621	11.720	12.170
GSGSPHVRGQNKQT	935	13.174	11.017	12.095
GSGSPHKTTQNQQT	936	9.722	13.381	11.552
CSGSPHSKAQNQQT	904	11.772	9.447	10.610
GSGPVRALRQNQQT	937	3.369	17.431	10.400
GSGSPHVRGQKQQT	938	7.573	12.498	10.036
GSGSPHRKAQNQQT	905	12.308	7.349	9.828
GRGSPHASRQNQQT	318	11.903	6.780	9.342
CSGSPHKTSQNQQT	939	11.167	6.631	8.899
CSHSPHKSGQNQQT	940	11.356	6.304	8.830
GSGSPHSKDQNQQT	604	3.492	10.236	6.864

[0502] Taken together, these results demonstrate that after 3 rounds of screening of this AAV9 variant library with loop IV modifications in NHP and mice, many AAV capsid variants outperformed the wild-type AAV9, for example, in penetration of the blood brain barrier (BBB) and spinal cord expression. These capsid variants were able to cross-species, evidenced by expression and tropism in the NHP brain/spinal cord as well as in the brain of two different mouse species.

Example 2. Individual Capsid Characterization in Mice

[0503] The goal of these experiments was to determine the transduction level, tropism, ability to cross the blood brain barrier, and overall spatial distribution in the central nervous system (CNS) of 2 capsid variants selected from the study described in Example 1 relative to AAV9 following intravenous injection in mice. The 2 capsid variants were TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2), as outlined in Table 3 above. The amino acid and DNA sequences of TTM-001 and TTM-002 are provided, e.g., in Tables 4 and 5, respectively.

[0504] AAV particles were generated with each of these capsid variants encapsulating a luciferase-EGFP transgene driven by a CMV/chicken beta actin promoter in a single stranded viral genome. Each capsid variant and AAV9 control were tested by intravenously administering by tail vein injection, the AAV particle formulation at 5e11 VG/dose (2.5E13 vg/kg) to three female BALB/c

mice. The in-life period was 28 days and then various CNS and peripheral tissues were collected for measuring transgene mRNA, transgene protein, and viral DNA (biodistribution).

[0505] At 28 days post-injection of the AAV particles encapsulated in the TTM-001 capsid variant (AAV_TTM-001), mice were injected with luciferin and their brains were harvested for IVIS imaging. Robust luciferase signal was observed in mice injected with AAV particles encapsulated in the TTM-001 capsid variant, and this was greatly increased relative to AAV particles encapsulated in the wild-type AAV9 control capsid.

[0506] The brains isolated from mice injected with the AAV particles encapsulated in the TTM-001 capsid variant (AAV_TTM-001) or the TTM-002 capsid variant (AAV_TTM-002) were assayed by qPCR for the presence of transgene RNA as a measure of transgene expression, and the presence of viral DNA as a measure of viral genome levels. Data were provided as fold over AAV9 (**Table 12**). As shown in **Table 12**, when compared to the wild-type AAV9 capsid control, TTM-001 and TTM-002 demonstrated a 30-fold and 66-fold increase, respectively, in transgene mRNA levels and expression in the brain, indicative of enhanced payload delivery. This correlated with a 32-fold (TTM-001) and 47-fold (TTM-002) increase, respectively, in viral genome (DNA) concentrations in the brain relative to the AAV9 capsid control, which is indicative of enhanced CNS tropism and transduction (**Table 12**).

Table 12. Transgene mRNA and viral genome levels (DNA) in mice relative to the AAV9 control

Measure	Tissue	AAV9	TTM-001	TTM-002
mRNA (transgene expression)	Brain	1.0	30.4503	66.2161
DNA (viral genome quantification)	Brain	1.0	32.0315	47.2810
mRNA (transgene expression)	Liver	1.0	1.2356	0.2016
DNA (viral genome quantification)	Liver	1.0	0.4802	0.0277

[0507] The brain tissues and spinal cords of the mice were also subjected to anti-GFP immunohistochemistry staining to evaluate overall CNS tropism and biodistribution. Immunohistochemical staining correlated with the qPCR analysis, as TTM-001 and TTM-002 showed significantly stronger staining and payload expression in the brain and spinal cord, as compared to the AAV9 control. More specifically, TTM-001 and TTM-02 demonstrated localization and strong payload expression and transduction in the mid-brain region, with increased staining observed in the hippocampus and thalamus, as well as in the brain stem, compared to AAV9. Less staining was observed in the cortical regions of the brain compared to the midbrain. However, staining in these cortical regions was stronger for TTM-001 and TTM-002 compared to the AAV9 control. It also appeared that the TTM-001 and TTM-002 capsid variants were able to transduce non-neuronal cells, including glial cells and oligodendrocytes. With respect to the spinal cord, staining and payload expression for TTM-01 and TTM-002 were localized to the ventral horns of the grey matter.

[0508] Peripheral tissues were also isolated from the mice intravenously injected with the AAV particles encapsulated in the TTM-001 capsid variant or the TTM-002 capsid variant for analysis by

qPCR and/or GFP immunohistochemical staining. Transgene mRNA levels and viral genome DNA levels were quantified in the liver by qPCR and the fold over AAV9 was calculated for each capsid variant (**Table 12**). TTM-001 resulted in similar levels of payload expression (mRNA levels) as compared to wild-type AAV9, but only half as much viral genome DNA was quantified in the liver compared to AAV9. TTM-002 demonstrated greatly reduced mRNA and viral genome DNA levels in the liver compared to AAV9. GFP immunohistochemical staining of the spleen, heart, skeletal muscle, kidneys, and lungs of mice injected with AAV particles encapsulated in the TTM-001 capsid variant or the TTM-002 capsid variant showed similar levels of payload expression as compared to those mice injected with AAV particles encapsulated in the wild-type AAV9 control capsid.

[0509] Taken together, these data demonstrate that TTM-001 and TTM-002 are enhanced CNS tropic capsids in mice that can infect non-neuronal cells. Additionally, these capsid variants were able to successfully penetrate the blood brain barrier following intravenous injection.

Example 3. Maturation of TTM-001 and TTM-002 Capsid in Mice

[0510] This Example describes maturation of the TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2) capsid variants to further enhance their transduction and biodistribution in the central nervous system and evolve the AAV capsid variants to provide further cross-species compatibility. Two approaches were used to mature the TTM-001 and TTM-002 capsid sequences in order to randomize and mutate within and around the peptide insert comprised within loop IV of the capsid variant. As many of the AAV capsid variants that demonstrated the greatest fold-enrichment in the NHP brain relative wild-type AAV9 comprised an SPH motif in the same position (e.g., immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138) (see **Example 1**), the SPH motif was not mutated in either approach to mature the TTM-001 and TTM-002 capsid variants. In the first maturation approach, sets of three contiguous amino acids were randomized across the mutagenesis region in the TTM-001 and TTM-002 sequences, which spanned from position 450 to position 466, numbered according to SEQ ID NO: 981 and 982. In the second maturation approach, mutagenic primers were used to introduce point mutations at a low frequency, scattered across the mutagenesis region in the TTM-001 and TTM-002 sequences ranging from position 449 to position 466, numbered according to SEQ ID NO: 981 and 982. AAV capsid variants arising from each maturation approach for TTM-001 were pooled together and AAV capsid variants arising from each maturation approach for TTM-002 were also pooled together, for subsequent testing and characterization in mice.

[0511] The library of pooled matured AAV capsid variants generated from TTM-001 or library of pooled matured AAV capsid variants generated from the TTM-002 matured AAV capsid variant each were intravenously injected into the tail vein of three female CD-1 Outbred mice (Charles River) at a dose of 1.0×10^{12} VG/dose. After 14-days in life, the brains of the mice were isolated and RNA was

extracted. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate the fold enrichment ratio relative to the corresponding TTM-001 or TTM-002 control, and the peptides comprised within the variants were identified. The data for the TTM-001 matured capsid variants is provided in **Table 13** and the data for the TTM-002 matured capsid variants is provided in **Table 14**.

[0512] As shown in **Table 13**, approximately 714 TTM-001 matured capsid variants demonstrated at least a 2-fold increase in expression relative to the non-matured TTM-001 control, and several variants demonstrated greater than a four-fold enrichment relative to the non-matured TTM-001 control. Also, across the peptides comprised within the TTM-001 matured capsid variants with the greatest fold-enrichment relative to the non-matured TTM-001 capsid in the brain, it was observed that the modifications in the variant sequences appeared in the region C-terminal to the SPH motif present within the capsid variant. This indicates that modifications that appeared to improve TTM-001 capsid tropism in the CNS of mice were skewed to the C-terminal portion of the peptide insertion in loop IV of the sequence. Additionally, a number of these C-terminal modifications were the incorporation of an arginine (R) or leucine (L) residue.

Table 13. NGS fold-enrichment of TTM-001 matured AAV capsid variants in the brain of CD-1 Outbred mice

Peptide Sequence	SEQ ID NO:	Fold enrichment over TTM-001	Peptide Sequence	SEQ ID NO:	Fold enrichment over TTM-001
KTINGSGSPHSLLWNQQT	1008	7.983	KTINGRPSPHVKAQNQQT	1365	2.419
KTINGSGSPHSKAQYYVT	1009	6.283	KTINGSGSPHSMFPNQQT	1366	2.419
KTINGSGSPHSKLRRQQT	1010	6.231	KTINGMKSPHSKAQNQQT	1367	2.413
KTINGSGSPHSIWQNQQT	1011	5.883	KTINGSGSPHSKVRAQQT	1368	2.412
KTINGSSSPHCTAQNQQT	1012	5.607	KGLVGSGSPHSKAQNQQT	1369	2.412
KTINGSGSPHSKAGCGQT	1013	5.341	KTINGSRSPHVRAQNQQT	1370	2.412
KSMNGSGSPHSRAQNQQT	1014	5.145	KTIRLRGSPHSKAQNQQT	1371	2.408
KTINGSGSPHSKRLRQQT	1015	5.034	KPRLGSGSPHSKAQNQQT	1372	2.406
KTINGSGSPHSLRRNQQT	1016	4.985	KTINGSGSPHSKAWYPQT	1373	2.406
KTINGSGSPHSRGRNQQT	1017	4.961	KTINGSCSPHVRAQNQQT	1374	2.406
KTINGSGSPHSEIVNQQT	1018	4.931	KTINGSGSPHSKRNVQQT	1375	2.404
KTINGSGSPHSSRRNQQT	1019	4.920	KTINGSGSPHSKRGLQQT	1376	2.402
KTINGSGSPHCLLQNQQT	1020	4.898	KTINGSGSPHSKRLAQQT	1377	2.397
KTIMRVGSPHSKAQNQQT	1021	4.875	KTINGSGSPHSKVTRQQT	1378	2.397
KTINGSGSPHSKAFRLQT	1022	4.849	KTINGSGSPHTWLQNQQT	1379	2.397
KTINGSGSPHCLAQNQQT	1023	4.847	KTINGSGSPHSKQRSQQT	1380	2.395
KTINGSCSPHRKAQNQQT	1024	4.801	KTINGSGSPHSKAQRCST	1381	2.393
KTINGSGSPHFLRQNQQT	1025	4.777	KTINGSPSPHYLAQNQQT	1382	2.391
KTINGSGSPHSLRFNQQT	1026	4.765	KTINGRRSPHLKAQNQQT	1383	2.391
KTINGSGSPHSYLRNQQT	1027	4.566	KTINGSGSPHWSLQNQQT	1384	2.389
KTINGSGSPHCSLQNQQT	1028	4.540	KTINGVTSPHWKAQNQQT	1385	2.387
KTINGSGSPHVLWQNQQT	1029	4.533	KTINGSGSPHSKAQTTRT	1386	2.385
KTINGSGSPHSKWLLQQT	1030	4.521	KTINGSGSPHSKAQLFKT	1387	2.385
KTINGSGSPHSLWSNQQT	1031	4.467	KTINGSGSPHSKARSYQT	1388	2.382

KTINGSGSPHSKRRLQQT	1032	4.451	KTINGSGSPHSKLSRQQT	1389	2.380
KTINGSGSPHSVYLNQQT	1033	4.426	KTINGSGSPHLVFQNQQT	1390	2.376
KTINGSGSPHSLWLNQQT	1034	4.412	KTINGSGSPHSKAQLVKT	1391	2.376
KTINGSGSPHSKAQRKLT	1035	4.339	KTINGSGSPHSKAQTGRT	1392	2.371
KTINGSGSPHSKALRRQT	1036	4.330	KTINGSGSPHSKARYSQT	1393	2.369
KTINGSGSPHSKAQRLRT	1037	4.322	KTINGRHSPHLKAQNQQT	1394	2.369
KYLSGSGSPHSKAQNQQT	1038	4.264	KTINGSGSPHSFARNQQT	1395	2.367
KTINGSGSPHSKAQRRLT	1039	4.227	KTINGSGSPHSSCQNQQT	1396	2.365
KTINGSGSPHSKARRQQT	1040	4.218	KTINGSGSPHSLFANQQT	1397	2.365
KTINGSGSPHSKARRLQT	1041	4.210	KTINGSGSPHSKRLTQQT	1398	2.363
KTINGSGSPHSKSRRQQT	1042	4.175	KTINGSGSPHSKAQTART	1399	2.363
KTINGLLSPHWKAQNQQT	1043	4.173	KTINGSGSPHFLNQNQQT	1400	2.359
KTINGSGSPHSKARLRQT	1044	4.155	KTINGSGSPHSKAQLILT	1401	2.358
KTINGSGSPHSKASKRQT	1045	4.117	KEVGGSGSPHSKAQNQQT	1402	2.358
KTINGSGSPHVRRQNQQT	1046	4.114	KTINGSRSPHIRAQNQQT	1403	2.356
KTINGSGSPHSKAQLYRT	1047	4.108	KTINGSGSPHSSWLNQQT	1404	2.354
KGLSGSGSPHSKAQNQQT	1048	4.056	KTINGSMSPHLYAQNQQT	1405	2.354
KTINGSGSPHSLFRNQQT	1049	4.037	KTINGMSSPHRKAQNQQT	1406	2.352
KTINGSGSPHSKAQLTVT	1050	4.026	KTINGSGSPHSKPRPQQT	1407	2.352
KTINGSRSPHTRAQNQQT	1051	3.989	KTINGSGSPHSGLWNQQT	1408	2.350
KTINGSGSPHSKAKLRQT	1052	3.976	KTINGSGSPHRWAQNQQT	1409	2.350
KTINGSGSPHSKLIRQQT	1053	3.968	KTINGSGSPHSKIRLQQT	1410	2.348
KTINGSGSPHSKALRFQT	1054	3.894	KTINGSGSPHSKFSCQQT	1411	2.348
KTINGSGSPHSKRTFQQT	1055	3.879	KTINGSGSPHSKSCAQQT	1412	2.348
KTINGSGSPHSKAQKRLT	1056	3.872	KTINGSGSPHSKRLMQQT	1413	2.345
KTINGSGSPHLWSQNQQT	1057	3.857	KTSRCSGSPHSKAQNQQT	1414	2.345
KTINGSGSPHLLWQNQQT	1058	3.855	KTINGSGSPHSFLLNQQT	1415	2.341
KTINGSGSPHSRLRNQQT	1059	3.851	KTINGSGSPHCSAQNQQT	1416	2.339
KTINGSGSPHSKRAAQQT	1060	3.838	KTINGSGSPHSKAQPSKT	1417	2.339
KTINGSGSPHSKRSWQQT	1061	3.838	KTINGSGSPHSYVRNQQT	1418	2.339
KTINGSGSPHSKAQLRRT	1062	3.825	KTINGSGSPHSKAQQSRT	1419	2.337
KTINGSGSPHYLVQNQQT	1063	3.819	KTINGSGSPHSFVVNQQT	1420	2.335
KTINGSGSPHSKLFRQQT	1064	3.806	KTINGFRSPHSKAQNQQT	1421	2.333
KTINGSGSPHSKRAMQQT	1065	3.801	KTINGSGSPHSKWLVQQT	1422	2.333
KTINGSGSPHSKTLRQQT	1066	3.788	KTINGSGSPHSKRTAQQT	1423	2.331
KTINGSGSPHSRSRNQQT	1067	3.784	KTINGLFSPHRKAQNQQT	1424	2.331
KTINGSGSPHRRRQNQQT	1068	3.754	KTINTIESPHSKAQNQQT	1425	2.331
KTINGSGSPHSKTCLQQT	1069	3.717	KRLFGSGSPHSKAQNQQT	1426	2.330
KTINGSGSPHSKSRWQQT	1070	3.698	KTINGSGSPHSKAPNHLT	1427	2.324
KTINGSGSPHRFRQNQQT	1071	3.698	KTINGSGSPHLFRQNQQT	1428	2.324
KTINGLRSPHRKAQNQQT	1072	3.676	KTINGSGSPHSKASRHQT	1429	2.324
KTRSRSGSPHSKAQNQQT	1073	3.669	KTINGSGSPHSKLSWQQT	1430	2.322
KTINGSGSPHSKAQLVVT	1074	3.654	KETAGSGSPHSKAQNQQT	1431	2.320
KTINGSGSPHSRKLNQQT	1075	3.646	KTINGHRSPHLKAQNQQT	1432	2.320
KTINGSGSPHSLLCNQQT	1076	3.644	KTINGSGSPHSKGCLQQT	1433	2.320
KTINGKRSPHSKAQNQQT	1077	3.611	KTINGSGSPHSKAQVLIT	1434	2.317
KTINGSRSPHLFAQNQQT	1078	3.601	KTINGSGSPHSKLRSQQT	1435	2.309
KSINGSGSPHSKAHDQQT	1079	3.592	KTINGTLSPHRKAQNQQT	1436	2.307
KTINGSGSPHSKAQRSRT	1080	3.585	KTINGSGSPHSTWTNQQT	1437	2.307
KTINGSGSPHSTWLNQQT	1081	3.583	KTINGSGSPHSKAQCRLT	1438	2.304
KTINGSGSPHSKASRRQT	1082	3.577	KTINNLRSPHSKAQNQQT	1439	2.302
KTINGSGSPHSKRSMQQT	1083	3.561	KTINGSGSPHSKARANQT	1440	2.300
KTINGSGSPHSCLQNQQT	1084	3.559	KTINGRQSPHTKAQNQQT	1441	2.294

KTINGSGSPHSKRLWQQT	1085	3.529	KTINSARSPHSKAQNQQT	1442	2.292
KTINGSGSPHSWLSNQQT	1086	3.495	KTINGCSSPHRKAQNQQT	1443	2.291
KTINGSGSPHLRRQNQQT	1087	3.493	KTINGSVSPHFMAQNQQT	1444	2.287
KTINGSGSPHSKARRSQT	1088	3.493	KTINGSGSPHSLCQNQQT	1445	2.285
KTINGSGSPHSKHLRQQT	1089	3.438	KTINGSLSPHLFAQNQQT	1446	2.285
KTINGSGSPHSCSQNQQT	1090	3.428	KTINGSGSPHSKACPLQT	1447	2.285
KTINGSGSPHSKSFRQQT	1091	3.426	KTINGRTSPHRKAQNQQT	1448	2.285
KTINGSGSPHLCLQNQQT	1092	3.425	KTINGSGSPHSKRATQQT	1449	2.283
KTINGSGSPHSKAQTSRT	1093	3.421	KTINGSGSPHSKARIMQT	1450	2.283
KTINGSGSPHSLCSNQQT	1094	3.413	KTINGSGSPHVTWQNQQT	1451	2.281
KTINGSRSPHLRAQNQQT	1095	3.410	KTINGSGSPHSKRLPQQT	1452	2.279
KTINGSGSPHSKAQVSKT	1096	3.406	KTINGSGSPHSKAQGFRT	1453	2.279
KTINGSGSPHSKAQRHVT	1097	3.404	KTINGSGSPHLYGQNQQT	1454	2.277
KTINGSSSPHLCAQNQQT	1098	3.402	KTINGSGSPHLSCQNQQT	1455	2.277
KTINGSGSPHSFLRNQQT	1099	3.384	KTINGSGSPHSKAQFTLT	1456	2.277
KTINGSGSPHSFVLNQQT	1100	3.382	KTINGSRSPHFKAQNQQT	1457	2.277
KTINGSGSPHSKMRAQQT	1101	3.382	KTINGRPSPHSKAQNQQT	1458	2.276
KTINGSGSPHRPRQNQQT	1102	3.380	KTINGFSSPHRKAQNQQT	1459	2.276
KTINGSGSPHSKCLLQQT	1103	3.374	KTINGRASPHVKAQNQQT	1460	2.272
KTINGSGSPHSKAQSRRT	1104	3.372	KTINGSGSPHSKAQNEVH	1461	2.272
KTINGSGSPHSRWQNQQT	1105	3.372	KTINGSGSPHSKRSLQQT	1462	2.270
KYSVGSGSPHSKAQNQQT	1106	3.365	KTINGSGSPHRSRQNQQT	1463	2.270
KTINGSGSPHSKRFLQQT	1107	3.359	KPPTGSGSPHSKAQNQQT	1464	2.270
KTINGSGSPHSLFLNQQT	1108	3.358	KTINGSGSPHSKAARTQT	1465	2.266
KTINGSGSPHSKAYLRQT	1109	3.356	KTINGSGSPHSSWANQQT	1466	2.264
KTINGSGSPHSKRNGQQT	1110	3.350	KTINGSGSPHSKAQRHAT	1467	2.264
KTINGSGSPHTRRQNQQT	1111	3.350	KTINSPRSPHSKAQNQQT	1468	2.264
KTINGSGSPHSKPRLQQT	1112	3.337	KTINGSGSPHSKSERQQT	1469	2.263
KTINFLRSPHSKAQNQQT	1113	3.331	KTINGSGSPHALFQNQQT	1470	2.261
KTINGSGSPHLLCQNQQT	1114	3.328	KTINGSGSPHSKAQCYVT	1471	2.261
KTINGSGSPHSKARIVQT	1115	3.287	KTINGVASPHRKAQNQQT	1472	2.261
KTINGSKSPHFKAQNQQT	1116	3.285	KTINGSGSPHSALWNQQT	1473	2.261
KTINGSGSPHSKAQIRLT	1117	3.279	KTINGSGSPHSKSVRQQT	1474	2.259
KTINGSSSPHWVAQNQQT	1118	3.277	KTINGSGSPHSHMENQQT	1475	2.259
KTINGSGSPHSKATRRQT	1119	3.277	KTINGSGTPHSKAQNQQT	1476	2.259
KTINGSLSPHSCAQNQQT	1120	3.268	KTINGSGSPHSKTGRQQT	1477	2.259
KTINGSGSPHSLYLNQQT	1121	3.264	KTINGSGSPHSKAQANRT	1478	2.255
KTINGSGSPHSKVGRQQT	1122	3.255	KTINGSGSPHSKARFSQT	1479	2.253
KTINGSGSPHSRRLNQQT	1123	3.251	KYLLGSGSPHSKAQNQQT	1480	2.253
KTINGSGSPHSKAQHSRT	1124	3.227	KTINGSGSPHCSSQNQQT	1481	2.253
KTINGSGSPHSKAFPRQT	1125	3.220	KTINGSMSPHRKAQNQQT	1482	2.251
KTINGSPSPHRRAQNQQT	1126	3.216	KTINGNLSPHRKAQNQQT	1483	2.250
KTINGSGSPHSKRNLQQT	1127	3.210	KEVAGSGSPHSKAQNQQT	1484	2.250
KTINGSGSPHSKPTRQQT	1128	3.201	KTINLSRSPHSKAQNQQT	1485	2.246
KTINGSGSPHSKLWLQQT	1129	3.199	KTINGSGSPHSKARQQQT	1486	2.244
KTINGSGSPHWLAQNQQT	1130	3.192	KTINGTPSPHRKAQNQQT	1487	2.244
KTINGSGSPHRTRQNQQT	1131	3.190	KTINGSGSPHSKFKLQQT	1488	2.244
KTINGSGSPHSKLNKQQT	1132	3.181	KTINGSGSPHSKAWLLQT	1489	2.240
KTINGSGSPHSSLWNQQT	1133	3.179	KTINGLRSPHSKAQNQQT	1490	2.238
KTINGSGSPHSKAQITLT	1134	3.177	KTINGRLSPHRKAQNQQT	1491	2.238
KTINGSGSPHSKFLFQQT	1135	3.173	KTINGSPSPHLFAQNQQT	1492	2.238
KTINGSGSPHSKRTPQQT	1136	3.169	KDLRGSGSPHSKAQNQQT	1493	2.238
KTINGSGSPHSKAQNSRR	1137	3.168	KTINGSGSPHSKAQLAKT	1494	2.238

KTINGSGSPHSRLKNQQT	1138	3.156	KTINGSGSPHSKPRSQQT	1495	2.235
KTINGSGSPHSCLLNQQT	1139	3.127	KTINGSGSPHSKKMSQQT	1496	2.235
KTINGSGSPHTLYQNQQT	1140	3.117	KTINGSGSPHSKAQLIVT	1497	2.235
KTINGSGSPHSKYPSQQT	1141	3.114	KTINGSGSPHSKARFTQT	1498	2.233
KTINGSGSPHSKLRNQQT	1142	3.112	KTINGSGSPHPLFQNQQT	1499	2.233
KTINGSGSPHLNWQNQQT	1143	3.112	KTINGSGSPHSKAQRGMT	1500	2.231
KTINGVVSPHRKAQNQQT	1144	3.106	KTINGSGSPHSKAQNLRR	1501	2.231
KTINGSGSPHSYRPNQQT	1145	3.095	KTINGSGSPHSKAQFRVT	1502	2.231
KTINGSTSPHRRAQNQQT	1146	3.089	KTINGSGSPHSKAFVRQT	1503	2.225
KTINGSCSPHPLAQNQQT	1147	3.086	KTINGSGSPHSKARLTQT	1504	2.223
KTINGSGSPHSKAFARQT	1148	3.082	KTINGSGSPHRFKQNQQT	1505	2.223
KTINGSGSPHSKALRYQT	1149	3.076	KTINGSGSPHSKEETQQT	1506	2.223
KTINGSKSPHRLAQNQQT	1150	3.073	KTINGSGSPHSKTRAQQT	1507	2.223
KTINMRVSPHSKAQNQQT	1151	3.073	KTINGSGSPHSVSWNQQT	1508	2.223
KTINGSGSPHMYLQNQQT	1152	3.061	KTINGSGSPHTKWQNQQT	1509	2.222
KTINGSGSPHSKLARQQT	1153	3.054	KTINGSNSPHRKAQNQQT	1510	2.218
KTINGSGSPHSKARPYQT	1154	3.050	KTINGSGSPHSKAQNKRS	1511	2.214
KTINGSGSPHSKRVPQQT	1155	3.048	KTINGSGSPHSTRQNQQT	1512	2.212
KTINGSGSPHLSWQNQQT	1156	3.047	KTINGTRSPHTKAQNQQT	1513	2.203
KTINGRSSPHGKAQNQQT	1157	3.035	KTINGSGSPHVLFQNQQT	1514	2.203
KTINGSGSPHLWTQNQQT	1158	3.034	KTINGSVSPHYLAQNQQT	1515	2.203
KTINGLLSPHRKAQNQQT	1159	3.026	KTINGALSPHRKAQNQQT	1516	2.203
KTINGSGSPHRLRQNQQT	1160	2.998	KTINGSGSPHSKARLYQT	1517	2.201
KTINGSCSPHSGAQNQQT	1161	2.994	KTINGSGSPHEHNQNQQT	1518	2.199
KTINGSGSPHSKAQRRST	1162	2.993	KTINGVLSPHWKAQNQQT	1519	2.199
KTINGSGSPHSKLCSQQT	1163	2.989	KTINGSGSPHSKASRQQT	1520	2.197
KTINGSGSPHSKAQLLKT	1164	2.985	KTINGSGSPHSKRSFQQT	1521	2.197
KTINGRKSPHSKAQNQQT	1165	2.985	KTINGSGSPHSKRVSQQT	1522	2.196
KTINGSGSPHLLYQNQQT	1166	2.983	KTINGSGSPHSYSRNQQT	1523	2.196
KTINGSGSPHSKLLRQQT	1167	2.981	KTINGSGSPHSTVWNQQT	1524	2.196
KTINGSGSPHSLRHNQQT	1168	2.980	KTINGSGSPHSVLFNQQT	1525	2.194
KTINGSGSPHSSKRNQQT	1169	2.978	KTINGPLSPHCKAQNQQT	1526	2.194
KTINGSGSPHSKARSRQT	1170	2.972	KTINGSGSPHSKRVGQQT	1527	2.190
KTINGRSSPHRKAQNQQT	1171	2.965	KTINGSGSPHSKLWSQQT	1528	2.190
KTINGSKSPHRTAQNQQT	1172	2.950	KTINGSGSPHSKAQGVRT	1529	2.188
KTINGMRSPHVKAQNQQT	1173	2.937	KTINGSVSPHRRAQNQQT	1530	2.186
KTINGSGSPHSKRMSQQT	1174	2.931	KTINGSGSPHLRFQNQQT	1531	2.186
KTINGSGSPHSKVPKQQT	1175	2.924	KTINGSASPHVFAQNQQT	1532	2.186
KTINLIRSPHSKAQNQQT	1176	2.920	KTWVRSGSPHSKAQNQQT	1533	2.186
KTINGSGSPHPFLQNQQT	1177	2.916	KTINGSGSPHSKARMQQT	1534	2.184
KTINGSGSPHSKARLWQT	1178	2.914	KTINGSGSPHSKASRGQT	1535	2.182
KTINGSGSPHSRTRNQQT	1179	2.912	KTINGSGSPHSKAQVCLT	1536	2.182
KTINGSGSPHSKRSNQQT	1180	2.886	KTINGSGSPHSKARGVQT	1537	2.181
KTINGSLSPHSWAQNQQT	1181	2.885	KTINGSGSPHGLWQNQQT	1538	2.181
KTINGSRSPHYKAQNQQT	1182	2.879	KTINGSGSPHSKAQVWFT	1539	2.181
KTINRHSSPHSKAQNQQT	1183	2.877	KTINGSGSPHSKAQVTLT	1540	2.179
KTINGSGSPHSKRRNQQT	1184	2.877	KTINGSGSPHSKAQLRIT	1541	2.179
KTINGSGSPHSKAKHLQT	1185	2.870	KDSLGSGSPHSKAQNQQT	1542	2.175
KTINGSGSPHSKRTYQQT	1186	2.870	KTINGSGSPHSKRASQQT	1543	2.173
KTINGVLSPHRKAQNQQT	1187	2.868	KTINGSGSPHSKRINQQT	1544	2.173
KTINGSGSPHSFITNQQT	1188	2.868	KTINGSGSPHSKASKNQT	1545	2.171
KTINGSGSPHSTRLNQQT	1189	2.860	KTINGSGSPHSKAQLPWT	1546	2.169
KTINGSGSPHSKRTSQQT	1190	2.857	KTINGSGSPHSKLTRQQT	1547	2.169

KTINGSGSPHSRRSNQQT	1191	2.851	KTINGSGSPHSKTNRQQT	1548	2.169
KTINGHLSPHRKAQNQQT	1192	2.851	KTINRVISPHSKAQNQQT	1549	2.169
KTINGSGSPHSKAQFSRT	1193	2.847	KTINGSGSPHTLWQNQQT	1550	2.168
KTINGSGSPHSKAQTFRT	1194	2.847	KTINGSGSPHSRRQNQQT	1551	2.166
KTINGSGSPHSKPLRQQT	1195	2.844	KTINGSGSPHSKGGRQQT	1552	2.164
KTINGSGSPHSKASCRQT	1196	2.840	KTINGSESPHDSAQNQQT	1553	2.164
KTINGSGSPHSKILWQQT	1197	2.838	KTINGSGSPHSRPRNQQT	1554	2.164
KTINGSGSPHSKALKRQT	1198	2.836	KTINGSGSPHSRKQNQQT	1555	2.162
KTINGSGSPHSKAHRSQT	1199	2.819	KTINGSGSPHSKAQEELT	1556	2.162
KTINGSGSPHSMLYNQQT	1200	2.808	KTINGWRSPHSKAQNQQT	1557	2.160
KTINGSGSPHSKCTLQQT	1201	2.808	KTINGSGSPHSLLYNQQT	1558	2.158
KTINGSGSPHSKAQNRMR	1202	2.804	KTINGSGSPHSFRLNQQT	1559	2.158
KTINGSGSPHSKLVRQQT	1203	2.801	KTINGSGSPHSKAQFLRT	1560	2.156
KTINGSGSPHSKRILQQT	1204	2.801	KTINGSGSPHSKQSRQQT	1561	2.156
KTINGSGSPHSKAQWLRT	1205	2.795	KTINGSRSPHSKAQNRQT	1562	2.155
KTINGSGSPHSLTCNQQT	1206	2.795	KTINGRPSPHIKAQNQQT	1563	2.155
KTINGIRSPHTKAQNQQT	1207	2.793	KTINGSGSPHSKRLVQQT	1564	2.151
KTINGSGSPHSKAQRWLT	1208	2.788	KGHEGSGSPHSKAQNQQT	1565	2.151
KTINGSGSPHSKAQLSIT	1209	2.784	KTINGSGSPHSKAQKRST	1566	2.151
KTINGSGSPHIYRQNQQT	1210	2.782	KTINGSGSPHSYLLNQQT	1567	2.147
KTINGSGSPHSLRSNQQT	1211	2.778	KTINGSGSPHSKPRGQQT	1568	2.147
KTINGSGSPHSKVKPQQT	1212	2.778	KTINGSGSPHSKTRLQQT	1569	2.145
KTINGSGSPHSKATRHQT	1213	2.777	KTINGSGSPHSKSHRQQT	1570	2.145
KTINGSLSPHLCAQNQQT	1214	2.775	KEIKGSGSPHSKAQNQQT	1571	2.140
KTINGSGSPHSKACASQT	1215	2.775	KTINGSGSPHSKARGIQT	1572	2.140
KWSPGSGSPHSKAQNQQT	1216	2.764	KTINGYRSPHSKAQNQQT	1573	2.140
KTINGYLSPHRKAQNQQT	1217	2.762	KTINGSGSPHSKLWTQQT	1574	2.140
KTINGSGSPHSKVIRQQT	1218	2.760	KTINGSGSPHSKPWLQQT	1575	2.138
KTINGSGSPHFLLQNQQT	1219	2.758	KTINGSGSPHWSVQNQQT	1576	2.138
KTINGSGSPHSKARSKQT	1220	2.758	KTINGSGSPHSKVARQQT	1577	2.136
KTINGVPSPHWKAQNQQT	1221	2.758	KTINGSGSPHTLFQNQQT	1578	2.136
KTINGSGSPHSKATRNQT	1222	2.756	KTINGSCSPHLAAQNQQT	1579	2.134
KTINGSGSPHSKACSAQT	1223	2.754	KTINGSGSPHSKTSRQQT	1580	2.132
KTINGSGSPHSKARYVQT	1224	2.749	KTINGSGSPHSKAQNARH	1581	2.127
KTINGSRSPHARAQNQQT	1225	2.745	KTINGSGSPHSKAQLKLT	1582	2.125
KTINGSGSPHSKAQHLRT	1226	2.741	KTINGSGSPHSKAQNWRT	1583	2.125
KTINGSGSPHSKAKSRQT	1227	2.739	KTINGSGSPHFLPQNQQT	1584	2.123
KTINGSGSPHSKIGRQQT	1228	2.739	KTINGSGSPHSKNVRQQT	1585	2.123
KTINGLASPHRKAQNQQT	1229	2.737	KTINGSGSPHFMRQNQQT	1586	2.123
KTINGSGSPHSKARTRQT	1230	2.737	KTINGSGSPHGWAQNQQT	1587	2.121
KTINGSGSPHSKSIRQQT	1231	2.728	KTINGSGSPHFHLQNQQT	1588	2.121
KTINGSGSPHSKRLYQQT	1232	2.721	KTINGSASPHWSAQNQQT	1589	2.121
KTINGLPSPHRKAQNQQT	1233	2.719	KTINGSSSPHSWAQNQQT	1590	2.119
KTINGSLSPHRRAQNQQT	1234	2.717	KTINGSGSPHSKAHRQQT	1591	2.117
KTINGKTSPHGKAQNQQT	1235	2.717	KTINGSGSPHSKQRVQQT	1592	2.117
KTINGSRSPHRLAQNQQT	1236	2.698	KTLRRSGSPHSKAQNQQT	1593	2.117
KTINGSGSPHSLTWNQQT	1237	2.698	KTINGSGSPHSKGVRQQT	1594	2.115
KTINGSKSPHRKAQNQQT	1238	2.696	KTINGSLSPHTWAQNQQT	1595	2.115
KTINGSGSPHSKAQLRKT	1239	2.689	KTINGSGSPHSKRALQQT	1596	2.114
KTINGSGSPHSKSRHQQT	1240	2.685	KTINGSGSPHCLSQNQQT	1597	2.114
KTINRRLSPHSKAQNQQT	1241	2.678	KTINGSGSPHSKAQSLKT	1598	2.110
KTINGSGSPHSRRVNQQT	1242	2.676	KTINGSGSPHSFVRNQQT	1599	2.110
KTINGSGSPHSHWQNQQT	1243	2.676	KTINGSGSPHSIFSNQQT	1600	2.110

KTTHCSGSPHSKAQNQQT	1244	2.672	KTINGSGSPHSKVSRQQT	1601	2.108
KTINGSGSPHSWLQNQQT	1245	2.665	KTINGSGSPHSKARNKQT	1602	2.108
KTINGSTSPHYLAQNQQT	1246	2.665	KTINASGSPHSKAQGQQT	1603	2.108
KTINGLTSPHRKAQNQQT	1247	2.663	KTINGSGSPHSKLRMQQT	1604	2.106
KTINGSGSPHSKRLLQQT	1248	2.659	KTINGSWSPHMLAQNQQT	1605	2.106
KTINGSGSPHSKLCVQQT	1249	2.659	KTINGSGSPHSLFPNQQT	1606	2.106
KTINGFLSPHRKAQNQQT	1250	2.654	KPPLGSGSPHSKAQNQQT	1607	2.102
KTINGSGSPHSKMRPQQT	1251	2.652	KTINGIASPHRKAQNQQT	1608	2.099
KTINGSGSPHSKQTRQQT	1252	2.650	KTINGSCSPHSLAQNQQT	1609	2.099
KTINGSGSPHSYLINQQT	1253	2.650	KTINGRLSPHFKAQNQQT	1610	2.097
KTINGSGSPHSKALRSQT	1254	2.648	KTINGSGSPHSKARMTQT	1611	2.091
KTINGMLSPHRKAQNQQT	1255	2.646	KTINGSGSPHSKARLQQT	1612	2.089
KTINGSGSPHSKCLTQQT	1256	2.644	KTINGSGSPHSKWVSQQT	1613	2.089
KTINGSGSPHSKAQLTLT	1257	2.641	KTINGSGSPHSKKVSQQT	1614	2.088
KTINGHSSPHRKAQNQQT	1258	2.639	KTINGSGSPHSKAQSYRT	1615	2.088
KTINGSGSPHLTWQNQQT	1259	2.637	KAFNGSGSPHSKAPNQQT	1616	2.088
KTINGSGSPHSKAQYCLT	1260	2.628	KTINGSGSPHSKAQYRLT	1617	2.088
KTINGSGSPHSFLVNQQT	1261	2.624	KTINGSWSPHLVAQNQQT	1618	2.084
KTINMSRSPHSKAQNQQT	1262	2.622	KTINGSGSPHSWTQNQQT	1619	2.084
KTINGSGSPHSKAQLHRT	1263	2.618	KTINGSGSPHSKAQSHRT	1620	2.084
KTINGSGSPHLYMQNQQT	1264	2.615	KGINGSGSPHGKAQNQQT	1621	2.084
KTINGSRSPHRRAQNQQT	1265	2.615	KTINGSGSPHSKAQNRKL	1622	2.084
KTINGSGSPHSKAQNRRS	1266	2.613	KTINGRYSPHSKAQNQQT	1623	2.080
KTINLRFSPHSKAQNQQT	1267	2.611	KTINGSGSPHSKGRSQQT	1624	2.080
KTINGSGSPHSKAQRLWT	1268	2.611	KTINGSGSPHCVAQNQQT	1625	2.080
KTINGSGSPHSKGRAQQT	1269	2.607	KTINGSGSPHSKIRPQQT	1626	2.080
KTINGSGSPHSLSCNQQT	1270	2.605	KTINGSGSPHSKAQSSKT	1627	2.078
KTINGLVSPHCKAQNQQT	1271	2.605	KTINGSGSPHSKRPFQQT	1628	2.076
KTINGSSSPHLWAQNQQT	1272	2.605	KTINGSSSPHCLAQNQQT	1629	2.074
KTINGSGSPHSKAHRLQT	1273	2.603	KTINGTRSPHAKAQNQQT	1630	2.071
KTINGSGSPHPYAQNQQT	1274	2.598	KTINGSGSPHLLFQNQQT	1631	2.069
KTINGSGSPHSTRPNQQT	1275	2.598	KTINGRRSPHTKAQNQQT	1632	2.069
KTINGRSSPHPKAQNQQT	1276	2.596	KTINGSGSPHSKASKQQT	1633	2.069
KTINGSGSPHSKAQSWRT	1277	2.596	KTINGSGSPHSKAQLGRT	1634	2.069
KTINGQRSPHVKAQNQQT	1278	2.596	KTINGSGSPHSVFLNQQT	1635	2.069
KTINGSGSPHSKAQFVRT	1279	2.596	KTINGSGSPHSKSARQQT	1636	2.067
KTINGSGSPHSKCLNQQT	1280	2.594	KTINGSGSPHSKLRLQQT	1637	2.065
KTINGSGSPHSSLCNQQT	1281	2.592	KTRKSSGSPHSKAQNQQT	1638	2.065
KTINGQRSPHSKAQNQQT	1282	2.590	KTINGFRSPHLKAQNQQT	1639	2.063
KTINGSGSPHSLSWNQQT	1283	2.588	KTINGSGSPHSKRSIQQT	1640	2.063
KTINGSGSPHSSRKNQQT	1284	2.585	KTINGSGSPHSKGRIQQT	1641	2.061
KTINGSGSPHSKRTLQQT	1285	2.583	KTINGSRSPHRPAQNQQT	1642	2.061
KTINGSLSPHCLAQNQQT	1286	2.577	KTINGSGSPHSKLRPQQT	1643	2.060
KTINGSGSPHSKAQSSRT	1287	2.575	KTINGSGSPHMYAQNQQT	1644	2.060
KTINGSGSPHLKRQNQQT	1288	2.564	KTINGRTSPHAKAQNQQT	1645	2.060
KTINGSGSPHSKARMGQT	1289	2.564	KTINGSGSPHSKAGRGQT	1646	2.058
KTINGSGSPHSKAQVKLT	1290	2.564	KTINGSGSPHSKLMRQQT	1647	2.056
KTINGSGSPHSKLPRQQT	1291	2.562	KTINGSGSPHSKANKSQT	1648	2.056
KTINGSGSPHSKLCLQQT	1292	2.562	KTINGSGSPHSKAVRQQT	1649	2.052
KTINGSGSPHSPLWNQQT	1293	2.562	KTINGSGSPHSKCLSQQT	1650	2.052
KTINGSVSPHSWAQNQQT	1294	2.562	KTINGSGSPHSKAQWVLT	1651	2.052
KTIRSKGSPHSKAQNQQT	1295	2.559	KTINGSGSPHSKAQFWVT	1652	2.050
KTINGSRSPHSWAQNQQT	1296	2.559	KTINGSGSPHSKALCRQT	1653	2.048

KTINGSGSPHSKILRQQT	1297	2.557	KEVMGSGSPHSKAQNQQT	1654	2.047
KTINGRQSPHVKAQNQQT	1298	2.557	KTINGSGSPHSKNTRQQT	1655	2.047
KTINGSGSPHSKAQSIKT	1299	2.555	KTINGSGSPHTWTQNQQT	1656	2.045
KTINGSGSPHSKAQASKT	1300	2.546	KTINGSTSPHWSAQNQQT	1657	2.043
KTINGSGSPHSRLFNQQT	1301	2.542	KTINGNVSPHRKAQNQQT	1658	2.043
KTINGSGSPHIYLQNQQT	1302	2.542	KTINGSTSPHLFAQNQQT	1659	2.041
KTINGSGSPHSRVRNQQT	1303	2.540	KTINGSGSPHSKAQNYRA	1660	2.039
KTINGSGSPHSKAVRAQT	1304	2.538	KTINGSGSPHSKARGQQT	1661	2.039
KTINGSGSPHSKPARQQT	1305	2.538	KTINGSGSPHSKAQNRIR	1662	2.039
KTINGSGSPHSRYSNQQT	1306	2.536	KTINGSGSPHSKWTLQQT	1663	2.039
KTINGSRSPHRSAQNQQT	1307	2.536	KTINGSGSPHSKAQMKCT	1664	2.039
KTINGSLSPHIYAQNQQT	1308	2.536	KTINGSGSPHSLWQNQQT	1665	2.037
KTINGSGSPHSKPVRQQT	1309	2.529	KTINGSGSPHSKAQLSKT	1666	2.035
KTINGMRSPHGKAQNQQT	1310	2.527	KTINLIWSPHSKAQNQQT	1667	2.035
KTINGSGSPHSKARITQT	1311	2.525	KTINGSGSPHSKRVLQQT	1668	2.035
KTINGSGSPHSWSLNQQT	1312	2.523	KTINGSGSPHSKVRVQQT	1669	2.034
KTINTSRSPHSKAQNQQT	1313	2.520	KTINSRFSPHSKAQNQQT	1670	2.032
KTINGSGSPHSKAFTRQT	1314	2.518	KRSKGSGSPHSKAQNQQT	1671	2.030
KTINGSGSPHSKAVRNQT	1315	2.516	KTINGSGSPHRRLQNQQT	1672	2.030
KTINGSGSPHSKAQTNRT	1316	2.510	KTINGSGSPHSCAQNQQT	1673	2.030
KTINGSGSPHSKANRMQT	1317	2.508	KTINGPLSPHRKAQNQQT	1674	2.028
KTINGSGSPHSKAQLVLT	1318	2.508	KTINGSVSPHLYAQNQQT	1675	2.028
KTINGSGSPHSKATRQQT	1319	2.505	KTINGRISPHLKAQNQQT	1676	2.028
KTINGSGSPHSKARGTQT	1320	2.505	KTINGSHSPHRKAQNQQT	1677	2.028
KTINGSGSPHSKAQWSVT	1321	2.501	KTINGSGSPHSKAQVSIT	1678	2.028
KTINGSGSPHSKAWLIQT	1322	2.499	KTINGSMSPHRRAQNQQT	1679	2.028
KTINGSGSPHSKAFRPQT	1323	2.499	KTINGRQSPHAKAQNQQT	1680	2.026
KTINGSGSPHRRSQNQQT	1324	2.497	KTINGSGSPHSKAVWRQT	1681	2.026
KTINGSGSPHSKGIRQQT	1325	2.497	KQPLGSGSPHSKAQNQQT	1682	2.024
KTINGSGSPHCTLQNQQT	1326	2.497	KTINGSGSPHSKAQNVKL	1683	2.024
KPLPGSGSPHSKAQNQQT	1327	2.495	KTINGSGSPHSKRGTQQT	1684	2.022
KTINGSGSPHLVCQNQQT	1328	2.493	KTINGSVSPHYVAQNQQT	1685	2.022
KTINGSGSPHSKARGYQT	1329	2.492	KTINGSGSPHSKNLRQQT	1686	2.022
KTINGRISPHGKAQNQQT	1330	2.492	KTINGSGSPHSKAQAFRT	1687	2.022
KTINGSSSPHWLAQNQQT	1331	2.490	KTINGSGSPHSKCSNQQT	1688	2.020
KTINGSGSPHSKARMAQT	1332	2.488	KELVGSGSPHSKAQNQQT	1689	2.019
KPLDGSGSPHSKAQNQQT	1333	2.488	KTINGSGSPHSLVFNQQT	1690	2.019
KPLRGSGSPHSKAQNQQT	1334	2.486	KTINGSGSPHSKAQATRT	1691	2.019
KTINGSGSPHSKAQNAKL	1335	2.486	KTINGTSSPHCKAQNQQT	1692	2.017
KTINGSGSPHSKLSKQQT	1336	2.482	KTINGSGSPHSKALWRQT	1693	2.015
KTINGSGSPHSKARNGQT	1337	2.482	KTINGSGSPHSKAQFSVT	1694	2.015
KTINGSGSPHSKAQRRQT	1338	2.479	KTINGSGSPHSKLYMQQT	1695	2.015
KTINGSGSPHSKWPGQQT	1339	2.477	KTINGSLSPHYMAQNQQT	1696	2.015
KTINGSGSPHAFLQNQQT	1340	2.475	KTINGSGSPHSKAWLMQT	1697	2.015
KTINGILSPHRKAQNQQT	1341	2.475	KTINGSGSPHSKSLKQQT	1698	2.013
KTINGSGSPHSWGSNQQT	1342	2.473	KTINGSGSPHSKAQNTRR	1699	2.013
KTINGSGSPHSSCLNQQT	1343	2.471	KTINGSGSPHYLLQNQQT	1700	2.011
KTINGSGSPHSKAQSVKT	1344	2.464	KTINGSGSPHTWSQNQQT	1701	2.011
KTINGSGSPHSLRYNQQT	1345	2.462	KTINGSGSPHSKTRMQQT	1702	2.011
KTINGSGSPHSKARKLQT	1346	2.458	KTINTRPSPHSKAQNQQT	1703	2.011
KHRSGSGSPHSKAQNQQT	1347	2.456	KTINGSGSPHSKAQILVT	1704	2.009
KTINGSGSPHSKWSLQQT	1348	2.453	KTINGSGSPHSKAQNAKS	1705	2.009
KTINGSGSPHSKAQTMRT	1349	2.453	KTINGSGSPHSRTYNQQT	1706	2.009

KTINGSGSPHSKTIRQQT	1350	2.453	KTINGSGSPHSKKGGQQT	1707	2.009
KTINGKLSPHMKAQNQQT	1351	2.449	KTINGYSSPHRKAQNQQT	1708	2.007
KTINGSGSPHSKARPFQT	1352	2.443	KTINGSGSPHWVSQNQQT	1709	2.007
KTINGSGSPHSKPRVQQT	1353	2.441	KTINGSGSPHSKARLAQT	1710	2.007
KTINGSGSPHSKAQVVLT	1354	2.438	KTINGMCSPHSKAQNQQT	1711	2.006
KTRRSSGSPHSKAQNQQT	1355	2.438	KTINGSGSPHSKSNKQQT	1712	2.006
KTINGSGSPHSKPSRQQT	1356	2.438	KTINGSGSPHSKAQFVLT	1713	2.006
KTINGSGSPHSVYRNQQT	1357	2.432	KTINGSISPHFVAQNQQT	1714	2.006
KTINGSGSPHSKTCSQQT	1358	2.430	KTINGSGSPHRRMQNQQT	1715	2.004
KTINGSPSPHRKAQNQQT	1359	2.430	KTINGSGSPHSKAWILQT	1716	2.004
KTINGRSSPHFKAQNQQT	1360	2.428	KTINGSGSPHSKAQGVKT	1717	2.002
KTINGSGSPHSKAQMVRT	1361	2.426	KTINGSGSPHSKAQFSLT	1718	2.000
KTINGSRSPHCSAQNQQT	1362	2.426	KTINMLRSPHSKAQNQQT	1719	2.000
KTINGSCSPHLRAQNQQT	1363	2.423	KTINGSGSPHSKAQLGKT	1720	2.000
KTINGSGSPHSKAQCLFT	1364	2.421	KTINGSGSPHSMYLNQQT	1721	2.000

[0513] As shown in **Table 14**, approximately 72 TTM-002 matured capsid variants demonstrated at least a 2-fold increase in expression relative to the non-matured TTM-002 control, with a few variants demonstrating greater than a three- to five-fold enrichment relative to the non-matured TTM-002 control. Also, across the peptides comprised within the TTM-002 matured capsid variants with the greatest fold-enrichment relative to the non-matured TTM-002 capsid in the brain, it was observed that the modifications in the variant sequences appeared in the region N-terminal to the SPH motif present within the capsid variant. This indicates that modifications that appeared to improve TTM-002 capsid tropism in the CNS of mice were skewed to the N-terminal portion of the peptide insertion in loop IV of the sequence. Additionally, a number of these N-terminal modifications that were incorporated into the matured TTM-002 capsid variants were negatively charged amino acids (particularly glutamic acid (E)).

Table 14. NGS fold-enrichment of TTM-002 matured AAV capsid variants in the brain of CD-1 Outbred mice

Peptide Sequence	SEQ ID NO:	Fold enrichment over TTM-002	Peptide Sequence	SEQ ID NO:	Fold enrichment over TTM-002
KTINGHDSPHVTDQNQQT	1722	5.20	KAEVGHDSPHKSGQNQQT	1760	2.15
KTINGHDSPHKRGQHRQT	1723	4.20	KMDAGHDSPHKSGQNQQT	1761	2.15
KMPEGHDSPHKSGQNQQT	1724	3.18	KVEWGHDSPHKSGQNQQT	1762	2.15
KMEGGHDSPHKSGQNQQT	1725	2.72	KAEQGHDSPHKSGQNQQT	1763	2.14
KMEYGHDSPHKSGQNQQT	1726	2.71	KLEWGHDSPHKSGQNQQT	1764	2.14
KAEWGHDSPHKSGQNQQT	1727	2.69	KTINGHPSPHYLGQNQQT	1765	2.14
KCEWGHDSPHKSGQNQQT	1728	2.68	KTINGHLSPHYYGQNQQT	1766	2.13
KANNGQDSPHKSGQNQQT	1729	2.67	KMELGHDSPHKSGQNQQT	1767	2.13
KTINGHDSPHLCGQNQQT	1730	2.59	KMETGHDSPHKSGQNQQT	1768	2.12
KIPEGHDSPHKSGQNQQT	1731	2.54	KMEAGHDSPHKSGQNQQT	1769	2.12
KADMGHDSPHKSGQNQQT	1732	2.48	KTINGHDSPHLLWQNQQT	1770	2.12
KTINGHLSPHYFGQNQQT	1733	2.41	KTINRQRSPHKSGQNQQT	1771	2.11
KIEYGHDSPHKSGQNQQT	1734	2.41	KIESGHDSPHKSGQNQQT	1772	2.11
KADYGHDSPHKSGQNQQT	1735	2.40	KTAKDHDSPHKSGQNQQT	1773	2.11

KIETGHDSPHKSGQNQQT	1736	2.38	KMEVGHDSPHKSGQNQQT	1774	2.11
KTINGHDSPHTNGQKQQT	1737	2.38	KCEIGHDSPHKSGQNQQT	1775	2.10
KMEWGHDSPHKSGQNQQT	1738	2.38	KATNGHDSPHKSGLNQQT	1776	2.10
KTINGHDSPHWLLQNQQT	1739	2.37	KMDGGHDSPHKSGQNQQT	1777	2.09
KCEYGHDSPHKSGQNQQT	1740	2.36	KQEVGHDSPHKSGQNQQT	1778	2.07
KRINGHDSPHKSGQKQQN	1741	2.34	KADQGHDSPHKSGQNQQT	1779	2.07
KMEIGHDSPHKSGQNQQT	1742	2.34	KTINGHESPHKSAQNHQT	1780	2.07
KLEYGHDSPHKSGQNQQT	1743	2.33	KTINGHDSPHKSAQNQWT	1781	2.07
KADWGHDSPHKSGQNQQT	1744	2.32	KNMNGHDSPHKSGQNTHS	1782	2.06
KIEIGHDSPHKSGQNQQT	1745	2.30	KTPWEHDSPHKSGQNQQT	1783	2.05
KTIKDNDSPHKSGQNQQT	1746	2.27	KTINGHSSPHYFGQNQQT	1784	2.05
KDIMGHDSPHKSGQNQQT	1747	2.23	KIEMGHDSPHKSGQNQQT	1785	2.05
KFEQGHDSPHKSGQNQQT	1748	2.22	KTANEHDSPHKSGQNQQT	1786	2.05
KMEFGHDSPHKSGQNQQT	1749	2.21	KTINGHDSPHKSGRRRQT	1787	2.04
KCDQGHDSPHKSGQNQQT	1750	2.21	KISNGHDSPHKSAQNQQT	1788	2.03
KLPEGHDSPHKSGQNQQT	1751	2.19	KTGNGHDSPHKSGQYQQT	1789	2.03
KIENGHDSPHKSGQNQQT	1752	2.19	KTINGHYSPHLFGQNQQT	1790	2.02
KMESGHDSPHKSGQNQQT	1753	2.18	KTINGNYSPHKIGQNQQT	1791	2.02
KAEIGHDSPHKSGQNQQT	1754	2.17	KTINGHDSPHKSRQNDQT	1792	2.01
KVEYGHDSPHKSGQNQQT	1755	2.17	KQQQGHDSPHKSGQNQQT	1793	2.01
KIINGHDSPHKSGLTQQT	1756	2.17	KTPQDHDSPHKSGQNQQT	1794	2.00
KTSNGDDSPHKSGRNQQT	1757	2.17	KHDWGHDSPHKSGQNQQT	1795	2.00
KIEVGHDSPHKSGQNQQT	1758	2.16	KIEGGHDSPHKSGQNQQT	1796	2.00
KMEMGHDSPHKSGQNQQT	1759	2.16			

[0514] These data demonstrate that following two maturation approaches, matured TTM-001 and TTM-002 capsid variants with loop IV modifications were generated with significantly enhanced CNS tropism in mice compared to the corresponding non-matured TTM-001 and TTM-002 capsid variants, which already exhibited a significant fold enrichment over AAV9 in the mouse brain.

Example 4. Maturation of TTM-001 and TTM-002 Capsid in NHPs

[0515] This Example describes maturation of the AAV9 capsid variants, TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941 (encoded by SEQ ID NO: 942)) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2 (encoded by SEQ ID NO: 944)) in NHPs to further enhance their transduction and biodistribution in the central nervous system as well as other tissues, and evolve the AAV capsid variants to provide further cross-species compatibility. Two approaches were used to mature the TTM-001 and TTM-002 capsid sequences in order to randomize and mutate within and around the peptide insert comprised within loop IV of the capsid variant. As many of the AAV capsid variants that demonstrated the greatest fold-enrichment in the NHP brain relative wild-type AAV9 comprised an SPH motif in the same position (e.g., immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138) (see **Example 1**), the SPH motif was not mutated in either approach to mature the TTM-001 and TTM-002 capsid variants. In the first maturation approach, sets of three contiguous amino acids were randomized across the mutagenesis region in the TTM-001 and TTM-002 sequences, which spanned from position 450 to position 466,

numbered according to SEQ ID NO: 981 and 982. In the second maturation approach, mutagenic primers were used to introduce point mutations at a low frequency, scattered across the mutagenesis region in the TTM-001 and TTM-002 sequences ranging from position 449 to position 466, numbered according to SEQ ID NO: 981 and 982. AAV capsid variants arising from each maturation approach for TTM-001 and TTM-002 were pooled together, for subsequent testing and characterization in NHPs.

[0516] The library of pooled matured AAV capsid variants generated using the first maturation approach and the second maturation approach for the TTM-001 and TTM-002 AAV capsid variants were injected into two NHPs. After a period in life, the brains, heart, liver, muscle, and DRG of the NHPs were isolated and RNA was extracted. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate the fold enrichment ratio relative to an AAV9 control, and the peptides comprised within the variants were identified.

[0517] Following the RNA recovery and NGS analysis from the second maturation approach, approximately 680,000 capsid variants were identified. The 680,000 matured capsid variants were then filtered based on samples with a raw virus count greater than 10 and a coefficient of variance (CV) of less than 1, which was calculated for each peptide across the brain samples taken from the two NHPs. Those that had a CV value <1 were identified, as these were the peptides that were reliably detected in the majority of samples isolated from the brains of the two NHPs. Using this filtering criteria, this led to approximately 64,000 matured capsid variants.

[0518] **Table 15** provides the peptide sequences of the matured capsid variants having a raw virus count greater than 10, a CV of less than 1 for the brain samples isolated, and that also demonstrated a 50-fold or greater fold-increase in expression in the brain relative to the AAV9 control in both mice and NHPs. The matured variants in **Table 15**, were also those variants that had a fold-change in expression that was less than 2 relative to the AAV9 control in the liver and the DRG. Applying these criteria, approximately 350 matured capsid variants were identified that demonstrated high transduction in the brain in NHPs and mice, cross-species compatibility in mice and NHPs, and were de-targeted in the liver and DRG, relative to the AAV9 control. Several variants as shown in **Table 15**, led to greater than 100-fold increase in expression relative to AAV9 in the NHP and/or mouse brain, with one variant resulting in a greater than 200-fold increase in expression relative to AAV9 in both species.

[0519] Fold-change in expression for the TTM-001 and TTM-002 matured variants in **Table 15** that showed increased expression in the brain of the NHPs and mice, were also calculated for the DRG, muscle, liver (RNA and DNA), and heart of the NHPs following each maturation approach. As shown in **Table 15**, many variants were de-targeted in the peripheral tissues with a lower fold-change in expression relative to the AAV9 control, demonstrating CNS-specific tropism and a preferential

transduction of the brain and CNS. Some variants demonstrated increased expression to AAV9 in multiple tissues, including the brain and peripheral tissues, demonstrating pan-tropism.

Table 15. NGS fold-enrichment of TTM-001 and TTM-002 matured AAV capsid variants in the brain of NHPs and mice

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9						
		Brain (NHP)	DRG (NHP)	Heart (NHP)	Muscle (NHP)	Liver RNA (NHP)	Liver DNA (NHP)	Brain (Mouse)
KTIIGSGSPHSKAQNRHT	3239	217.176	0.000	0.000	0.000	0.000	0.000	210.515
KTFPGSGSPHSKVQNQQT	3240	199.720	0.000	0.000	0.000	0.000	0.967	97.703
KTEKMSGSPHSKAQNQQT	3241	169.461	0.523	0.000	0.000	0.000	0.158	109.161
KEINGRGSPHSKAQNQQT	3527	134.390	0.239	0.000	0.000	0.000	0.232	52.311
KTVNRNGSPHSKAQNQQT	3528	133.016	0.000	0.416	0.000	0.000	0.000	85.361
KTVNGSGSPHSKARDQQT	3242	124.789	0.123	0.039	0.312	0.569	0.454	132.137
KTFNGSGSPHSKAPNLQT	3243	121.436	0.000	0.167	0.000	0.000	0.015	168.920
KTEKTSGSPHSKAQNQQT	3244	120.337	0.000	0.355	0.000	0.000	0.119	101.467
KTINGSGSPHSKAHVRQT	3245	119.798	0.000	0.000	0.262	0.694	1.039	165.590
KTVNGSGSPHSKAPNQHT	3246	117.207	0.000	0.109	0.000	0.000	0.074	51.008
KTEKISGSPHSKAQNQQT	3247	116.603	0.000	0.000	0.000	0.000	0.426	102.978
KTINGPGSPHSKAHNQQT	3529	115.742	0.146	0.000	0.235	0.000	0.513	52.508
KTVNGSGSPHSKTQSQQT	3248	115.086	0.000	0.726	0.000	0.000	0.340	63.248
TTINGSGSPHSKAQNQQT	3249	114.856	1.340	14.856	0.827	1.281	0.957	72.058
KSINESGSPHSKAQNQQI	3250	113.833	0.000	0.000	0.000	0.000	0.000	67.649
KTERTSGSPHSKAQNQQT	3251	112.957	0.000	0.009	0.000	1.128	0.207	117.374
KTINGSGSPHSKAQPAKT	3252	111.472	0.331	0.000	1.089	0.044	1.796	215.275
KTEKSSGSPHSKAQNQQT	3253	107.470	0.000	0.016	0.014	0.977	0.179	100.177
KTSYGNGSPHSKAQNQQT	3530	105.937	0.000	0.000	0.000	0.000	0.114	105.894
KTEKGSGSPHSKAQNQQT	3254	105.614	0.053	0.031	0.000	0.586	0.169	84.653
KTINGSGSPHSKSQTQQN	3255	104.474	0.000	0.131	0.000	0.084	0.038	54.021
KTERISGSPHSKAQNQQT	3256	103.692	0.000	0.000	0.000	0.062	0.370	89.637
KTERASGSPHSKAQNQQT	3257	103.669	0.000	0.000	0.000	0.127	0.070	115.550
KELHGSGSPHSKAQNQQT	3258	102.680	0.000	0.000	0.000	1.634	0.592	96.554
KAINGSGSPHSKAQNLAT	3259	101.954	0.000	10.954	8.655	0.298	0.239	116.685
KTVNGSGSPHSKSQNQLT	3260	101.327	0.000	0.035	0.000	0.000	0.025	80.716
KTERNSGSPHSKAQNQQT	3261	99.892	0.000	0.000	0.000	0.000	0.107	87.392
KSVNGNGSPHSKAQNQQT	3531	99.385	0.000	1.329	0.000	0.359	0.079	51.016
KTFNGSGSPHSKAQGQQT	3262	99.253	0.000	0.208	0.000	0.128	0.099	81.459
KTINGSGSPHGWVQNQQT	3532	97.122	0.000	0.000	0.000	1.240	1.975	290.720
KTERVSGSPHSKAQNQQT	3263	96.943	0.000	0.000	0.000	0.000	0.144	135.438
KTINGSGSPHSKALNRQS	3264	96.843	0.136	0.532	0.000	0.042	0.178	55.945
KTERLSGSPHSKAQNQQT	3265	95.857	0.000	0.004	0.005	0.126	0.260	102.372
KTDNGSGSPHSKAHNQQT	3266	95.164	0.000	0.000	0.000	0.000	0.027	55.313
KTFHGSGSPHSKTQNQQT	3267	94.714	0.000	0.210	0.120	0.000	0.000	51.119
KTINGGGSPHSKAQTQQI	3533	92.345	0.000	0.000	0.000	0.000	0.023	54.199
KTSNGSGSPHSKAQNPPT	3268	91.528	0.000	0.000	0.000	0.000	0.039	51.541
ETINGSGSPHSKAQNLQT	3269	90.969	0.221	1.023	0.197	0.179	0.813	107.216
KTVHGNGSPHSKAQNQQT	3534	90.073	0.000	0.000	0.000	0.000	0.304	97.003
NTINGSGSPHSKAQNQQT	3270	90.017	1.712	1.261	1.171	0.923	0.540	55.179
KTINGGGSPHSKAQNQQC	3535	89.301	0.219	0.000	0.000	0.287	0.319	53.840
KTENMSGSPHSKAQNQQT	3271	89.247	0.000	0.000	0.000	0.000	0.260	130.568
KTENVSGSPHSKAQNQQT	3272	88.506	0.000	0.000	0.000	0.964	0.112	108.591
KTSSGSGSPHSKAQYQQT	3273	87.304	0.000	0.000	0.000	0.000	0.299	58.143
KTIDGGGSPHSKAQNKQT	3536	85.019	0.000	0.000	0.000	0.000	0.477	55.517
KTEKVSGSPHSKAQNQQT	3274	84.558	0.000	0.022	0.000	0.873	0.424	112.185

KAINGSGSPHSKAQDQET	3275	84.080	0.000	0.000	0.000	0.194	0.027	87.637
KTCNKSGSPHSKAQNQQT	3276	83.992	0.000	0.000	0.165	0.283	0.000	119.496
KTINGGGSPHSKAQNQLI	3537	83.881	0.000	0.000	0.000	0.046	0.387	78.383
KNINGGGSPHSKAQNQQT	3538	83.083	0.000	0.042	0.000	0.000	0.000	75.913
KTEHLSGSPHSKAQNQQT	3277	83.080	0.000	0.000	0.012	0.021	0.189	69.494
KAEMGSGSPHSKAQNQQT	3278	83.049	0.000	0.020	0.000	0.768	0.112	135.019
KATNGSGSPHSKAQNHQT	3279	82.627	0.000	0.176	0.000	0.155	0.057	66.207
KAIKGSGSPHSKAQDQQT	3280	82.258	0.000	0.000	0.000	0.108	0.000	85.178
KTINGGGSPHSKSQNQLT	3539	82.231	0.000	0.070	0.000	0.000	0.498	126.986
KTVNGNGSPHSKAQNKQT	3540	81.481	0.000	0.000	0.000	0.000	0.122	69.455
KTINGSGSPHSKGHWQQT	3281	81.434	0.000	0.000	0.000	0.000	1.011	65.252
KTDKTSGSPHSKAQNQQT	3282	81.430	0.000	0.000	0.000	1.362	0.291	169.515
KTFKGSGSPHSKAPNQQT	3283	80.890	0.000	0.000	0.000	0.000	0.017	71.144
KTVNGSGSPHSKAQNQLI	3284	80.509	0.000	0.000	0.000	0.000	0.166	71.156
KTINGSGSPHSKRPEQQT	3285	80.418	0.000	0.013	0.000	0.149	0.361	50.319
KTINGSGSPHSKAQRTMT	3286	80.388	0.000	0.022	0.170	1.812	1.025	100.248
KTEKASGSPHSKAQNQQT	3287	80.285	0.000	0.041	0.000	0.000	0.261	90.390
KSDQGSGSPHSKAQNQQT	3288	80.076	0.000	0.000	0.000	0.993	0.124	151.911
KTEITSGSPHSKAQNQQT	3289	79.620	0.000	0.163	0.000	0.332	0.074	76.686
KTDKSSGSPHSKAQNQQT	3290	79.470	0.055	0.012	0.000	1.437	0.367	141.351
KTIDGSGSPHSKAQNQQH	3291	79.090	0.000	0.000	0.000	0.136	0.049	57.914
KTVNGNGSPHSKAQNQHT	3541	78.849	0.000	0.000	0.000	0.000	0.045	54.086
KNTNGSGSPHSKAQNQQT	3292	78.445	0.000	0.000	0.000	0.571	0.177	89.719
KTETHSGSPHSKAQNQQT	3293	77.974	0.000	0.067	0.000	0.000	0.512	57.287
KTINGGGSPHSKALNQQN	3542	77.822	0.000	0.131	0.000	0.000	0.274	69.884
KTINGSGSPHSKALHQHT	3294	77.502	0.000	0.052	0.041	0.000	0.188	68.196
KTINGTGSPHSKAQNHQI	3543	77.089	0.171	0.000	0.000	0.000	0.166	54.281
KTINGSGSPHSKAQHRIT	3295	76.849	0.105	0.499	0.170	1.424	0.214	127.000
KTINGSGSPHSKAQYIHT	3296	76.170	0.000	0.014	0.033	1.523	0.168	59.649
KTENISGSPHSKAQNQQT	3297	76.072	0.000	0.000	0.000	0.115	0.132	83.118
KTIIGGGSPHSKAHNQQT	3544	75.872	0.000	0.050	0.000	0.000	0.235	65.492
KTINGSGSPHSKAQKFET	3298	75.788	0.000	0.000	0.028	0.108	0.093	65.588
KTSNESGSPHSKAQNHQT	3299	75.720	0.000	0.000	0.000	0.169	0.217	70.590
KTINGSGSPHSKAQFPST	3300	75.677	0.000	0.004	0.000	0.849	0.127	119.712
KTERPSGSPHSKAQNQQT	3301	75.669	0.000	0.029	0.000	0.000	0.156	73.894
KTINGNGSPHSKAQNPLT	3545	75.269	0.000	0.000	0.000	0.366	0.000	53.583
KSIKGNGSPHSKAQNQQT	3546	75.196	0.000	0.000	0.000	0.000	0.000	90.251
KTERMSGSPHSKAQNQQT	3302	74.910	0.000	0.000	0.000	0.100	0.151	122.812
KTERSSGSPHSKAQNQQT	3303	74.853	0.000	0.071	0.000	1.036	0.056	125.538
KTELHSGSPHSKAQNQQT	3304	74.620	0.000	0.000	0.000	0.021	0.089	53.124
KTELTSGSPHSKAQNQQT	3305	74.548	0.000	0.000	0.000	0.537	0.421	100.311
KTINGSGSPHSKAHNQQR	3306	74.272	0.562	0.486	0.047	0.956	0.057	107.301
KTINGGGSPHSKAQSQQI	3547	74.264	0.000	0.000	0.000	0.000	0.235	67.651
KTINGSGSPHSKAQAIKT	3307	74.261	0.255	0.000	0.000	0.186	0.132	73.560
KTENTSGSPHSKAQNQQT	3308	74.061	0.000	0.000	0.218	0.233	0.730	96.249
KTIDGSGSPHSKGQNRQT	3309	73.930	0.000	0.000	0.000	0.106	0.091	63.626
KNINGSGSPHSKAQSQQT	3310	73.757	0.000	0.000	0.000	0.000	0.041	57.432
KTINGSVSPHGKAQNQLT	3548	73.525	0.000	0.061	0.067	0.000	0.053	51.358
KTSNASGSPHSKAQNQLT	3311	73.501	0.000	0.000	0.297	0.000	0.313	150.401
KTEARSGSPHSKAQNQQT	3312	73.349	0.000	0.000	0.000	0.695	0.118	62.903
KTEKNSGSPHSKAQNQQT	3313	73.347	0.000	0.000	0.044	0.159	0.021	74.393
KTANGSGSPHSKAQYQQT	3314	73.038	0.000	0.000	0.000	0.153	0.160	139.451
KTVNGSGSPHSKAQYQHT	3315	72.847	0.000	0.000	0.000	0.000	0.130	54.158
KTINGSGSPHTKAQNPQS	3316	72.594	0.000	0.000	0.000	0.000	0.130	62.508
KTINGSGSPHSKGQNPPT	3317	72.339	0.000	0.206	0.000	0.000	0.041	134.808
KTIIGSGSPHSKAQHQLT	3318	72.291	0.000	0.000	0.000	0.000	0.000	100.144
KTINGSGSPHSKAQSPPT	3319	71.632	0.069	0.047	0.274	0.179	0.425	97.111
NTIYGSGSPHSKAQNQQT	3320	71.267	1.739	0.000	273.69	0.000	0.209	59.707

KTINGSGSPHSKAQAKLT	3321	71.154	0.000	0.273	0.017	1.591	0.777	130.132
KTDKNSGSPHSKAQNQQT	3322	70.964	0.000	0.000	0.000	0.070	0.123	62.932
KTINGSGSPHSKTKSQQT	3323	70.891	0.000	0.568	0.045	0.418	0.496	83.923
KTINGSGSPHSKAQDRPT	3324	70.831	0.132	0.006	0.000	0.039	0.379	66.800
KTINGIGSPHSKAQNLGT	3549	70.543	0.000	0.071	0.000	0.000	0.135	104.769
KTINGSGSPHSKAQSQQL	3325	70.539	0.000	0.000	0.000	0.000	0.041	51.126
KTENLSGSPHSKAQNQQT	3326	70.303	0.070	0.000	0.000	0.395	0.470	107.385
KTINGSGSPHSKAQAFHT	3327	70.159	0.033	0.000	0.058	0.762	0.119	86.268
KTINGSGSPHSKAQKQQD	3328	70.116	0.000	0.024	0.000	0.064	0.083	110.196
KTFSGSGSPHSKAQNLQT	3329	70.035	0.000	0.327	0.303	0.000	0.228	70.917
KAINGSGSPHSKAQNAQT	3330	69.651	0.000	0.000	0.000	0.023	0.142	72.160
KTESWSGSPHSKAQNQQT	3331	69.144	0.000	0.000	0.000	0.000	0.019	67.699
KTTNGSGSPHSKAHNQLT	3332	69.062	0.000	0.000	0.000	0.708	0.000	65.505
KTVNGNGSPHSKAQNHQT	3550	68.889	0.000	0.000	0.000	0.000	0.030	52.482
KTEDKSGSPHSKAQNQQT	3333	68.813	0.000	0.000	0.000	0.000	0.000	70.071
KTESASGSPHSKAQNQQT	3334	68.651	0.000	0.000	0.000	0.274	0.084	80.500
KTNNGSGSPHSKAQNQQY	3335	68.530	0.000	0.040	0.000	0.000	0.059	82.656
KTSNGGGSPHSKAQNLQT	3551	68.311	0.000	0.052	0.000	0.000	0.000	124.871
KTDKMSGSPHSKAQNQQT	3336	68.167	0.000	0.000	0.000	0.017	0.205	88.234
KEVHGSGSPHSKAQNQQT	3337	67.901	0.000	0.000	0.000	0.727	0.000	100.111
KTINGSGSPHSKAQKLNT	3338	67.782	0.073	0.092	0.000	1.232	0.201	68.637
KTINGGGSPHSKSQNQHT	3552	67.773	0.000	0.057	0.000	0.000	0.220	100.748
KTVNGGGSPHSKAQSQQT	3553	67.634	0.000	0.055	0.000	0.000	0.210	160.711
KTTNGSGSPHSKAQYQHT	3339	67.325	0.000	0.000	0.000	1.378	0.080	83.337
KTISGSGSPHSKAQYQHT	3340	66.739	0.000	0.000	0.000	0.000	0.191	59.822
KTESTSGSPHSKAQNQQT	3341	66.649	0.000	0.009	0.000	1.688	0.176	95.861
KTINGSGSPHSKSQNVQT	3342	66.627	0.000	0.190	0.000	0.202	0.188	56.672
KSINGSGSPHSKAQAQQT	3343	66.464	0.000	0.711	0.000	0.148	0.111	78.451
KTVNGSGSPHSKAQNLQA	3344	66.379	0.000	0.000	0.000	0.000	0.132	50.934
KTVRDSGSPHSKAQNQQT	3345	66.056	0.000	0.025	0.000	0.129	0.461	142.600
KTFNASGSPHSKAPNQQT	3346	65.392	0.208	0.000	0.000	0.215	0.156	66.275
KTDRMSGSPHSKAQNQQT	3347	65.143	0.000	0.000	0.000	0.332	0.103	104.890
KTINGSGSPHSKAQTPPT	3348	64.657	0.010	0.015	0.014	0.200	0.207	54.179
ETIKGSGSPHSKAQNQQT	3349	64.609	0.000	0.000	0.144	0.000	0.024	67.201
KNHIGSGSPHSKAQNQQT	3350	64.535	0.000	0.000	0.000	1.253	0.187	70.356
KTINGSGSPHSKAQYQHA	3351	64.435	0.000	0.000	0.024	0.993	0.097	57.278
KTIPIDGSPHSKAQNQQT	3554	64.421	0.000	0.047	0.000	0.234	0.936	76.826
KTINGSGSPHSKAQGQQA	3352	64.128	0.000	0.185	0.000	0.063	0.195	64.116
KTFNGSGSPHNKAQNHQT	3353	64.060	0.000	0.000	0.035	0.094	0.317	67.757
KESDGSGSPHSKAQNQQT	3354	63.766	0.000	0.000	0.000	0.567	0.146	115.231
KTINGSGSPHSKAQPPAT	3355	63.510	0.048	0.030	0.031	0.126	0.302	117.453
KTINGSGSPHSKAQERPT	3356	63.460	0.000	0.011	0.000	0.810	0.173	57.506
KTIKGSGSPHSKAQDLQT	3357	63.260	0.000	0.000	0.000	0.000	0.218	58.576
KTDLKSGSPHSKAQNQQT	3358	63.152	0.000	0.000	0.012	0.285	0.377	62.687
KTINGGGSPHSKAQNPPT	3555	63.041	0.000	0.082	0.000	0.000	0.057	64.045
KTINGSGSPHSKAQAMHT	3359	62.756	0.000	0.000	0.010	0.976	0.393	84.056
KTVPNSGSPHSKAQNQQT	3360	62.540	0.000	0.000	0.011	0.202	0.161	93.793
KTVIGSGSPHSKALNQQT	3361	62.358	0.000	0.310	0.000	0.062	0.245	60.369
KTINGSGSPHSKAQHPST	3362	62.255	0.000	0.044	0.000	1.345	0.301	101.103
KTINGLGSPHSKSQNQQT	3556	62.170	0.000	0.157	0.000	0.146	0.107	64.139
KTINGTGSPHSKAQNQQM	3557	62.151	0.000	0.000	0.000	0.000	0.000	62.376
KTINGSGSPHSKAPGLQT	3363	62.043	0.007	0.000	0.005	0.651	0.210	144.610
KTINGSGSPHSKAQGIRT	3364	61.952	0.041	0.000	0.012	0.897	0.502	155.013
KTESHSGSPHSKAQNQQT	3365	61.947	0.000	0.000	0.000	1.480	0.106	52.506
KTINGSGSPHSKAQAPAT	3366	61.934	0.000	0.169	0.015	0.696	0.197	127.420
KTINGSGSPHSKSQSQQI	3367	61.870	0.000	0.000	0.000	0.200	0.175	64.027
KAEHGSGSPHSKAQNQQT	3368	61.830	0.000	0.000	0.000	0.772	0.184	116.201
KTEDRSGSPHSKAQNQQT	3369	61.756	0.000	0.000	0.000	1.004	0.408	66.887

KNCLGSGSPHSKAQNQQT	3370	61.442	0.000	0.036	0.000	1.849	0.026	82.488
KTDRGSGSPHSKAQNQQT	3371	61.419	0.000	0.004	0.000	0.211	0.316	74.256
KTINGSGSPHSKAQIPPT	3372	61.258	0.000	0.000	0.000	0.758	0.115	87.661
KTVKGSGSPHSKAQDQQT	3373	61.175	0.000	0.041	0.000	0.432	0.090	58.114
KNADGSGSPHSKAQNQQT	3374	60.944	0.000	0.000	0.000	1.239	0.085	104.503
KTDKVSGSPHSKAQNQQT	3375	60.935	0.000	0.015	0.000	0.765	0.128	146.657
KTITGSGSPHSKAQTQLT	3376	60.846	0.160	8.992	0.000	0.000	0.000	55.640
KTINGSGSPHSKAQAPST	3377	60.696	0.200	0.005	0.000	0.751	0.263	115.528
KNCVGSGSPHSKAQNQQT	3378	60.535	0.000	0.000	0.000	0.018	0.282	96.175
KTIRDAGSPHSKAQNQQT	3558	60.346	0.000	0.000	0.000	0.141	0.251	113.179
KTVKDSGSPHSKAQNQQT	3379	60.216	0.000	0.019	0.000	0.443	0.251	87.334
KNALGSGSPHSKAQNQQT	3380	60.014	0.000	0.003	0.000	0.682	0.213	137.222
KVINGSGSPHSKGQNQQT	3381	60.001	0.000	0.000	0.031	0.264	0.157	68.532
KTVNGGGSPHSKAQNQQS	3559	59.871	0.062	0.020	0.000	0.080	0.185	61.847
KTIQDGGSPHSKAQNQQT	3560	59.865	0.000	0.000	0.116	1.435	0.789	87.522
KTISGGGSPHSKAQNQQN	3561	59.801	0.000	0.000	0.000	0.722	0.039	87.761
KTSNASGSPHSKAHNQQT	3382	59.607	0.000	0.078	0.067	0.031	0.050	67.967
KTINGSGSPHSKAQNTYA	3383	59.603	0.000	0.000	0.000	0.425	0.346	101.715
KTINGSGSPHSKSQNQHI	3384	59.438	0.000	0.099	0.000	0.111	0.108	76.025
KTINGGGSPHSKAQDKQT	3562	59.322	0.000	0.000	0.000	0.000	0.093	50.764
KTEFVSGSPHSKAQNQQT	3385	59.306	0.000	0.000	0.000	0.196	0.276	69.788
KTVNGSGSPHSKAQNHLT	3386	59.239	0.133	0.034	0.000	0.000	0.156	70.786
KTREISGSPHSKAQNQQT	3387	59.027	0.000	0.042	0.224	0.356	0.269	51.696
KTINGSGSPHSKAQIGMT	3388	59.013	0.081	106.528	0.000	1.003	0.248	134.585
KTIDGSGSPHSKALNKQT	3389	58.992	0.000	0.267	0.000	0.000	0.056	74.626
KTIIGGGSPHSKAQNPQT	3563	58.924	0.000	0.202	0.000	0.000	0.126	53.992
KQGEGSGSPHSKAQNQQT	3390	58.752	0.000	0.000	0.000	0.000	0.151	135.300
KTINGTGSPHSKAPNQLT	3564	58.738	0.000	0.000	0.000	0.229	0.035	86.939
KTVNGSGSPHSKAQLQQT	3391	58.681	0.315	0.465	0.045	0.529	0.333	81.201
KTFNGGGSPHSKAQYQQT	3565	58.609	0.000	0.000	0.000	0.163	0.045	72.618
KSINGSGSPHSKTQSQQT	3392	58.608	0.000	3.017	0.000	0.155	0.017	71.397
KTVNGGGSPHSKAQHQQT	3566	58.602	0.729	0.000	0.000	0.000	0.043	138.544
KSEKGSGSPHSKAQNQQT	3393	58.566	0.000	0.010	0.011	1.601	0.059	158.931
KNVNGSGSPHSKAQNQQT	3394	58.481	0.000	0.000	0.000	0.917	0.166	53.379
KGGEGSGSPHSKAQNQQT	3395	58.472	0.000	0.034	0.000	0.037	0.066	91.023
KTINGSGSPHSKAQRMST	3396	58.435	0.192	0.037	0.000	1.707	0.882	53.414
KTINGSGSPHSKAQGILT	3397	58.418	0.000	0.005	0.010	0.569	0.192	102.631
KEFVGSGSPHSKAQNQQT	3398	58.374	0.000	0.046	0.000	0.088	0.326	128.675
KTIIGSGSPHSKAQDRQT	3399	58.258	1.393	0.230	0.219	0.000	0.045	53.981
KSDKGSGSPHSKAQNQQT	3400	58.248	0.000	0.000	0.000	0.076	0.166	146.566
KTEQVSGSPHSKAQNQQT	3401	58.247	0.000	0.000	0.000	0.000	0.081	88.487
KTEHVSGSPHSKAQNQQT	3402	58.228	0.000	0.024	0.000	0.433	0.141	71.410
KTINGSGSPHSKARDWQT	3403	58.216	0.000	0.005	0.000	0.800	0.259	120.704
KTENASGSPHSKAQNQQT	3404	58.187	0.000	0.038	0.000	0.371	0.129	88.439
KEVQGSGSPHSKAQNQQT	3405	58.125	0.000	0.000	0.000	0.657	0.000	168.220
KTINGSGSPHSKAQNTHD	3406	58.108	0.000	0.027	0.000	0.410	0.126	81.189
KTINGSGSPHSKAPNLQI	3407	58.022	0.000	0.044	0.000	1.548	0.243	55.714
KTINGSGSPHSKAQERST	3408	58.021	0.000	0.011	0.005	0.829	0.409	87.656
KTSNGSGSPHSKAQNYQT	3409	57.894	0.000	0.082	0.000	0.000	0.110	63.681
KTEYISGSPHSKAQNQQT	3410	57.891	0.000	0.000	0.000	0.076	0.075	57.620
KTINGSGSPHSKAQRTCT	3411	57.863	0.000	0.140	0.129	1.855	1.716	90.146
KTINGSGSPHSKAQIGHT	3412	57.769	0.024	0.000	0.000	0.281	0.154	99.262
KNCWGSGSPHSKAQNQQT	3413	57.756	0.000	0.000	0.000	0.000	0.092	59.888
KTINGSGSPHSKAQGAIT	3414	57.627	0.000	0.000	0.000	0.594	0.161	95.696
KTDVNSGSPHSKAQNQQT	3415	57.593	0.000	0.000	0.000	0.000	0.331	66.127
KSDIGSGSPHSKAQNQQT	3416	57.592	0.000	0.000	0.000	0.844	0.128	107.342
KTINGSGSPHSKAQVPPT	3417	57.316	0.000	0.012	0.000	0.257	0.200	90.220
KTINGSGSPHSKAQVQQI	3418	57.308	0.000	1.113	0.000	0.000	0.113	61.957

KTINGSGSPHSKALMRQT	3419	57.234	0.060	0.036	0.100	1.798	0.517	81.332
KTINGSGSPHSKAQYSVT	3420	57.130	0.000	0.093	0.000	1.235	0.302	60.023
KNSIGSGSPHSKAQNQQT	3421	57.101	0.000	0.052	0.000	0.083	0.074	97.381
KTINGSGSPHSKVPNLQT	3422	57.046	0.000	0.029	0.000	0.459	0.082	50.474
KAINGSGSPHSKAQSQQI	3423	56.976	0.000	0.000	0.000	0.000	0.000	57.052
KTINGSGSPHSKAQAITT	3424	56.924	0.000	0.000	0.000	1.239	0.438	75.250
KTINGSGSPHSKAQKTLT	3425	56.844	0.000	0.017	0.009	1.800	1.400	66.415
KTVNGSGSPHSKAQNQWT	3426	56.823	0.000	0.000	0.299	0.000	0.219	69.906
KTINGSGSPHSKAQLHHT	3427	56.815	0.025	0.000	0.010	0.712	0.368	58.418
KTEQTSGSPHSKAQNQQT	3428	56.683	0.000	0.045	0.000	0.792	0.430	59.360
KTINGSGSPHSKAQNIII	3429	56.630	0.000	0.062	0.123	0.099	0.056	76.742
KNSLGSGSPHSKAQNQQT	3430	56.621	0.000	0.028	0.000	0.308	0.162	101.942
KTIPMEGSPHSKAQNQQT	3567	56.560	0.000	0.000	0.000	1.824	0.371	89.951
KTINGSGSPHSKAQGHHT	3431	56.559	0.000	0.000	0.000	0.632	0.117	71.050
KTDRTSGSPHSKAQNQQT	3432	56.466	0.000	0.000	0.000	0.062	0.160	148.498
KTINGSGSPHSKAQSKVT	3433	56.373	0.000	0.050	0.014	1.021	0.390	76.115
KEVVGSGSPHSKAQNQQT	3434	56.371	0.000	0.000	0.000	0.000	0.323	116.964
KTINGSGSPHSKAQLPST	3435	56.238	0.005	4.258	0.001	1.040	0.185	84.918
KTINGSGSPHSKAIGKQT	3436	56.158	0.000	0.000	0.000	0.887	0.088	110.132
KTEPTSGSPHSKAQNQQT	3437	56.134	0.000	0.000	0.000	0.061	0.527	143.397
KTVNGGGSPHSKSQNQQT	3568	56.114	0.116	0.000	0.000	0.000	0.040	170.548
KTINGSGSPHSKAQAIHT	3438	56.047	0.000	0.000	212.32	0.887	0.890	81.908
KTINGSGSPHSKAQHGLT	3439	55.999	0.000	0.000	0.101	1.913	0.244	117.191
KSELGSGSPHSKAQNQQT	3440	55.997	0.000	0.005	0.000	0.881	0.239	120.521
KTINGSGSPHSKAQFMCT	3441	55.916	0.000	0.000	0.000	0.078	0.448	81.959
KTINVSGSPHSKAQGQQT	3442	55.870	0.000	0.191	0.000	0.592	0.040	87.211
KTINGGGSPHSKAQNQMT	3569	55.778	0.000	0.000	0.000	0.866	0.012	73.177
KTVNGSGSPHSKAQHLQT	3443	55.739	0.091	0.036	0.000	0.062	0.409	62.743
KTIRENGSPHSKAQNQQT	3570	55.605	0.000	0.000	0.016	0.131	0.257	95.931
KTINGSGSPHSKTQNHQN	3444	55.551	0.000	0.048	0.000	0.000	0.099	64.846
KTINGSGSPHSKAQPART	3445	55.513	0.000	0.000	0.328	1.294	0.991	127.301
KTVNGSGSPHSKAQSLQT	3446	55.497	0.000	0.060	0.000	0.000	0.143	69.033
KTINGSGSPHSKSQSQLT	3447	55.430	0.000	0.035	0.000	0.050	0.013	125.577
KTINGSASPHSKAHSQQT	3571	55.293	0.000	0.000	0.000	0.000	0.166	66.252
KTWQNSGSPHSKAQNQQT	3448	55.245	0.000	0.000	0.000	0.111	0.265	114.258
KTINGSGSPHSKAQDRQS	3449	55.137	1.146	0.016	0.106	0.644	0.086	55.701
KTINGSGSPHSKAQMPST	3450	54.986	1.691	0.039	0.028	0.450	0.202	114.331
KTNNGGGSPHSKAQNLQT	3572	54.963	0.000	0.000	0.000	0.000	0.089	80.506
KTINGSGSPHSKAQGSLT	3451	54.717	0.000	0.006	0.013	0.480	0.298	142.786
KTEVTSGSPHSKAQNQQT	3452	54.663	0.000	0.000	0.000	0.323	0.185	81.482
KSINGGGSPHSKAQYQQT	3573	54.612	0.000	0.000	0.000	0.105	0.010	65.952
KTVIGSGSPHSKSQNQQT	3453	54.603	0.000	0.000	0.000	0.000	0.106	69.121
KAVNVSGSPHSKAQNQQT	3454	54.586	0.000	0.000	0.000	0.000	0.023	57.835
KTVNGNGSPHSKSQNQQT	3574	54.586	0.000	0.000	0.000	0.256	0.168	95.384
KTDRNSGSPHSKAQNQQT	3455	54.495	0.000	0.000	0.000	0.823	0.241	85.823
KTINGSGSPHSKAQVPAT	3456	54.475	0.000	0.002	0.000	0.782	0.223	137.743
KGVLGSGSPHSKAQNQQT	3457	54.472	0.000	0.007	0.027	0.359	0.189	145.740
KTLNGNGSPHSKAQNLQT	3575	54.458	0.668	0.000	0.000	0.161	0.172	159.134
KAINGSGSPHSKAQDKQT	3458	54.452	0.000	0.000	0.057	0.044	0.223	56.004
KTSNGSGSPHSKAHYQQT	3459	54.414	0.000	0.251	0.000	0.249	0.204	54.162
KTINGSGSPHSKAQVPST	3460	54.366	0.000	1.001	0.000	0.202	0.139	117.223
KTINGSGSSHSKAQNQQT	3576	54.292	1.709	1.870	1.287	1.075	0.458	67.731
KTELRSGSPHSKAQNQQT	3461	54.289	0.000	0.007	0.040	0.790	0.239	57.814
KNINGSGSPHSKAQNHQT	3462	54.248	0.000	0.034	0.000	0.340	0.075	74.979
KTVNGGGSPHSKAQNHQT	3577	54.246	0.375	0.024	0.000	0.000	0.146	67.188
KTINGSGSPHSKARGEQT	3463	54.207	0.025	0.006	0.005	0.309	0.327	128.098
KTINGGGSPHSKAQYQHT	3578	54.188	0.000	0.000	0.000	0.000	0.223	82.256
KTEDLSGSPHSKAQNQQT	3464	54.156	0.000	0.000	0.000	1.193	0.132	70.198

KTINGSGSPHSKAPGQQT	3465	54.071	0.065	0.000	0.004	0.542	0.179	73.440
KTIPKNGSPHSKAQNQQT	3579	53.824	0.000	0.032	0.000	0.115	0.178	77.458
KTINGSGSPHSKAQSLQI	3466	53.778	0.000	0.186	0.000	0.022	0.047	51.543
KTINGSGSPHSKRLEQQT	3467	53.512	0.000	0.118	0.003	0.161	0.292	71.704
KTERGSGSPHSKAQNQQT	3468	53.475	0.000	0.030	0.000	1.416	0.175	85.368
KTVNGSGSPHSKAPNQQT	3469	53.444	0.833	2.206	0.006	0.156	0.178	58.080
KTSNGSGSPHSKAQNQST	3470	53.353	0.000	0.000	0.000	0.000	0.014	120.897
KTINGSGSPHSKAQKVIT	3471	53.273	0.000	0.000	0.000	0.357	0.402	95.147
KTEGISGSPHSKAQNQQT	3472	53.270	0.000	0.000	0.000	0.000	0.010	78.303
KTINGSGSPHSKAQNNDQ	3473	53.226	0.000	0.000	0.000	0.593	0.046	59.664
KTINGSGSPHSKAQSVHT	3474	53.226	0.000	0.004	0.000	0.446	0.217	76.110
KTINGSGSPHSKAQPLGT	3475	53.049	0.015	0.004	0.001	0.515	0.222	68.656
KTINKEGSPHSKAQNQQT	3580	53.006	0.000	0.029	0.000	0.177	0.111	64.520
KTCNASGSPHSKAQNQQT	3476	52.998	0.000	0.011	0.000	0.897	0.141	67.934
KAINGSGSPHSKAHNQET	3477	52.973	0.000	0.030	0.000	0.035	0.058	71.809
KTEGLSGSPHSKAQNQQT	3478	52.891	0.000	0.000	0.020	0.104	0.155	104.529
KTRDASGSPHSKAQNQQT	3479	52.861	0.000	0.000	0.010	1.062	0.402	52.089
KTSNGSGSPHSKAQNLQI	3480	52.843	0.000	0.000	1.605	0.178	0.214	74.823
KTGNGSGSPHSKAQIQQT	3481	52.809	0.000	0.000	0.000	0.000	0.012	98.291
KTVNGGGSPHSKAQNLQT	3581	52.788	0.000	0.031	0.000	0.000	0.165	83.215
KTDRSSGSPHSKAQNQQT	3482	52.737	0.000	0.000	0.000	0.995	0.085	123.421
KTINGSGSPHSKAQVRNT	3483	52.735	0.000	0.101	0.011	0.230	0.423	68.893
KTINGSGSPHSKAPSNQT	3484	52.680	1.494	4.762	0.003	0.330	0.208	87.951
KTINGSGSPHSKAQVGHT	3485	52.624	0.000	0.000	0.006	0.535	0.192	106.448
KNAIGSGSPHSKAQNQQT	3486	52.516	0.000	0.000	0.000	0.165	0.198	117.939
KAENGSGSPHSKAQNQQT	3487	52.487	0.000	0.157	0.029	0.000	0.242	120.256
KTINGSGSPHSKAQRDIT	3488	52.415	0.098	0.000	0.008	1.784	0.605	88.122
KTINGSGSPHSKAQMPNT	3489	52.408	0.084	0.036	0.025	0.057	0.359	66.040
KTVNGSGSPHSKSQNQQT	3490	52.395	0.033	0.077	0.013	0.105	0.175	58.000
KTIPAIGSPHSKAQNQQT	3582	52.346	0.000	0.009	0.000	0.034	0.134	51.949
KTINGSGSPHSKARGLQT	3491	52.275	0.000	0.000	0.036	1.235	1.425	169.881
KTELGSGSPHSKAQNQQT	3492	52.232	0.000	0.007	0.006	0.532	0.088	87.314
KAETGSGSPHSKAQNQQT	3493	52.219	0.000	0.047	0.581	0.009	0.188	132.940
KTINGSGSPHSKLQKQQT	3494	52.144	0.615	0.477	1.071	1.113	0.429	61.833
KTINGSGSPHSKAPSLQT	3495	52.137	0.041	1.614	0.002	0.902	0.222	70.363
KTINGSGSPHSKAQRDQT	3496	51.897	0.069	0.014	0.040	0.867	0.554	102.317
KTDVGSGSPHSKAQNQQT	3497	51.849	0.000	0.007	0.000	0.385	0.560	115.774
KTINGSGSPHSKNRDQQT	3498	51.830	0.000	0.008	0.000	0.480	0.138	100.300
KSINGSGSPHSKAPNLQT	3499	51.812	0.000	0.256	0.000	0.085	0.139	59.270
KTINGSGSPHSKAQAKGT	3500	51.727	0.048	0.016	0.000	0.271	0.525	104.917
KTVNGSGSPHSKAQDKQT	3501	51.580	0.428	0.000	0.069	0.041	0.063	69.225
KTINGGGSPHSKAQNPQA	3583	51.574	0.000	0.000	0.000	0.192	0.000	102.792
KTINGSGSPHSKAQSAHT	3502	51.569	0.068	0.070	0.000	0.589	0.249	79.498
KTINGNGSPHSKSQNQHT	3584	51.379	0.000	0.054	0.000	0.000	0.082	56.614
KTVPTSGSPHSKAQNQQT	3503	51.348	0.013	0.000	0.000	1.017	0.338	102.651
KTIDGSGSPHSKSQNHQT	3504	51.307	0.000	0.000	0.000	0.000	0.269	63.174
KTDVKSGSPHSKAQNQQT	3505	51.296	0.000	0.000	0.000	0.515	0.224	53.601
KAINRSGSPHSKAQDQQT	3506	51.262	0.000	0.000	0.000	0.000	0.036	54.631
KTINGSGSPHSKAQSTMT	3507	51.249	0.018	0.002	0.002	0.321	0.341	73.213
KTVNASGSPHSKAQNQLT	3508	51.249	0.000	0.000	0.000	0.000	0.268	99.559
KTINGSGSPHSKAQREMT	3509	51.076	0.000	24.900	143.49	1.564	0.476	70.961
KTVHGSGSPHSKAQSQQT	3510	51.057	0.000	0.000	0.000	0.143	0.146	54.185
KTINGGGSPHSKSQNRQT	3585	51.017	0.000	0.000	0.000	0.000	0.421	149.370
KTINGSGSPHSKAQYRAT	3511	51.008	0.000	0.158	0.000	0.690	0.120	50.650
KTINGGGSPHSKAQRQQT	3586	50.998	0.000	0.041	0.000	0.991	0.142	147.942
KTEPMSGSPHSKAQNQQT	3512	50.960	0.203	0.000	0.000	1.816	0.415	126.322
KTINGSGSPHSKNQWQQT	3513	50.800	0.000	0.044	0.047	0.111	0.324	65.506
KETAGSGSPHSKAQNQQT	3514	50.762	0.000	0.027	0.000	1.706	0.054	212.795

KTINGSGSPHSKAQRMNT	3515	50.686	0.000	108.747	0.019	0.943	0.264	97.975
KNNLGSGSPHSKAQNQQT	3516	50.670	0.000	0.019	0.000	0.406	0.121	102.408
KTINGSGSPHAKAQNHQT	3517	50.667	0.211	0.140	0.051	0.101	0.090	80.603
KTIIKNGSPHSKAQNQQT	3587	50.587	0.000	0.000	0.000	0.000	0.751	75.547
KTINGSGSPHSYHVNQQT	3588	50.486	0.000	0.056	0.059	0.528	0.275	179.489
KTINGSGSPHSKAGDSQT	3518	50.457	0.614	0.236	0.008	1.062	0.071	74.355
KTINGSGSPHSKLKSQQT	3519	50.368	0.000	0.296	0.000	1.796	1.096	95.240
KTINGSGSPHSKAQKIST	3520	50.285	0.000	0.000	0.088	0.108	0.302	51.115
KTEYNSGSPHSKAQNQQT	3521	50.256	0.000	0.000	0.000	0.000	0.009	62.679
KTINGSGSPHSKAPSMQT	3522	50.249	0.000	0.000	0.004	0.941	0.460	75.504
EAINGSGSPHSKAQNQQT	3523	50.243	0.629	0.094	0.000	0.057	1.519	117.305
KTINGSGSPHSKASPRQT	3524	50.227	0.088	0.005	0.068	1.761	0.530	67.241
KTINGSGSPHSKRMEQQT	3525	50.177	0.000	0.000	0.000	1.327	0.208	81.769
KTINGSGSPHSKAQYQNT	3526	50.099	0.000	0.008	0.000	0.017	0.119	71.846
KTERVSGSPHSKAQNQQT	3589	96.943	0.000	0.000	0.000	0.000	0.144	135.438
KAEIGHDSPHKSGQNQQT	1754	63.249	0.000	0.000	0.000	0.060	0.024	27.173

[0520] **Table 16** provides the peptide sequence of 341 matured capsid variants, and the fold enrichment of these matured capsid variants relative to the AAV9 control that demonstrated a 75-fold or greater increase in expression in the brain of NHPs relative to the AAV9 control and had a fold-change in expression that was less than 2 relative to the AAV9 control in the liver and the DRG.

Table 16. NGS fold-enrichment of TTM-001 and TTM-002 matured AAV capsid variants in the brain of NHPs

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9						
		Brain (NHP)	DRG (NHP)	Heart (NHP)	Muscle (NHP)	Liver RNA (NHP)	Liver DNA (NHP)	Brain (Mouse)
KTFNRSGSPHSKAQNQQI	3591	86.359	0.000	113.67	0.000	0.000	0.092	25.568
KTIIGSGSPHSKAQNRHT	3239	217.176	0.000	0.000	0.000	0.000	0.000	210.515
KTFPGSGSPHSKVQNQQT	3240	199.720	0.000	0.000	0.000	0.000	0.967	97.703
KTEKMSGSPHSKAQNQQT	3241	169.461	0.523	0.000	0.000	0.000	0.158	109.161
KAINGHDSPHKSGQIRQT	3606	108.510	0.000	23.908	0.000	0.132	0.261	8.862
KTINGHDSPHKIGQNQHA	3607	77.321	0.000	18.836	0.028	0.220	0.132	7.578
KEINGRGSPHSKAQNQQT	3527	134.390	0.239	0.000	0.000	0.000	0.232	52.311
KTVNRNGSPHSKAQNQQT	3528	133.016	0.000	0.416	0.000	0.000	0.000	85.361
KAINGYDSPHKSGQKQQT	3608	83.803	0.041	9.491	0.000	0.031	0.150	13.057
KTVNGSGSPHSKARDQQT	3242	124.789	0.123	0.039	0.312	0.569	0.454	132.137
KTESGHDSPHKSGQNQQT	3609	86.513	0.000	7.414	0.000	0.000	0.038	13.163
KTINGHDSPHKSGQSVQT	3610	75.748	0.010	6.808	0.000	0.165	0.058	9.321
KTFNGSGSPHSKAPNLQT	3243	121.436	0.000	0.167	0.000	0.000	0.015	168.920
KTEKTSGSPHSKAQNQQT	3244	120.337	0.000	0.355	0.000	0.000	0.119	101.467
TTINGHDSPHKSGQNQQT	3611	108.963	1.512	3.445	0.869	0.659	1.109	14.788
KTINGHESPHKSGRSQQT	3612	97.106	0.000	3.329	0.022	0.000	0.181	9.378
KTINGSGSPHSKAHVRQT	3245	119.798	0.000	0.000	0.262	0.694	1.039	165.590
KTVNGSGSPHSKAPNQHT	3246	117.207	0.000	0.109	0.000	0.000	0.074	51.008
KTEKISGSPHSKAQNQQT	3247	116.603	0.000	0.000	0.000	0.000	0.426	102.978
KTINGPGSPHSKAHNQQT	3529	115.742	0.146	0.000	0.235	0.000	0.513	52.508
KTINGHDSPHKSGQNKLE	3613	76.204	0.000	1.430	0.000	0.015	0.031	12.419
KTVNGSGSPHSKTQSQQT	3248	115.086	0.000	0.726	0.000	0.000	0.340	63.248
TTINGSGSPHSKAQNQQT	3249	114.856	1.340	14.856	0.827	1.281	0.957	72.058
KSINESGSPHSKAQNQQI	3250	113.833	0.000	0.000	0.000	0.000	0.000	67.649
KTINGHDSPHKTGQNQQK	3614	77.562	0.000	1.056	0.000	0.000	0.000	6.379
KTERTSGSPHSKAQNQQT	3251	112.957	0.000	0.009	0.000	1.128	0.207	117.374
KTINGSGSPHSKAQPAKT	3252	111.472	0.331	0.000	1.089	0.044	1.796	215.275

KTINGRGSPHKRGQNQQT	3837	120.889	0.100	0.814	0.434	0.458	0.614	13.988
KTINGSGSPHTKAQNPPT	3592	147.061	0.000	0.727	0.000	0.000	0.000	34.425
KTEKSSGSPHSKAQNQQT	3253	107.470	0.000	0.016	0.014	0.977	0.179	100.177
KAINGHDNPHKSGQNQQT	3615	88.906	0.297	0.721	0.482	0.222	0.130	9.702
KTSYGNGSPHSKAQNQQT	3530	105.937	0.000	0.000	0.000	0.000	0.114	105.894
KTINGQDSPHKSGQHQQA	3616	85.657	1.127	0.579	0.000	0.193	0.557	5.582
KTEKGSGSPHSKAQNQQT	3254	105.614	0.053	0.031	0.000	0.586	0.169	84.653
KTINGSGSPHSKSQTQQN	3255	104.474	0.000	0.131	0.000	0.084	0.038	54.021
KTERISGSPHSKAQNQQT	3256	103.692	0.000	0.000	0.000	0.062	0.370	89.637
KTERASGSPHSKAQNQQT	3257	103.669	0.000	0.000	0.000	0.127	0.070	115.550
KSINGHDSPHKSGQIQHT	3617	87.598	0.000	0.480	0.000	0.714	0.347	13.872
KELHGSGSPHSKAQNQQT	3258	102.680	0.000	0.000	0.000	1.634	0.592	96.554
KAINGSGSPHSKAQNLAT	3259	101.954	0.000	10.954	8.655	0.298	0.239	116.685
KTVNGSGSPHSKSQNQLT	3260	101.327	0.000	0.035	0.000	0.000	0.025	80.716
KAINGHDSPHKSGPRQQT	3618	145.142	0.000	0.408	0.000	0.000	0.000	8.259
KTVNGHDSPHKSGHTQQT	3619	82.246	0.000	0.378	1.142	0.000	0.123	6.160
KSINGHDSPHKSGQRQHT	3620	80.132	0.000	0.357	0.000	0.000	0.000	9.851
KTERNSGSPHSKAQNQQT	3261	99.892	0.000	0.000	0.000	0.000	0.107	87.392
KSLNGSGSPHTKAQNQQT	3593	81.515	0.197	0.333	0.000	0.000	0.085	45.140
KSVNGNGSPHSKAQNQQT	3531	99.385	0.000	1.329	0.000	0.359	0.079	51.016
KAINGHDSPHKSAQSQQT	3621	95.204	0.146	0.310	0.000	0.699	0.058	14.595
KSIYGHESPHKSGQNQQS	3622	90.947	0.817	0.310	0.000	0.000	0.243	8.064
KTFNGSGSPHSKAQGQQT	3262	99.253	0.000	0.208	0.000	0.128	0.099	81.459
KTVNGHDSPHKSLQNQQT	3623	112.925	0.000	0.301	0.059	0.000	0.322	16.726
KTINGSGSPHGWVQNQQT	3532	97.122	0.000	0.000	0.000	1.240	1.975	290.720
KTINGHGSPHSKAQNPQT	3838	83.478	0.000	0.288	0.219	0.000	0.260	11.001
KTSNGYDSPHKSGQKQQT	3624	77.001	0.032	0.286	0.000	0.000	0.016	8.813
KTVNGHDSPHKSGRNQET	3625	102.695	0.000	0.286	0.000	0.000	0.027	11.958
KTTNGHDSPHKSGQTQLT	3626	115.637	0.000	0.283	0.000	0.052	0.321	17.885
KAINGHDSPHKSEKNQQT	3627	77.103	0.000	0.274	0.000	0.000	0.000	26.868
KTERVSGSPHSKAQNQQT	3263	96.943	0.000	0.000	0.000	0.000	0.144	135.438
KTINGSGSPHSKALNRQS	3264	96.843	0.136	0.532	0.000	0.042	0.178	55.945
KTERLSGSPHSKAQNQQT	3265	95.857	0.000	0.004	0.005	0.126	0.260	102.372
KIINGRDSPHKSGQDQQT	3628	78.773	0.000	0.254	0.000	0.000	0.156	16.132
KTDNGSGSPHSKAHNQQT	3266	95.164	0.000	0.000	0.000	0.000	0.027	55.313
KTFHGSGSPHSKTQNQQT	3267	94.714	0.000	0.210	0.120	0.000	0.000	51.119
KTISGHDSPHKTGHNQQT	3629	92.490	0.000	0.233	0.057	0.730	0.000	8.823
KTVNAHDSPHKSGQNQLT	3630	79.137	0.000	0.233	0.178	0.753	0.045	29.254
KTINGGGSPHSKAQTQQI	3533	92.345	0.000	0.000	0.000	0.000	0.023	54.199
KSINGYDSPHKSGQTQQT	3631	79.227	1.817	0.226	0.000	0.000	1.148	4.497
KTINGHESPHKSGQTQQI	3632	86.089	0.000	0.222	0.000	0.000	0.024	3.989
KTINGHDSPHKSGQSKQA	3633	101.460	0.000	0.222	0.000	0.185	0.114	7.510
KTSNGSGSPHSKAQNPPT	3268	91.528	0.000	0.000	0.000	0.000	0.039	51.541
ETINGSGSPHSKAQNLQT	3269	90.969	0.221	1.023	0.197	0.179	0.813	107.216
KTVHGNGSPHSKAQNQQT	3534	90.073	0.000	0.000	0.000	0.000	0.304	97.003
NTINGSGSPHSKAQNQQT	3270	90.017	1.712	1.261	1.171	0.923	0.540	55.179
KTINGGGSPHSKAQNQQC	3535	89.301	0.219	0.000	0.000	0.287	0.319	53.840
KTENMSGSPHSKAQNQQT	3271	89.247	0.000	0.000	0.000	0.000	0.260	130.568
KTENVSGSPHSKAQNQQT	3272	88.506	0.000	0.000	0.000	0.964	0.112	108.591
KTSSGSGSPHSKAQYQQT	3273	87.304	0.000	0.000	0.000	0.000	0.299	58.143
KTIDGGGSPHSKAQNKQT	3536	85.019	0.000	0.000	0.000	0.000	0.477	55.517
KTEKVSGSPHSKAQNQQT	3274	84.558	0.000	0.022	0.000	0.873	0.424	112.185
KAINGSGSPHSKAQDQET	3275	84.080	0.000	0.000	0.000	0.194	0.027	87.637
KTCNKSGSPHSKAQNQQT	3276	83.992	0.000	0.000	0.165	0.283	0.000	119.496
KTINGGGSPHSKAQNQLI	3537	83.881	0.000	0.000	0.000	0.046	0.387	78.383
KNINGGGSPHSKAQNQQT	3538	83.083	0.000	0.042	0.000	0.000	0.000	75.913
KTEHLSGSPHSKAQNQQT	3277	83.080	0.000	0.000	0.012	0.021	0.189	69.494
KAIIGHESPHKSGQNQQT	3634	88.563	0.000	0.150	0.000	0.062	0.145	8.530

KTINGHDSPHKTGQNQPP	3635	77.357	0.000	0.149	0.000	0.000	0.096	8.865
KAINGHDSPHKSGQSPQT	3636	75.734	0.095	0.148	0.000	0.000	0.238	14.195
KAEMGSGSPHSKAQNQQT	3278	83.049	0.000	0.020	0.000	0.768	0.112	135.019
KATNGSGSPHSKAQNHQT	3279	82.627	0.000	0.176	0.000	0.155	0.057	66.207
KTIKGNDSPHKSVQNQQT	3637	85.986	0.000	0.135	0.000	0.263	0.000	8.603
KAIKGSGSPHSKAQDQQT	3280	82.258	0.000	0.000	0.000	0.108	0.000	85.178
KTINGGGSPHSKSQNQLT	3539	82.231	0.000	0.070	0.000	0.000	0.498	126.986
KTEFGHDSPHKSGQNQQT	3638	77.245	0.000	0.124	0.000	0.561	0.063	16.337
KTINGHDSPHKSAQNYQT	3639	130.375	0.000	0.124	0.000	0.097	0.123	19.443
KTFNGSASPHSKALNQQT	3839	84.258	0.000	0.122	0.000	0.104	0.037	31.855
KTINGCGSPHASGQNQQT	3840	132.540	0.000	0.121	0.042	0.000	0.059	1.857
KTINAHDSPHKIGQNHQT	3640	106.832	0.000	0.121	0.000	0.000	0.231	5.074
KTVNGNGSPHSKAQNKQT	3540	81.481	0.000	0.000	0.000	0.000	0.122	69.455
KTINGHESPHKSAQNRQT	3641	95.531	0.000	0.113	0.000	0.130	0.082	4.815
KTINGSGSPHSKGHWQQT	3281	81.434	0.000	0.000	0.000	0.000	1.011	65.252
KTTNGHDSPHKSGQNQQG	3642	85.113	0.000	0.107	0.000	0.000	0.017	10.555
KTIKGQDSPHKIGQNQQT	3643	110.357	0.000	0.103	0.058	0.166	0.135	11.829
KTDKTSGSPHSKAQNQQT	3282	81.430	0.000	0.000	0.000	1.362	0.291	169.515
KTVNGHDSPHKSGQNHLT	3644	81.516	0.000	0.100	0.017	0.000	0.028	16.096
KTFKGSGSPHSKAPNQQT	3283	80.890	0.000	0.000	0.000	0.000	0.017	71.144
KSINGHDSPHKSGQYQHT	3645	88.195	0.000	0.099	0.000	0.000	0.149	14.485
KTINGNDSPHKSVQNHQT	3646	120.002	0.000	0.099	0.788	0.000	0.000	7.920
KTVNGSGSPHSKAQNQLI	3284	80.509	0.000	0.000	0.000	0.000	0.166	71.156
KTINGSGSPHSKRPEQQT	3285	80.418	0.000	0.013	0.000	0.149	0.361	50.319
KTINGSGSPHSKAQRTMT	3286	80.388	0.000	0.022	0.170	1.812	1.025	100.248
KTITGHDSPHKSGQNQWT	3647	81.658	0.000	0.090	0.000	0.936	0.000	7.744
KTNNGHDSPHKSVQNQHT	3648	115.172	0.000	0.083	0.000	0.000	0.062	7.934
KTEKASGSPHSKAQNQQT	3287	80.285	0.000	0.041	0.000	0.000	0.261	90.390
KTIDGHDSPHKSGQNQHA	3649	91.058	0.000	0.082	0.000	0.000	0.000	10.781
KSDQGSGSPHSKAQNQQT	3288	80.076	0.000	0.000	0.000	0.993	0.124	151.911
KTVNGHDSPHKSGQTRQT	3650	133.276	0.251	0.080	0.093	0.034	0.129	7.174
KTVNGHDSPHKSGQNLHT	3651	88.080	0.000	0.080	0.000	0.000	0.039	11.363
KAISGHDSPHKSGLNQQT	3652	78.846	0.000	0.079	0.000	0.000	0.015	11.045
KTEITSGSPHSKAQNQQT	3289	79.620	0.000	0.163	0.000	0.332	0.074	76.686
KTDKSSGSPHSKAQNQQT	3290	79.470	0.055	0.012	0.000	1.437	0.367	141.351
KAINGHDSPHKSAQNQET	3653	90.402	0.000	0.073	0.000	0.746	0.000	10.674
KTITGHDSPHKSGQHLQT	3654	137.945	0.000	0.072	0.000	0.000	0.000	4.187
KTIDGSGSPHSKAQNQQH	3291	79.090	0.000	0.000	0.000	0.136	0.049	57.914
KTVNGNGSPHSKAQNQHT	3541	78.849	0.000	0.000	0.000	0.000	0.045	54.086
KNTNGSGSPHSKAQNQQT	3292	78.445	0.000	0.000	0.000	0.571	0.177	89.719
KTINGHDSPHKSRLNQPT	3655	92.883	0.000	0.070	0.050	0.904	1.075	5.598
KTETHSGSPHSKAQNQQT	3293	77.974	0.000	0.067	0.000	0.000	0.512	57.287
KTVDGHDSPHKSGQKQQT	3656	78.802	0.000	0.069	0.000	0.157	0.342	7.200
KTINGQDSPHKSGQNQDT	3657	82.075	0.000	0.067	0.000	0.225	0.144	9.626
KTINGGGSPHSKALNQQN	3542	77.822	0.000	0.131	0.000	0.000	0.274	69.884
KTIEGHDSPHKSGRNQQT	3658	75.838	0.000	0.065	0.017	0.000	0.079	7.818
KTTNGHDSPHKSGQNLLT	3659	77.738	0.130	0.064	0.185	0.424	0.326	15.192
KTINGHDSPHKSGQLVIT	3660	76.781	0.089	0.064	0.000	0.338	0.475	11.323
KTVNGHDSPHKSRQSQQT	3661	76.458	0.000	0.063	0.000	0.000	0.021	8.136
KTINGSGSPHSKALHQHT	3294	77.502	0.000	0.052	0.041	0.000	0.188	68.196
KTINGHDSPHKSGRTQET	3662	81.599	0.000	0.062	0.000	0.000	0.137	7.270
KTINGHDSPHKSVQTHQT	3663	77.309	0.237	0.062	0.000	0.000	0.116	7.519
KTINGTGSPHSKAQNHQI	3543	77.089	0.171	0.000	0.000	0.000	0.166	54.281
KTINGSGSPHSKAQHRIT	3295	76.849	0.105	0.499	0.170	1.424	0.214	127.000
KTINGSGSPHSKAQYIHT	3296	76.170	0.000	0.014	0.033	1.523	0.168	59.649
KTSNGHDSPHKSGQNQPA	3664	75.834	0.000	0.056	0.000	0.000	0.000	8.501
KTEGKHDSPHKSGQNQQT	3665	98.384	0.000	0.056	0.000	0.000	0.000	10.345
KTENISGSPHSKAQNQQT	3297	76.072	0.000	0.000	0.000	0.115	0.132	83.118

KVINGHDSPHKSGQTQQT	3666	91.665	0.000	0.055	1.526	0.311	0.000	7.391
KTIIGGGSPHSKAHNQQT	3544	75.872	0.000	0.050	0.000	0.000	0.235	65.492
KTINGPDSPHKIGQNQQS	3667	85.726	0.000	0.055	0.171	0.000	0.063	10.055
KTINGSGSPHSKAQKFET	3298	75.788	0.000	0.000	0.028	0.108	0.093	65.588
KTSNESGSPHSKAQNHQT	3299	75.720	0.000	0.000	0.000	0.169	0.217	70.590
KTINGSGSPHSKAQFPST	3300	75.677	0.000	0.004	0.000	0.849	0.127	119.712
KTERPSGSPHSKAQNQQT	3301	75.669	0.000	0.029	0.000	0.000	0.156	73.894
KAVNGHDSPHKSVQNQQT	3668	81.051	0.448	0.051	0.000	0.665	0.091	11.288
KTINGNGSPHSKAQNPLT	3545	75.269	0.000	0.000	0.000	0.366	0.000	53.583
KSIKGNGSPHSKAQNQQT	3546	75.196	0.000	0.000	0.000	0.000	0.000	90.251
KTINGHDSPHKSRQDQHT	3669	75.595	0.000	0.049	0.118	0.030	0.045	8.540
KAINGPDSPHKSGQKQQT	3670	78.213	0.464	0.047	0.000	0.323	0.162	10.395
KTINGHDSPHKSRQSQHT	3671	88.544	0.499	0.046	0.000	0.059	0.032	8.324
KTIYGHDSPHKSVQNQLT	3672	92.381	0.000	0.043	0.000	0.103	0.016	12.323
KTVNGHDSPHKSGQNLLT	3673	83.969	0.114	0.040	0.023	0.000	0.035	18.894
KTESAHDSPHKSGQNQQT	3674	80.810	0.000	0.039	0.000	0.000	0.000	13.338
KTENKSGSPHSKAQNQQT	3594	103.854	0.000	0.037	0.000	0.000	0.119	31.182
KTTNGQDSPHKSGQNQQS	3675	92.419	0.000	0.037	0.043	0.000	0.079	7.592
KTDKGSGSPHSKAQNQQT	3595	94.572	0.000	0.037	0.000	0.951	0.367	47.888
KTIDGHDSPHKSGRNQQI	3676	80.240	0.000	0.037	0.000	0.040	0.144	10.363
KTINGYDSPHKSGQYQHT	3677	81.534	0.000	0.036	0.000	0.000	0.000	10.524
KTDNGHDSPHKSRQNQQT	3678	105.312	0.000	0.033	0.000	0.000	0.018	7.931
KTINGHDSPHKSWVRQQT	3679	125.537	0.000	0.033	0.000	0.291	0.174	11.687
KTINGHESPHKSGQNQHS	3680	92.248	0.000	0.032	0.012	0.090	0.088	9.720
KTVNGHDSPHKIGHNQQT	3681	120.985	0.000	0.029	0.000	0.000	0.009	10.167
KTCNGHDSPHKSGRNQQT	3682	94.616	0.000	0.025	0.000	0.000	0.128	12.496
KTINGNGSPHSKAQNHQA	3841	88.274	0.000	0.024	0.000	0.000	0.041	36.754
KNVVGHDSPHKSGQNQQT	3683	75.330	0.000	0.024	0.000	0.063	0.049	8.077
KTELWHDSPHKSGQNQQT	3684	85.323	0.057	0.020	0.000	0.000	0.243	9.915
KTELRHDSPHKSGQNQQT	3685	98.098	0.000	0.019	0.000	0.000	0.007	6.588
KTINGHDSPHKSNAWQQT	3686	84.825	0.000	0.016	0.000	0.000	0.132	15.788
KTDAGHDSPHKSGQNQQT	3687	88.924	0.000	0.013	0.000	1.076	0.070	18.107
KTEVGHDSPHKSGQNQQT	3688	112.457	0.000	0.011	0.000	0.000	0.138	13.125
KTESRHDSPHKSGQNQQT	3689	81.766	0.000	0.011	0.000	0.052	0.036	6.975
KSELGHDSPHKSGQNQQT	3690	107.059	0.000	0.005	0.000	0.000	0.055	13.285
KTINGHDSPHKSGQSVPT	3691	77.840	0.000	0.003	0.000	0.136	0.061	6.768
KTINGHESPHKSGQNIQP	3692	253.840	0.000	0.000	0.000	0.000	0.000	14.042
KTEMKHDSPHKSGQNQQT	3693	240.075	0.000	0.000	0.000	0.000	0.000	3.183
KTINGHDSPHKSVQNHLN	3694	196.758	0.000	0.000	0.000	0.000	0.000	14.557
KTINGHDSPHKIGLDQQT	3695	165.627	0.000	0.000	0.000	1.942	0.000	5.469
KTSNASGSPHSKAQHQQT	3596	165.206	0.000	0.000	0.000	0.000	0.082	40.558
KTINGHDSPHKRGPDQQS	3696	160.084	0.000	0.000	0.000	0.000	0.000	2.923
KTINGMGSPHSKTQNQQT	3842	158.728	0.000	0.000	0.000	0.000	0.638	47.809
KTIKGHDSPHKSGESQQT	3697	142.264	0.000	0.000	0.000	0.000	0.218	4.176
KTEGWHDSPHKSGQNQQT	3698	142.064	0.000	0.000	0.000	0.000	0.264	11.785
KTINGHDSPHKHGQNHQT	3699	141.405	0.191	0.000	0.000	0.000	0.000	10.214
KTEQLHDSPHKSGQNQQT	3700	138.345	0.000	0.000	0.000	0.000	0.000	12.606
KTVNGTGSPHSKAQNQLT	3843	137.639	0.000	0.000	0.000	0.000	0.277	48.950
KTIIGHDSPHKSGQYQHT	3701	131.825	0.000	0.000	0.000	0.000	0.211	5.762
KTSNGHDSPHKSVQNKQT	3702	130.640	0.000	0.000	0.000	0.172	0.039	11.850
KIVNGQVSPHKSGQNQQT	3703	129.649	0.000	0.000	0.000	0.000	0.031	16.942
KTVNGHDSPHKSGQRQLT	3704	129.641	0.000	0.000	0.000	0.000	0.487	20.145
KTVNGHDSPHKIGQNQLT	3705	128.582	0.000	0.000	0.499	0.027	0.199	20.957
KTINGHDSPHKSGQIIVT	3706	125.245	0.000	0.000	0.151	0.000	0.379	6.808
KTEKIHDSPHKSGQNQQT	3707	125.178	0.000	0.000	0.000	0.000	0.000	17.604
KTENAHDSPHKSGQNQQT	3708	124.477	0.000	0.000	0.000	0.000	0.062	15.805
KIGNGHESPHKSGQNQQT	3709	123.324	0.000	0.000	0.000	0.000	0.000	11.198
KEVMGHDSPHKSGQNQQT	3710	121.107	0.000	0.000	0.000	0.000	0.000	17.191

KTEVKHDSPHKSGQNQQT	3711	119.733	0.000	0.000	0.000	0.000	0.000	5.550
KTINGYDSPHKSGQKQST	3712	119.615	0.000	0.000	0.000	0.000	0.000	7.970
KTIHGNGSPHSKAQNQET	3844	117.388	0.000	0.000	0.000	0.000	0.000	38.874
KYQVGHDSPHKSGQNQQT	3713	112.797	0.000	0.000	0.000	0.000	0.542	9.335
KTEAMHDSPHKSGQNQQT	3714	111.765	0.000	0.000	0.000	0.000	0.000	16.142
KTIKGDDSPHKSVQNQQT	3715	109.397	0.000	0.000	0.000	0.000	0.000	19.125
KTINGHDSPHKSVQSHQT	3716	109.375	0.107	0.000	0.319	0.000	0.547	12.617
KTINGHDSPHKSGQFVVT	3717	108.725	0.000	0.000	0.000	0.124	0.406	10.179
KTVNGHDSPHKSRQNLQT	3718	107.496	0.205	0.000	0.000	1.934	0.062	8.616
KATNGHNSPHKSGQNQET	3719	106.806	0.000	0.000	0.000	0.000	0.000	10.566
KAINGHDSPHKSAQNQQI	3720	106.539	0.000	0.000	0.000	0.000	0.113	21.786
KTEHGHDSPHKSGQNQQT	3721	106.486	0.000	0.000	0.000	0.000	0.006	14.956
KTVENHDSPHKSGQNQQT	3722	106.468	0.000	0.000	0.000	0.000	0.156	9.246
KTIYGHDSPHKSGQSQPT	3723	106.431	0.000	0.000	0.000	0.155	0.137	6.562
KTISGHESPHKSGQNEQT	3724	105.740	0.000	0.000	0.000	0.378	1.384	9.156
KAIIGHDSPHKSAQNQQT	3725	105.292	0.000	0.000	0.000	0.000	0.553	16.793
KAIDGHDSPHKSGQNQLT	3726	104.701	0.331	0.000	0.000	0.201	0.638	16.109
KTIMGHDSPHKSVQNQQT	3727	104.683	0.000	0.000	0.000	0.000	0.000	8.029
KEVGGHDSPHKSGQNQQT	3728	103.896	0.000	0.000	0.000	0.000	0.000	16.899
KTINGHDSPHKSAQNLLT	3729	103.332	0.000	0.000	0.256	0.194	0.000	15.722
KTEFTHDSPHKSGQNQQT	3730	102.052	0.062	0.000	0.000	0.439	0.047	12.527
KTINASGSPHSKAINQQT	3597	101.122	0.233	0.000	0.000	0.000	0.145	47.196
KAINGNGSPHKRGQNQQT	3845	100.925	0.000	0.000	0.000	0.000	0.159	10.011
KSEMGHDSPHKSGQNQQT	3731	100.539	0.000	0.000	0.000	0.000	0.000	18.356
KAQQGHDSPHKSGQNQQT	3732	100.395	0.000	0.000	0.000	0.000	0.057	3.954
KTEVMHDSPHKSGQNQQT	3733	99.473	0.000	0.000	0.000	0.000	0.000	12.400
KAINGHDSPHKSGQSLQT	3734	99.310	0.058	0.000	1.439	0.254	0.056	17.323
KTINGSGSPHSKAPNQQH	3598	99.300	0.252	0.000	0.000	0.000	0.038	39.297
KCGEGHDSPHKSGQNQQT	3735	99.298	0.000	0.000	0.000	0.000	0.000	13.147
KTVNGHDSPHKSAQNHQT	3736	99.257	0.000	0.000	0.078	0.000	0.027	17.639
KTVNGHDSPHKSGQTQLT	3737	98.524	0.000	0.000	0.313	0.183	0.172	14.883
KTNNGHDSPHKSGRNRQT	3738	98.307	0.000	0.000	0.124	0.000	0.037	5.840
KTCNEHDSPHKSGQNQQT	3739	97.092	0.000	0.000	0.000	0.000	0.000	8.364
KTINGHDSPHKYGQNEQT	3740	96.960	0.000	0.000	0.000	0.000	0.000	4.613
KASNRHDSPHKSGHNQQT	3741	96.283	0.000	0.000	0.000	0.000	0.675	8.045
KTINGNGSPHSKAPNLQT	3846	95.963	0.000	0.000	0.000	0.000	0.247	36.341
KTETKHDSPHKSGQNQQT	3742	95.121	0.000	0.000	0.000	0.000	0.083	4.831
KSINGHDSPHKSQQNQQT	3743	94.479	0.000	0.000	0.000	0.000	1.696	9.633
KTIGGHDSPHKSGQNQQI	3744	94.420	0.000	0.000	0.000	0.000	0.333	19.324
KTDPQHDSPHKSGQNQQT	3745	93.931	0.000	0.000	0.000	0.906	0.019	11.749
KTINRHDSPHKIVQNQQT	3746	93.409	0.000	0.000	0.000	0.000	0.000	3.064
KTEQYHDSPHKSGQNQQT	3747	93.065	0.000	0.000	0.000	0.000	0.047	15.190
KTINGHDSPHKSVQSKQT	3748	92.445	0.000	0.000	0.078	0.000	0.047	4.263
KELVGHDSPHKSGQNQQT	3749	92.262	0.000	0.000	0.000	0.594	0.000	14.890
KTENRHDSPHKSGQNQQT	3750	91.675	0.000	0.000	0.000	0.000	0.000	13.282
KELMGHDSPHKSGQNQQT	3751	91.191	0.000	0.000	0.000	0.000	0.000	14.032
KTINGNDSPHKIGHNQQT	3752	91.183	0.000	0.000	0.117	0.000	0.270	11.095
KTIKGGGSPHSKAQDQQT	3847	91.172	0.000	0.000	0.000	0.064	0.085	49.580
KTEGHHDSPHKSGQNQQT	3753	89.922	0.000	0.000	0.000	0.000	0.000	23.929
KTEGYHDSPHKSGQNQQT	3754	89.891	0.000	0.000	0.000	0.000	0.000	15.116
KTVNGHDSPHKSGQTQQI	3755	89.801	0.000	0.000	0.657	0.000	0.542	11.543
KTINGQDSPHKSGQNPLT	3756	89.726	0.000	0.000	0.000	0.363	0.000	15.561
KTVNASGSPHSKAQNHQT	3599	89.467	0.091	0.000	0.000	0.045	0.252	39.024
KTINGHDSPHKSGRDQKT	3757	88.871	0.000	0.000	0.000	0.350	0.181	12.117
KTINGHDSPHKSVHNQQN	3758	88.715	0.089	0.000	0.081	0.000	0.143	10.787
KTINGHDSPHKSGQWKRT	3759	88.633	0.000	0.000	0.000	0.202	0.094	5.186
KTIDGSGSPHSKAENRQT	3600	87.993	0.092	0.000	0.000	0.139	0.054	40.629
KNEIGHDSPHKSGQNQQT	3760	87.758	0.000	0.000	0.000	0.000	0.055	14.110

KAINGHDSPHKSGQSQQI	3761	87.585	0.000	0.000	5.310	0.000	0.000	12.864
KIINGHDSPHKSRQAQQT	3762	86.966	0.000	0.000	0.000	0.000	0.000	9.193
KTPNGHDSPHKSGQNQQI	3763	86.683	0.000	0.000	0.000	0.000	0.109	21.278
KITNGHDSPHKSGQTQQT	3764	86.443	0.000	0.000	0.000	0.192	0.190	17.479
KTINGHDSPHKSVQNHQI	3765	86.395	0.000	0.000	0.000	0.000	0.000	9.148
KTINGHDSPHKSKQNQQA	3766	86.265	0.000	0.000	0.000	0.123	0.041	5.768
KTINGHDSPHKSAQNQLN	3767	86.153	0.000	0.000	0.000	0.050	0.019	15.587
KTDITHDSPHKSGQNQQT	3768	85.876	0.000	0.000	0.000	0.000	0.013	9.076
KTVNGHDSPHKSGQTQPT	3769	85.680	0.000	0.000	1.301	1.064	0.000	8.067
KTEKFHDSPHKSGQNQQT	3770	85.358	0.000	0.000	0.000	0.000	0.026	7.229
KTDQGHDSPHKSGQNQQT	3771	85.267	0.000	0.000	0.000	0.000	0.000	16.042
KTINGHDSPHKLWINQQT	3772	85.132	0.000	0.000	1.154	0.000	0.017	12.704
KGINGPDSPHKSGQNQQT	3773	85.080	0.000	0.000	0.084	0.000	0.054	13.750
KSEIGHDSPHKSGQNQQT	3774	84.789	0.000	0.000	0.000	0.000	0.013	15.955
KTINGHDSPHKSVQKQLT	3775	84.351	0.000	0.000	0.000	0.038	0.103	11.890
KTINGHPSPHWKGQNQQT	3848	84.153	0.000	0.000	0.000	0.000	0.058	3.280
KTVNGHDSPHKSGRNQLA	3776	83.858	0.000	0.000	0.000	0.000	0.132	21.252
KTNNVHDSPHKSGQNQQS	3777	83.697	0.000	0.000	0.000	0.176	0.000	7.117
KTIKGSGSPHSKVQDQQT	3601	83.077	0.000	0.000	0.034	0.000	0.107	21.001
KSEKGHDSPHKSGQNQQT	3778	82.982	0.000	0.000	0.000	0.000	0.105	16.662
KWSAGHDSPHKSGQNQQT	3779	82.949	0.000	0.000	0.000	0.000	0.211	12.499
KELAGHDSPHKSGQNQQT	3780	82.876	0.000	0.000	0.000	0.000	0.093	18.063
KTINGHDSPHKMGRNQQS	3781	82.787	0.000	0.000	0.000	0.000	0.000	6.467
KTDQAHDSPHKSGQNQQT	3782	82.402	0.000	0.000	0.141	0.000	0.000	13.397
KTETQHDSPHKSGQNQQT	3783	82.316	0.000	0.000	0.000	0.000	0.198	10.823
KTEMTHDSPHKSGQNQQT	3784	82.221	0.000	0.000	0.000	0.000	0.000	8.431
KTINGHDSPHKSGISIQT	3785	82.019	0.000	0.000	0.000	0.191	0.044	7.310
KTDAVHDSPHKSGQNQQT	3786	81.968	0.000	0.000	0.000	0.297	0.107	13.596
KTSNGHDSPHKSVQNLQT	3787	81.921	0.000	0.000	0.072	0.000	0.330	11.544
KTEKYHDSPHKSGQNQQT	3788	81.637	0.000	0.000	0.000	0.000	0.013	7.580
KQTQGHDSPHKSGQNQQT	3789	81.581	0.000	0.000	0.000	0.000	0.133	15.225
KTINGHDSPHKMAHNQQT	3790	81.329	0.000	0.000	0.000	0.000	0.094	15.949
KAINGSGSPHSKAQTQQA	3602	81.207	0.000	0.000	0.000	0.000	0.016	40.435
KTINGHDSPHKHGQNQQN	3791	81.065	0.000	0.000	0.000	0.000	0.000	4.110
KGADGHDSPHKSGQNQQT	3792	80.981	0.000	0.000	0.000	0.000	0.074	11.423
KVGEGHDSPHKSGQNQQT	3793	80.775	0.084	0.000	0.000	0.000	0.019	16.378
KANEGHDSPHKSGQNQQT	3794	80.470	0.000	0.000	0.000	0.000	0.000	12.818
KTDTMHDSPHKSGQNQQT	3795	80.364	0.000	0.000	0.000	0.000	0.000	13.166
KTEAKSGSPHSKAQNQQT	3603	80.088	0.192	0.000	0.000	0.000	0.613	47.130
KTINGHDSPHKSVQSQQS	3796	80.000	0.000	0.000	0.000	1.055	0.082	17.620
KTIPGSGSPHSKAQNLQT	3604	79.973	0.871	0.000	0.000	0.000	0.000	32.693
KTCIAHDSPHKSGQNQQT	3797	79.857	0.000	0.000	0.066	0.000	0.093	1.930
KTINGHDSPHKSGQTVCT	3798	79.730	0.000	0.000	0.000	0.050	0.030	7.873
KELRGHDSPHKSGQNQQT	3799	79.596	0.000	0.000	0.000	0.000	0.006	22.001
KCQIGHDSPHKSGQNQQT	3800	79.359	0.000	0.000	0.000	0.000	0.000	2.614
KGVMGHDSPHKSGQNQQT	3801	79.170	0.000	0.000	0.000	0.138	0.086	17.287
KACDGHDSPHKSGQNQQT	3802	78.648	0.000	0.000	0.000	0.000	0.128	17.767
KTINGQDSPHKSGQYQQI	3803	78.585	0.000	0.000	0.000	0.286	0.672	5.664
KTINGHDSPHKSGQQIMT	3804	78.534	0.000	0.000	0.000	0.000	0.058	7.067
KTINGHDSPHKSRQNEQS	3805	78.534	0.000	0.000	0.000	0.112	0.188	13.388
KASNGHDSPHKSGLNHQT	3806	78.451	0.000	0.000	0.000	0.000	0.000	17.975
KTVNGHDSPHKSGQSQPT	3807	78.309	0.000	0.000	0.000	0.000	0.231	10.627
KNELGHDSPHKSGQNQQT	3808	78.135	0.000	0.000	0.000	0.000	0.182	17.457
KTETFHDSPHKSGQNQQT	3809	78.070	0.000	0.000	0.000	0.782	0.007	4.693
KAAEGHDSPHKSGQNQQT	3810	77.793	0.000	0.000	0.000	0.000	0.060	13.552
KGQNGHDSPHKSGQNQQT	3811	77.770	0.000	0.000	0.000	0.107	0.056	13.618
KNEFGHDSPHKSGQNQQT	3812	77.740	0.000	0.000	0.000	0.000	0.029	16.318
KTSIGYDSPHKSGQNQQT	3813	77.730	0.000	0.000	0.000	0.057	0.178	4.831

KTDNGHDSPHKSGQNLQT	3814	77.565	0.504	0.000	0.000	0.000	0.000	16.184
KTEGQHDSPHKSGQNQQT	3815	77.423	0.000	0.000	0.000	0.000	0.748	20.310
KTITGHDSPHKSRQDQQT	3816	77.127	0.000	0.000	0.000	0.000	0.000	6.250
KAEHGHDSPHKSGQNQQT	3817	77.026	0.000	0.000	0.000	0.000	0.017	20.937
KTINGDDSPHKSGQKQLT	3818	76.968	0.000	0.000	0.000	0.163	0.014	15.820
KCDQGHDSPHKSGQNQQT	3819	76.887	0.000	0.000	0.000	0.193	0.013	27.317
KEILGHDSPHKSGQNQQT	3820	76.770	0.000	0.000	0.000	0.804	0.009	10.771
KTIHGSGSPHSKAQNQAT	3605	76.765	0.000	0.000	0.000	0.000	0.215	43.969
KTERNHDSPHKSGQNQQT	3821	76.751	0.000	0.000	0.000	0.000	0.000	14.979
KAINGDDSPHKSGHNQQT	3822	76.578	0.000	0.000	0.000	0.032	0.059	17.755
KTSNGHNSPHKSGQNQET	3823	76.515	0.000	0.000	0.000	0.000	0.000	4.764
KTINGHDSPHKSGQMIHT	3824	76.364	0.000	0.000	0.000	0.000	0.000	9.486
KNAIGHDSPHKSGQNQQT	3825	76.289	0.000	0.000	0.000	0.009	0.072	15.178
KTDKFHDSPHKSGQNQQT	3826	76.204	0.000	0.000	0.000	0.000	0.000	7.096
KTEGFHDSPHKSGQNQQT	3827	76.191	0.000	0.000	0.000	0.000	0.080	13.163
KVINGHDSPHKSGRNHQS	3828	75.961	0.000	0.000	0.000	0.000	0.000	13.568
KTITGHDSPHKSVQNRQT	3829	75.940	0.000	0.000	0.000	0.621	0.000	4.310
KTPDMHDSPHKSGQNQQT	3830	75.871	0.659	0.000	0.000	0.000	0.048	11.277
KTINGHDSPHKSGQKMNT	3831	75.820	0.000	0.000	0.000	0.000	0.167	6.373
KTELQHDSPHKSGQNQQT	3832	75.814	0.000	0.000	0.000	0.105	0.000	11.798
KTIHGHDSPHKSGQSQQN	3833	75.777	0.000	0.000	0.059	0.000	0.166	7.426
KTEIGHDSPHKSGQNQQT	3834	75.525	0.000	0.000	0.016	0.012	0.000	9.593
KTINGHDSPHKSGQYQHA	3835	75.308	0.000	0.000	0.000	0.000	0.017	17.081
KTELYHDSPHKSGQNQQT	3836	75.235	0.000	0.000	0.000	0.000	0.042	10.354

[0521] **Table 17** provides the sequences of 216 matured capsid variants having a CV of less than 1 for the liver RNA samples isolated and a 10-fold or greater increase in expression relative to AAV9 in the liver of NHPs. These matured variants showed preferential transduction of the liver over other tissues as shown by a low value for fold-enrichment relative to AAV9 in the other tissues investigated including the brain, DRG, heart and muscle. As such, **Table 17** provides TTM-001 and TTM-002 matured AAV capsid variants with liver-specific tropism. Across the peptides within the matured capsid variants in **Table 17,** approximately 175 of them comprised the sequence GSGSPH (SEQ ID NO: 4695) and further comprised additional modifications in the C-terminal region of the sequence.

Table 17. NGS fold-enrichment of TTM-001 and TTM-002 matured AAV capsid variants in the liver of NHPs

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9					
		Liver RNA (NHP)	Liver DNA (NHP)	Brain (NHP)	DRG (NHP)	Heart (NHP)	Muscle (NHP)
KTQRKSGSPHSKAQNQQT	4011	119.659	1.439	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQARKT	4681	96.557	3.644	0.000	0.000	0.000	0.000
KYIVGSGSPHSKAQNQQT	4682	94.721	4.480	0.000	0.000	0.000	0.000
KTINGSGSPHSMYMNQQT	4683	81.106	5.840	0.000	0.000	0.000	0.000
KTINGSGSPHSKAFYRQT	4684	77.541	3.577	0.000	0.000	0.000	0.000
KTINGSGSPHSKLKRQQT	4685	76.103	6.884	0.000	0.000	0.000	0.000
KTINGSGSPHSKRHRQQT	4686	73.225	4.648	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQKCIT	4687	69.547	1.887	0.000	0.000	0.000	0.000
KTINGSGSPHSKWRLQQT	4688	68.083	2.037	0.000	0.000	0.000	0.000
KTINGSGSPHSRCRNQQT	4689	64.416	5.150	0.000	0.000	0.000	0.000
KTINGSGSPHSFTCNQQT	4690	63.936	2.155	0.000	0.000	0.000	271.289
KTINGSGSPHSKFFIQQT	4691	63.255	6.916	0.000	0.000	0.000	0.000
KTINGSYSPHCLAQNQQT	4012	62.942	0.309	0.000	0.000	0.000	0.000

KTINGSGSPHSKAQYSRT	4692	60.119	0.956	0.000	0.000	0.000	0.000
KTINGSGSPHIVWQNQQT	3849	58.021	6.056	0.000	0.000	0.000	0.000
KTINGSGSPHSKYFMQQT	3850	57.350	2.993	0.000	0.000	0.000	0.000
KTINGSGSPHSKARQRQT	3851	56.775	2.205	0.000	0.000	0.000	0.000
KTINGSGSPHSCHQNQQT	3852	56.242	8.562	0.000	0.000	0.000	0.000
KTINGSGSPHFPWQNQQT	3853	53.587	1.731	0.000	0.000	0.000	0.000
KTINGSGSPHSKIRRQQT	3854	53.528	1.388	0.000	0.000	0.000	0.000
KTINGSGSPHVYYQNQQT	3855	53.294	2.173	0.944	0.000	0.246	5.268
KTINGSGSPHSLYWNQQT	3856	53.262	0.000	0.000	0.000	0.000	0.000
KTINGSGSPHSKPKRQQT	3857	52.881	2.832	0.000	0.000	0.000	0.000
KPRWGSGSPHSKAQNQQT	3858	51.637	0.386	0.000	0.000	375.537	0.000
KTINGSGSPHSKAFSWQT	3859	51.304	1.805	0.000	0.000	0.000	0.000
KTINGSGSPHSRFWNQQT	3860	51.225	6.955	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQCLKT	3861	49.565	1.453	0.000	0.000	0.000	0.000
KTINGSGSPHSRMRNQQT	3862	48.902	2.816	0.000	0.000	0.000	0.000
KTINGSGSPHSVKKNQQT	3863	48.475	3.908	0.000	0.000	0.000	0.000
KTINGSGSPHSWAPNQQT	3864	47.897	1.789	0.000	0.000	0.000	0.000
KTINGSGSPHSLWKNQQT	3865	45.796	4.010	0.000	0.000	0.000	0.000
KTINGSGSPHSKARWQQT	3866	45.017	2.377	0.000	0.000	0.000	0.000
KTINGSGSPHSFRPNQQT	3867	44.801	9.191	0.000	0.000	0.000	0.000
KTINGSGSPHSKKVFQQT	3868	43.747	4.480	0.000	0.000	0.000	0.000
KTINGSLSPHFWAQNQQT	4013	43.190	2.041	0.000	0.000	0.000	0.000
KTINGSGSPHSYAFNQQT	3869	43.037	1.742	0.000	0.000	0.000	0.000
KTINHRISPHSKAQNQQT	4014	42.998	1.876	0.000	0.000	0.000	0.000
KTINGSGSPHSKACSRQT	3870	42.696	2.468	0.000	0.000	0.000	0.000
KTRRPSGSPHSKAQNQQT	4015	42.374	2.384	0.000	0.000	0.000	0.000
KYSAGSGSPHSKAQNQQT	3871	41.310	1.824	0.000	0.000	0.000	0.000
KTINGAYSPHRKAQNQQT	4016	40.969	1.283	0.000	0.000	0.000	0.000
KTINGSGSPHSKRLWQQT	3872	40.932	4.801	0.000	0.000	0.000	0.000
KTINGSGSPHSCSRNQQT	3873	40.372	4.293	0.000	0.000	0.000	0.000
KTINGSGSPHSRCPNQQT	3874	39.529	4.890	0.000	0.000	0.000	0.000
KTINGSGSPHSGACNQQT	3875	39.163	3.215	0.000	0.000	0.000	4733.916
KYYTGSGSPHSKAQNQQT	3876	38.777	1.199	0.000	0.000	0.000	0.000
KTINGSGSPHSKFRQQQT	3877	38.665	3.260	0.000	0.000	0.000	0.000
KTINGSGSPHSFPFNQQT	3878	38.584	4.693	0.000	0.000	0.000	0.000
KTINGSGSPHSFFGNQQT	3879	38.088	6.101	0.000	0.000	0.000	0.000
KTINGRRSPHGKAQNQQT	4017	37.728	3.259	0.000	0.000	0.000	0.000
KTINGSGSPHSMCQNQQT	3880	37.209	1.348	0.000	0.000	0.000	0.000
KTINGSGSPHSKLFWQQT	3881	37.022	4.178	0.000	0.000	0.000	0.000
KTINGSGSPHSKTRKQQT	3882	36.010	2.858	0.000	0.000	0.000	0.000
KTINGRTSPHRKAQNQQT	4018	35.792	5.682	0.000	0.000	0.000	0.000
KTINGSGSPHSGKRNQQT	3883	35.120	5.396	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQNFKR	3884	32.291	0.964	0.000	0.000	0.000	0.000
KTINGSGSPHFYRQNQQT	3885	31.724	9.342	0.000	0.000	0.000	0.000
KTINGSRSPHAWAQNQQT	4019	31.146	6.838	0.000	0.000	0.000	0.000
KTINGSGSPHCRVQNQQT	3886	31.043	1.203	0.000	0.000	0.000	0.000
KTINGSGSPHYGIQNQQT	3887	30.908	1.076	0.000	0.000	0.000	0.000
KTINKCLSPHSKAQNQQT	4020	30.667	5.097	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQRFKT	3888	30.363	0.139	0.000	0.000	0.000	0.000
KTINGSGSPHVNCQNQQT	3889	30.010	6.122	0.000	0.000	0.000	0.000
KTINGSGSPHSKPFRQQT	3890	29.842	8.700	0.000	0.000	0.000	0.000
KTINGSGSPHSLAWNQQT	3891	29.015	4.746	0.000	0.000	0.000	0.000
KTINGSGSPHSKRSYQQT	3892	28.973	2.116	0.000	0.000	0.000	0.000
KTINGSSSPHRCAQNQQT	4021	28.887	1.829	0.000	0.000	0.000	0.000
KTINGSGSPHWSYQNQQT	3893	28.607	3.751	0.000	0.000	0.000	0.000
KTINCRTSPHSKAQNQQT	4022	28.301	1.117	0.000	0.000	0.000	0.000
KTINGSGSPHRWLQNQQT	3894	28.147	6.882	0.000	0.000	0.000	0.000
KTIFDCGSPHSKAQNQQT	4023	27.844	1.602	0.000	0.000	0.000	0.000

KTINGSGSPHPSCQNQQT	3895	27.796	2.790	0.000	0.000	0.000	0.000
KTINGSGSPHSSWLNQQT	3896	27.318	3.271	0.000	0.000	0.000	0.000
KTINSPRSPHSKAQNQQT	4024	27.240	1.554	0.000	0.000	0.000	0.000
KPRFGSGSPHSKAQNQQT	3897	27.203	0.657	0.000	0.000	0.000	0.000
KWLTGSGSPHSKAQNQQT	3898	26.975	2.388	0.364	0.000	0.000	2578.486
KTINGSGSPHSKRRAQQT	3899	26.523	5.906	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQTCRT	3900	26.472	6.369	0.000	0.000	0.000	0.000
KTINGLDSPHRSRQNQQT	4025	26.403	0.321	0.000	0.000	0.000	0.000
KTINGSGSPHSKGCTQQT	3901	26.068	0.529	0.000	0.000	0.000	0.000
KTRTRSGSPHSKAQNQQT	4026	25.852	6.894	0.000	0.000	0.000	0.000
KTINGSGSPHVPWQNQQT	3902	25.294	3.435	0.000	0.000	0.000	0.000
KTINGSGSPHSKRYTQQT	3903	25.267	9.412	0.000	0.000	0.000	0.000
KTINGSISPHCPAQNQQT	4027	24.932	0.556	0.000	0.000	0.000	0.000
KTINGSGSPHSGCQNQQT	3904	24.818	1.981	0.000	0.000	0.000	0.000
KTINGSGSPHSFTPNQQT	3905	24.227	1.036	0.000	0.000	0.000	0.000
KTINGSGSPHSTTCNQQT	3906	23.771	3.315	0.000	0.000	0.000	0.000
KTINGSGSPHSKARMYQT	3907	23.424	0.313	0.000	0.000	0.000	0.000
KTINGLVSPHRKAQNQQT	4028	23.417	2.739	0.000	0.000	0.000	0.000
KTINGSGSPHPKRQNQQT	3908	23.055	2.355	0.000	0.000	0.000	0.000
KTINGSGSPHSKCFLQQT	3909	22.987	1.434	0.000	0.000	0.000	0.000
KTINGSGSPHWVPQNQQT	3910	22.907	3.219	0.000	0.000	0.000	0.000
KTINGSGSPHSFWSNQQT	3911	22.857	1.345	0.000	0.000	0.000	0.000
KRSYGSGSPHSKAQNQQT	3912	22.474	2.841	0.000	0.000	0.000	0.000
KYVFGSGSPHSKAQNQQT	3913	22.232	2.346	0.000	0.000	0.000	0.000
KTINGSGSPHSKFKNQQT	3914	21.951	1.074	0.000	0.000	0.000	0.000
KTINGSGSPHRIKQNQQT	3915	21.720	3.064	0.000	0.000	0.000	0.000
KTINGSGSPHSKAPRRQT	3916	21.645	3.940	0.000	0.000	0.000	0.000
KTINGSGSPHSFRYNQQT	3917	21.097	4.148	0.000	0.000	0.000	0.000
KTINGSGSPHSKMICQQT	3918	21.036	0.144	0.000	0.000	0.000	0.000
KTINGSGSPHLRWQNQQT	3919	21.014	9.649	0.000	0.000	0.000	0.000
KTINGSGSPHLPTQNQQT	3920	20.704	3.127	0.000	0.000	0.000	0.000
KTINGSGSPHSKWKSQQT	3921	20.390	1.239	0.000	0.000	4.904	0.163
KTINALRSPHSKAQNQQT	4029	20.053	1.655	0.000	0.000	0.000	0.000
KTINGSGSPHSYMRNQQT	3922	20.007	2.293	0.000	0.000	0.000	0.000
KTINGSGSPHSKAARRQT	3923	19.998	6.633	0.000	0.000	0.000	0.000
KTINGSGSPHLLCQNQQT	3924	19.796	3.484	0.673	0.000	0.000	1.309
KTINGSGSPHRCCQNQQT	3925	19.084	2.213	0.000	0.000	0.000	0.000
KTINGSGSPHLCVQNQQT	3926	19.030	1.428	0.000	0.000	0.000	0.000
KTINGSGSPHSKLTRQQT	3927	19.004	2.712	0.000	0.000	0.000	0.000
KTICGRGSPHSKAQNQQT	4030	18.923	2.171	0.000	0.000	0.000	0.000
KTTRKSGSPHSKAQNQQT	4031	18.849	2.617	0.000	0.000	0.000	0.000
KTINGSGSPHSKLCTQQT	3928	18.674	1.269	0.000	0.000	0.000	0.000
KKHLGSGSPHSKAQNQQT	3929	18.521	0.658	0.000	0.000	0.000	0.000
KTINGSGSPHSKIRGQQT	3930	18.150	1.584	0.000	0.000	0.000	0.000
KTMQRSGSPHSKAQNQQT	4032	18.020	3.159	0.000	0.000	0.000	0.000
KTINGSGSPHSYLVNQQT	3931	17.766	1.267	0.000	0.000	0.000	0.000
KTINGSGSPHQGCQNQQT	3932	17.676	1.037	0.000	0.000	0.000	0.000
KTINGSGSPHMAFQNQQT	3933	17.644	0.542	0.000	0.000	0.000	0.000
KTINGSGSPHSKACQFQT	3934	17.640	8.562	0.000	0.000	0.000	9.605
KTINGSGSPHSKWGLQQT	3935	17.543	2.639	0.000	0.000	0.000	0.000
KTINGSGSPHSKILRQQT	3936	17.419	2.546	0.000	0.000	0.000	0.000
KTINGSGSPHSFQINQQT	3937	17.418	0.269	0.308	0.000	1.568	0.000
KTINGSGSPHSKACISQT	3938	17.371	0.240	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQTHRT	3939	17.290	2.917	0.000	0.000	0.000	0.000
KTINGSGSPHSKALRCQT	3940	17.283	1.892	0.000	0.000	0.000	0.000
KTINGSGSPHSKAFYIQT	3941	17.172	0.239	0.000	0.000	0.000	0.000
KTINGSGSPHSKAHARQT	3942	17.075	1.800	0.000	0.000	0.000	0.000
KTINGSGSPHSLCLNQQT	3943	17.028	1.790	0.000	0.000	0.000	0.000

KTINGSGSPHSKAFVRQT	3944	16.935	1.985	0.000	0.000	0.000	0.000
KPPLGSGSPHSKAQNQQT	3945	16.897	0.805	0.000	0.000	0.000	0.000
KTINGSGSPHRPWQNQQT	3946	16.869	4.936	0.000	0.000	0.000	0.000
KPARGSGSPHSKAQNQQT	3947	16.793	1.391	0.000	0.000	0.000	0.000
KTINGSGSPHRPRQNQQT	3948	16.784	5.206	0.000	0.000	0.000	0.000
KTINGSGSPHSCPQNQQT	3949	16.701	1.776	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQFILT	3950	16.650	4.208	0.000	0.000	0.000	0.000
KTINGSSSPHWMAQNQQT	4033	16.635	2.390	0.000	0.000	0.000	0.000
KTRKRSGSPHSKAQNQQT	4034	16.603	2.075	0.000	0.000	0.000	0.000
KTINGSGSPHSVRYNQQT	3951	16.390	1.413	0.000	0.000	0.000	0.000
KSRRGSGSPHSKAQNQQT	3952	16.131	1.446	0.000	0.000	0.000	0.000
KTINGSGSPHSVRCNQQT	3953	15.860	3.912	0.000	0.000	0.000	0.000
KFFHGSGSPHSKAQNQQT	3954	15.412	0.897	0.000	0.000	0.000	0.000
KTINGSGSPHSKMPCQQT	3955	15.343	1.063	0.000	0.000	0.000	0.000
KTINGSGSPHSKKTSQQT	3956	15.244	1.344	0.000	0.000	0.000	0.000
KRYNGSGSPHSKAQNQQT	3957	15.160	0.806	0.000	0.000	0.000	0.000
KTINFTRSPHSKAQNQQT	4035	14.908	3.751	0.000	0.000	0.000	0.000
KTINGSGSPHSLPYNQQT	3958	14.792	2.048	0.000	0.000	0.000	0.000
KTINGSGSPHVYHQNQQT	3959	14.770	1.733	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQSRKT	3960	14.589	2.715	0.000	0.000	0.000	0.000
KTINGSGSPHSYTRNQQT	3961	14.535	1.986	0.000	0.000	0.000	0.000
KTINNLRSPHSKAQNQQT	4036	14.514	1.354	0.000	0.000	0.000	0.000
KTINGRPSPHGKAQNQQT	4037	14.442	0.705	0.000	0.000	0.000	0.000
KTINWSRSPHSKAQNQQT	4038	14.399	5.624	0.000	0.000	0.000	0.000
KTINGSGSPHLVYQNQQT	3962	14.196	1.045	0.000	0.000	0.000	0.000
KTINGTRSPHKKAQNQQT	4039	14.173	1.152	0.700	0.225	0.052	4.082
KTINGSGSPHSKALRWQT	3963	14.118	5.252	0.000	0.000	0.000	0.000
KTINGSGSPHYRYQNQQT	3964	14.107	1.027	0.000	0.000	0.000	0.000
KTINGSGSPHSWLKNQQT	3965	13.995	0.603	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQMQIT	3966	13.990	0.371	0.000	0.000	0.000	0.000
KTINGSVSPHCTAQNQQT	4040	13.502	2.955	0.000	0.000	0.000	0.000
KTINGSGSPHCPAQNQQT	3967	13.359	1.409	0.000	0.000	0.000	0.000
KTINGSGSPHSMCTNQQT	3968	13.114	0.392	0.000	0.000	0.000	0.000
KTINGSGSPHSPPDNQQT	3969	12.973	0.033	0.000	0.000	0.000	0.000
KTINGSGSPHSKRNYQQT	3970	12.781	5.528	0.000	0.000	0.000	0.000
KTTRCSGSPHSKAQNQQT	4041	12.639	8.168	0.000	0.000	0.000	0.000
KTKLCSGSPHSKAQNQQT	4042	12.570	2.139	0.000	0.000	0.000	0.000
KTINLGCSPHSKAQNQQT	4043	12.564	0.654	0.000	0.000	0.000	0.000
KTINGSGSPHRWTQNQQT	3971	12.490	0.844	0.000	0.000	0.000	0.000
KTISGHDSPHISGQYQQT	4044	12.395	0.420	0.000	0.000	0.074	1214.588
KTINGSGSPHSKACRLQT	3972	12.297	6.537	0.000	0.000	0.000	0.000
KTINGSGSPHPRKQNQQT	3973	12.249	3.248	0.000	0.000	0.000	0.000
KTINGSGSPHSKCSVQQT	3974	12.246	1.465	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQYVRT	3975	12.239	3.275	0.000	0.000	0.000	0.000
KTINGSGSPHSKARISQT	3976	12.142	1.565	0.000	0.000	0.000	0.000
KTINGRRSPHMKAQNQQT	4045	12.136	3.510	0.000	0.000	0.000	0.000
KTINGPWSPHRKAQNQQT	4046	12.103	0.434	0.000	0.000	0.000	0.000
KTINGSGSPHPFVQNQQT	3977	12.091	1.286	0.000	0.000	0.000	0.000
KTINGSGSPHSKLPKQQT	3978	11.856	0.274	0.000	0.000	0.000	0.000
KTINSCFSPHSKAQNQQT	4047	11.847	1.016	0.000	0.000	0.000	0.000
KTINGSGSPHSKSEQQQT	3979	11.785	1.769	0.000	0.000	0.000	0.000
KTINGSGSPHWVAQNQQT	3980	11.703	3.634	0.000	0.000	0.000	0.000
KTINGSGSPHSLYQNQQT	3981	11.590	1.503	0.000	0.000	0.000	0.000
KTINGSGSPHSKVRMQQT	3982	11.572	1.835	0.000	0.000	0.000	0.000
KTINYTRSPHSKAQNQQT	3983	11.514	0.431	0.000	0.000	0.000	0.000
KTIKRYGSPHSKAQNQQT	4048	11.461	2.022	0.000	0.000	0.000	0.000
KTINGSGSPHCALQNQQT	4693	11.404	3.867	0.000	0.000	0.000	0.000
KTINGSGSPHSSCTNQQT	3984	11.382	3.363	0.000	0.000	0.000	0.000

KTINGSGSPHSKNSRQQT	3985	11.280	1.093	0.000	0.000	0.000	0.000
KTINGSGSPHSKRKRQQT	3986	11.215	3.027	0.000	0.000	0.000	0.000
KTINGSGSPHLCTQNQQT	3987	11.176	2.489	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQSAKT	3988	11.162	4.200	0.000	0.000	0.000	0.000
KTINGSGSPHSTCLNQQT	3989	11.132	4.762	0.000	0.000	0.000	0.000
KTINGSGSPHSYARNQQT	3990	11.131	0.996	0.000	0.000	0.000	0.000
KTINGSGSPHSKQRPQQT	3991	11.130	2.347	0.000	0.000	0.000	0.000
KTINGSGSPHSKRVVQQT	3992	11.094	1.639	0.000	0.000	0.000	0.000
KRFSGSGSPHSKAQNQQT	3993	11.024	1.358	0.000	0.000	0.000	0.000
KTINGSGSPHKSGQNPQT	3994	11.014	11.790	0.000	0.000	0.000	0.000
KTINRYSSPHSKAQNQQT	4049	10.926	1.544	0.000	0.000	0.000	0.000
KTTGRSGSPHSKAQNQQT	4050	10.863	0.126	0.000	0.000	0.000	0.000
KTINGSGSPHSKALRHQT	3995	10.774	4.532	0.000	0.000	0.000	0.000
KTINGSGSPHSYYSNQQT	3996	10.680	2.856	0.000	0.000	0.000	0.000
KTINGSGSPHSLTCNQQT	3997	10.658	2.214	0.490	0.000	0.163	1.398
KTINGSGSPHSCQSNQQT	3998	10.631	1.468	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQSIKT	3999	10.544	1.355	0.000	0.000	0.000	0.000
KYSMGSGSPHSKAQNQQT	4000	10.478	1.587	0.000	0.000	0.000	0.000
KTINGSGSPHSKAKGWQT	4001	10.450	1.827	0.000	0.000	0.000	0.000
KTIVGSGSPHSKPQNQQT	4002	10.381	0.894	0.000	0.000	0.000	0.000
KTINGSGSPHFPFQNQQT	4003	10.322	3.715	0.000	0.000	0.000	0.000
KPFLGSGSPHSKAQNQQT	4004	10.318	1.328	0.000	0.000	0.000	0.000
KTINGSGSPHSKCTSQQT	4005	10.311	5.821	0.493	0.232	1.413	2.353
KTINRQFSPHSKAQNQQT	4051	10.275	4.480	0.000	0.000	0.000	0.000
KTINGSGSPHSVFENQQT	4006	10.218	0.224	0.000	0.000	0.000	0.000
KTINGSGSPHSKAKKVQT	4007	10.102	3.974	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQRCST	4008	10.084	0.762	0.000	0.000	0.000	0.000
KTINGSGSPHSKAQFCLT	4009	10.065	3.371	0.000	0.000	0.000	0.000
KTINGSGSPHGRYQNQQT	4010	10.028	0.778	0.000	0.000	0.000	0.000

[0522] **Table 18** provides the peptide sequences of 43 matured capsid variants having a raw virus count greater than 10, a CV of less than 1 for the heart samples isolated, and that also demonstrated a 4-fold or greater fold-increase in expression in the heart relative to the AAV9 control. A number of the matured variants shown in **Table 18** also demonstrated increased expression in other tissues isolated from the NHPs, including the brain, muscle, and/or liver, and are therefore pan-tropic.

Table 18. NGS fold-enrichment of TTM-001 and TTM-002 matured AAV capsid variants in the heart of NHPs

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9						
		Heart (NHP)	Brain (NHP)	DRG (NHP)	Muscle (NHP)	Liver RNA (NHP)	Liver DNA (NHP)	Brain (Mouse)
KTITGHDSPHSKAQNQQT	4052	34.375	230.437	4.338	1.378	19.165	5.672	0.000
KTINGSGSPHKSGQYQQT	4053	33.208	851.414	8.704	17.754	17.342	9.915	12.911
KTINGSGSPHKSGQDQQT	4054	31.166	218.057	34.358	33.372	27.081	8.836	24.849
KTINGSGSPHKSGQIQQT	4055	27.293	201.467	48.033	12.706	17.874	13.192	10.912
KTINGSGSPHKSGRNQQT	4056	27.283	313.826	8.723	36.593	15.252	12.352	21.595
KTINGSGSPHKSGKNQQT	4057	25.992	230.621	6.343	97.671	15.369	7.226	31.282
KTIYGHDSPHSKAQNQQT	4058	25.673	269.879	3.694	8.391	11.895	6.197	0.000
KTINGSGSPHKSGQNQQS	4059	24.783	244.030	16.675	26.058	18.059	9.809	29.751
KTINGSGSPHKSGQNLQT	4060	24.464	392.519	15.629	0.371	29.977	18.332	30.446
KAINGHDSPHSKAQNQQT	4061	22.460	640.466	7.358	9.986	9.358	8.490	0.000
KTVNGHDSPHSKAQNQQT	4062	21.066	614.034	3.392	30.908	21.560	11.933	121.235
KTIKGHDSPHSKAQNQQT	4063	20.803	213.564	24.646	12.361	15.379	6.551	13.319

KSINGHDSPHSKAQNQQT	4064	20.698	246.819	7.592	28.235	11.773	6.888	280.630
KTINGSGSPHKSGQTQQT	4065	19.925	466.459	55.454	15.485	15.473	6.446	15.179
KTINGSGSPHKSGHNQQT	4066	19.548	287.922	12.159	20.851	17.821	10.084	21.011
KTFNGHDSPHSKAQNQQT	4067	19.301	239.922	9.109	17.215	12.193	6.413	30.747
KTINGSGSPHKSGLNQQT	4068	19.136	319.093	3.083	4.096	14.009	7.446	9.340
KTINGHDSPHSKALNQQT	4069	18.542	605.641	13.375	1.902	12.621	7.054	51.283
KTINGSGSPHKSGQNQLT	4070	18.454	317.452	33.967	28.952	18.533	8.992	36.272
KTLNGHDSPHSKAQNQQT	4071	18.236	195.734	19.341	9.266	25.732	13.333	0.000
KTINGSGSPHKSGQNQHT	4072	14.269	313.837	7.125	39.273	29.714	7.797	25.119
KTIDGHDSPHSKAQNQQT	4073	13.836	242.100	1.731	12.555	17.223	7.439	0.000
KTNNGHDSPHSKAQNQQT	4074	12.872	134.488	0.504	3.877	17.044	5.982	22.358
KTINGSGSPHKSGQKQQT	4075	12.357	323.373	10.936	1172.3	12.604	7.970	48.699
KTINGSGSPHKSGQNRQT	4076	11.563	145.363	36.865	3.855	11.403	7.667	16.860
KTINGSGSPHKSGQNQQN	4077	11.507	156.385	582.38	8.559	9.273	7.668	18.138
KTINGSGSPHKSGQNQQA	4078	11.313	135.164	12.425	12.699	9.714	6.077	17.265
KTINGHDSPHSKAHNQQT	4079	10.024	236.106	19.495	5.258	2.406	3.316	45.691
KTINGHDSPHSKAQNQQT	4080	8.954	186.839	9.457	5.507	5.929	3.651	31.453
KTINGSGSPHKSGQNQQP	4081	8.744	261.947	43.435	10.217	6.468	4.265	19.828
KTINGHDSPHCKAQNQQT	4082	8.417	15.165	0.887	2.368	3.328	0.771	148.172
KTINGHDSPHSKAQNQQS	4083	5.678	603.027	7.280	0.670	4.301	4.307	65.271
KTINGSGSPHKSGQNQQT	4084	5.586	115.994	28.397	4.326	5.307	3.569	24.908
KTINGHDSPDKSGQNQQT	4085	5.569	30.854	4.934	1.112	0.671	0.781	14.499
KPINGHDSPHKSGQNHQS	4086	5.203	36.266	0.000	0.258	4.478	0.521	28.786
KTSNGSGSPHKSGQNQQT	4087	4.746	197.282	4.177	4.466	3.972	7.425	75.623
KTVNGSGSPHKSGQNQQT	4088	4.610	200.076	2.739	2.873	2.725	3.478	43.548
KTINGHDSTHKSGHNQQT	4089	4.369	27.630	2.883	1.302	0.421	0.176	12.973
KTINGHDSPHSKAQNQQN	4090	4.271	319.610	1.163	5.173	3.406	4.995	50.220
KTIYGSGSPHKSGQNQQT	4091	4.140	110.329	2.603	2.545	4.488	4.110	29.293
KTINGLDSQHKSGQNQQT	4092	4.055	12.958	3.240	3.205	0.645	0.296	5.608

[0523] **Table 19** provides the peptide sequences of 14 matured capsid variants having a raw virus count greater than 10, a CV of less than 1 for the muscle samples isolated (e.g., quadriceps), and that also demonstrated a 4-fold or greater fold-increase in expression in the muscle relative to the AAV9 control. A number of the matured variants shown in **Table 19** also demonstrated increased expression in other tissues isolated from the NHPs, including the brain, heart, and/or liver, and are therefore pan-tropic.

Table 19. NGS fold-enrichment of TTM-001 and TTM-002 matured AAV capsid variants in the muscle (e.g., quadriceps) of NHPs

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9						
		Muscle (NHP)	Brain (NHP)	DRG (NHP)	Heart (NHP)	Liver RNA (NHP)	Liver DNA (NHP)	Brain (Mouse)
KTINGSGSPHKSGRNQQT	4056	36.593	313.826	8.723	27.283	15.252	12.352	21.595
KTIIGHDSPHSKAQNQQT	4095	27.271	341.528	5.423	26.154	18.305	6.293	0.000
KTIYGHDSPHSKAQNQQT	4058	8.391	269.879	3.694	25.673	11.895	6.197	0.000
KTINGSGSPHKSGQNQQS	4059	26.058	244.030	16.675	24.783	18.059	9.809	29.751
KTVNGHDSPHSKAQNQQT	4062	30.908	614.034	3.392	21.066	21.560	11.933	121.235
KTIKGHDSPHSKAQNQQT	4063	12.361	213.564	24.646	20.803	15.379	6.551	13.319
KSINGHDSPHSKAQNQQT	4064	28.235	246.819	7.592	20.698	11.773	6.888	280.630
KTINGSGSPHKSGHNQQT	4066	20.851	287.922	12.159	19.548	17.821	10.084	21.011
KTFNGHDSPHSKAQNQQT	4067	17.215	239.922	9.109	19.301	12.193	6.413	30.747
KTSNGHDSPHSKAQNQQT	4096	18.580	507.189	7.777	17.770	21.537	8.789	70.219

KTINGHDSPHSKAQNQQT	4080	5.507	186.839	9.457	8.954	5.929	3.651	31.453
KTINGSGSPHKSGQNQQT	4084	4.326	115.994	28.397	5.586	5.307	3.569	24.908
KTINGHDSPHSKAQNQQN	4090	5.173	319.610	1.163	4.271	3.406	4.995	50.220
KTINGSGSPHSKAQNRRR	4097	4.237	8.348	0.291	0.636	1.597	5.396	158.853

[0524] These data demonstrate that following two maturation approaches, matured TTM-001 and TTM-002 capsid variants (AAV9 capsid variants) with loop IV modifications were generated with significantly enhanced CNS tropism over wild-type AAV9 controls in both NHPs and mice, while also exhibiting de-targeting in peripheral tissues (e.g., the liver and DRG). These resulting matured variants therefore demonstrated cross-species CNS tropism in both NHPs and mice. Matured TTM-001 and TTM-002 capsid variants with liver-specific tropism were also generated with at least 10 times the expression compared to wild-type AAV9 in the liver of NHPs. Several matured variants were also generated with increased expression in the heart and skeletal muscle (e.g., quadriceps) relative to wild-type AAV9 in NHPs.

Example 5. Evaluation of TTM-001 and TTM-002 AAV capsid variants in Diverse Primate Species

[0525] This Example evaluates the tropism and cross-species compatibility of the TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2) capsid variants in two diverse primate species, marmosets (*Callithrix jacchus*) and African green monkeys (*Chlorocebus sabaeus*), as compared to their tropism in cynomolgus macaques (*Macaca fascicularis*) provided in Example 1. The cross-species compatibility and tropism of an AAV9 capsid variant comprising the amino acid sequence of SPHKYG (SEQ ID NO: 966) was also investigated in this example. The amino acid and DNA sequences of TTM-001 and TTM-002 are provided, e.g., in Tables 4 and 5, respectively.

[0526] To investigate tropism in African green monkeys, AAV particles comprising the TTM-001 capsid variant, the TTM-002 capsid variant, an AAV9 capsid variant comprising SEQ ID NO: 966, or an AAV9 control under the control of a synapsin promoter, were intravenously injected into NHPs (n=2, 3-12 years of age) at a dose of 2E13 vg/kg. After 14-days in life, the brains and tissues (liver, DRG, quadriceps, and heart) of the NHPs were collected and RNA was extracted. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate the fold enrichment ratio relative to the AAV9 wild-type control.

[0527] To investigate tropism in marmoset monkeys, AAV particles comprising the TTM-001 capsid variant, the TTM-002 capsid variant, an AAV9 capsid variant comprising SEQ ID NO: 966, or an AAV9 control, were intravenously injected into NHPs (n=2, >10 months of age) at a dose of 2E13 vg/kg (8.75E12 vg/mL). After 28-days in life, the brains and tissues (liver quadriceps, and heart) of the NHPs were collected and RNA was extracted. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate the fold enrichment ratio relative to the AAV9 wild-type control.

[0528] As provided in **Table 20** (African green monkeys) and **Table 21** (marmosets), both the TTM-001 and TTM-002 capsid variants demonstrated increased CNS tropism in diverse primate species. The TTM-001 capsid variant demonstrated a 73.6-fold increase in expression relative to AAV9 in the brain of cynomolgus macaques (**Table 9**, Example 1), a 43.5-fold increase in expression relative to AAV9 in the brain of African green monkeys, and a 703.3-fold increase in expression relative to AAV9 in the brain of marmosets. The TTM-002 capsid variant demonstrated a 62.6-fold increase in expression relative to AAV9 in the brain of cynomolgus macaques (**Table 9**), a 13.8-fold increase in expression relative to AAV9 in the brain of African green monkeys, and a 366.6-fold increase in expression relative to AAV9 in the brain of marmosets. Both TTM-001 and TTM-002 led to a significant increase in expression relative to AAV9 in the heart of both African green monkeys and marmosets (**Table 20** and **Table 21**). The AAV9 capsid variant comprising SEQ ID NO: 966 also demonstrated in increase in expression relative to AAV9 in the brain and heart of both African green monkeys and marmosets. Furthermore, TTM-001, TTM-002, and the AAV9 capsid variant comprising SEQ ID NO: 966, also all led to increased expression in the brain of both BALB/c and C57Bl/6 mice (**Table 11**, Example 1), demonstrating an average fold change in expression relative to AAV9 across both species of mice of 63.1, 66.8, and 126.97, respectively.

Table 20. NGS-fold enrichment of TTM-001 (comprises SEQ ID NO: 941), TTM-002 (comprises SEQ ID NO: 2), and an AAV9 capsid variant comprising SEQ ID NO: 966 in African green monkeys

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9					
		Brain	DRG	Heart	Liver DNA	Liver RNA	Muscle
SPHSKA	941	43.525	1.010	184.789	0.242	1.547	1.715
HDSPHK	2	13.779	0.678	35.991	0.084	0.087	0.144
SPHKYG	966	9.805	0.071	44.865	0.085	0.136	0.234

Table 21. NGS-fold enrichment of TTM-001 (comprises SEQ ID NO: 941), TTM-002 (comprises SEQ ID NO: 943)), and an AAV9 capsid variant comprising SEQ ID NO: 966 in marmosets

Sequence	SEQ ID NO:	Fold Enrichment relative to AAV9				
		Brain	Heart	Liver DNA	Liver RNA	Muscle
SPHSKA	941	703.610	48.979	0.268	0.779	0.425
HDSPHK	2	366.625	18.572	0.075	0.276	0.229
SPHKYG	966	150.209	17.232	0.045	0.014	0.146

[0529] Taken together, these data demonstrate that the AAV9 capsid variants of TTM-001 and TTM-002 demonstrated increased CNS tropism relative to the AAV9 control in the CNS across three diverse primate species and two species of mice, providing evidence of strong cross-species capacity. The AAV9 capsid variant comprising the amino acid sequence of SEQ ID NO: 966 also demonstrated strong CNS expression relative to the AAV9 control in two species of NHPs and two species of mice, also showing strong cross-species capacity.

Example 6. Advanced maturation of TTM-002 capsid variant in mice

[0530] This Example describes additional maturation of the TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2) capsid variant in mice. In order to mature the TTM-002 capsid variant, sets of three contiguous amino acids were randomized across the mutagenesis region in TTM-002 sequence, which spanned from position 450 to position 466, numbered according to SEQ ID NO: 982. Unlike the maturation performed in in Example 3, where the SPH motif that was observed in the AAV capsid variants that demonstrated the greatest fold-enrichment in the NHP brain relative wild-type AAV9 was not disrupted, in the maturation approach used in this Example, the SPH motif was not held constant to further explore the role of this motif in the capsid variant. The matured TTM-002 capsid variants that resulted from the maturation approach were pooled together for subsequent testing and characterization in mice.

[0531] The library of matured AAV capsid variants generated from the TTM-002 matured AAV capsid variant were intravenously injected into the tail vein of three CD-1 Outbred mice (Charles River; 6-8 weeks of age) at a dose of 1.0×10^{12} VG/dose. After about 28 days in life, the brains of the mice were isolated, and RNA was extracted. Following RNA recovery and RT-PCR amplification, a systematic NGS enrichment analysis was performed to calculate the fold enrichment ratio relative to the corresponding TTM-002 non-matured control, and the peptides comprised within the variants were identified. Variants were filtered by those with a raw virus count in the sample above 10 and a coefficient of variance (CV) that was greater than 1 (identifies the peptides/variants reliably detected in the majority of the samples isolated from the three mice).

[0532] Following the advanced maturation screen and filtering of the variants, 1302 variants demonstrated an increase in expression relative to the non-matured TTM-002 capsid variant in the brain of the outbred mice. Of the 1302 variants with improved tropism relative to the non-matured TTM-002, 1283 comprised the SPH motif in the same position as the non-matured TTM-002 capsid variant (e.g., immediately subsequent to position 455, relative to a reference sequence numbered according to the amino acid sequence of SEQ ID NO: 138 or 982). Mutations in the region of the SPH motif present in the non-matured TTM-002 capsid variant only consistently appear in those variants with a fold change of 0.2 or 0.1 or lower relative to the non-matured TTM-002 control in the brain of the mice. This indicates that the SPH motif may be important to the increased brain tropism that observed for the TTM-002 capsid variant. In instances when the SPH motif was disrupted, the fold change of the matured variants of TTM-002 decreased considerably in relation to the non-matured TTM-002 variant which comprised the SPH motif.

Example 7. Tropism of TTM-002 AAV capsid variant

[0533] This Example further investigates the tropism and CNS cells transduced by the TTM-002 capsid variant (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2), as

outlined in Table 3 above. The amino acid and DNA sequences of TTM-002 are provided, e.g., in Tables 4 and 5, respectively.

[0534] AAV particles were generated with the TTM-002 capsid variant encapsulating a GFP transgene (AAV_TTM-002.GFP) or a payload driven by a heterologous CBA constitutive promoter (AAV_TTM-002.Payload).

[0535] Two tandem single cell RNA sequencing runs (scRNA-Seq) of mouse cells derived from the midbrain area were performed. In the first run, cells were pooled from two mice at day 28 post treatment with AAV_TTM-002.Payload particles. In the second run we treated with AAV_TTM-002.GFP particles, in the same manner but without xenografts. Orthotopic xenografts of MDA-MB-361-Luc#1 high passage cells grown as tumorspheres (in tumorsphere media; Sigma # C-28070) were injected (250,000 cells/2µL/mouse) intracranially into 2-month old female SCID CB17 (Mutation: Icr-Prkdcscid/IcrIcoCrl) congenic immunodeficient mice (Charles River Laboratories). The injections were 2.5mm (lateral), -1mm (posterior) with respect to bregma, lowered -3mm ventral and raised +.5 mm dorsal to a final -2.5mm ventral position. Two days later, dilutions of the AAV_TTM-002.Payload particles (run 1), or in the case without xenografts, dilutions of AAV_TTM-002.GFP particles (run 2) were prepared. IV injections of 100µL (2.5e11 VG/animal) of the AAV_TTM-002.payload particles or AAV_TTM-002.GFP particles were administered through the tail veins of mice (n=5 mice per groups). At 7 days post-injection, mice from run 1 were imaged in an AmiHTX (Spectral Imager) for bioluminescence of the human tumor cells due to expression of luciferase in response to intraperitoneal luciferin injections.

[0536] At 28 days post-injection with the AAV_TTM-002.payload particles or AAV_TTM-002.GFP particles, two mice from each run were necropsied, brain samples were isolated, and the midbrain was dissected and isolated. The midbrain samples were then exposed to a cold protease inhibitor (Creative Biomart #NATE-0633) and were dissociated at 6 degrees centigrade. For the samples collected from the mice of run 1 (AAV_TTM-002.Payload particles), myelin depletion was performed (Miltenyi, #130-096-731), cells were filtered through a 40µM mesh to filter out neurons) and loaded on a 10X chromium G chip. scRNA-Seq was performed (10X Genomics) and samples were sequenced on a NextGen500 Sequencing machine (Illumina). For the samples collected from run 2 (AAV_TTM-002.GFP particles and no xenografts), the cells were not myelin depleted or filtered through 40µM mesh to include neurons. The cells isolated after run 2 were FACS sorted for GFP+/7AAD- (live GFP+ cells). The resultant cells were loaded on a 10X chromium G chip and the scRNA-Seq was run and processed (10X Genomics).

[0537] For run 1, the scRNA-Seq data was filtered to include cells with only greater than 1000 genes per cell and less than 5000, and less than 20 percent mitochondrial gene expression. For run 2, the scRNA-Seq data was filtered to include cells with only greater than 200 genes per cell and less than 5000, and less than 20 percent mitochondrial gene expression. The data were normalized, scaled,

and integrated into one combined dataset. Clusters were generated with a resolution of 0.3 and each cluster identity was determined using a panel of cell type specific genes (e.g., as described in Brown *et al.*, 2021. "Deep Parallel Characterization of AAV Tropism and AAV-Mediated Transcriptional Changes via Single-Cell RNA Sequencing". *Front. Immunol.* 12:730825; the contents of which are hereby incorporated by reference in its entirety). The percentage of GFP sorted cells per cluster was calculated as was the percentage of payload expressing genes per cluster as parallel measures of TTM-002 transduction.

[0538] For payload expressing cells, endothelial cells had the highest proportion of payload positive cells, followed by astrocytes (**Table 22**). For GFP+ sorted cells, endothelial cells had the highest proportion of GFP positive cells, and astrocytes were the third highest cell type when sorting by proportion of cells expressing GFP (**Table 22**). These data indicate TTM-002 transduction exhibits an endothelial and astrocytic tropism. Furthermore, the astrocytic cluster had the second highest level of expression of Olig2 (oligodendrocytes demonstrated the greatest Olig2 expression). IHC staining was performed on brain samples isolated from AAV_TTM-002.GFP infected mice and demonstrated that GFP co-localized with some but not all Olig2+ cells. No co-staining was observed with mylein basic protein (MBP), a marker of oligodendrocytes. Co-staining with GFP was also not observed in NeuN positive cells (neurons), GFAP positive cells (astrocytes), and Iba1 positive cells (microglia). GFP staining was observed throughout the sagittal section of the mouse brain, which was demonstrative of increased staining in the midbrain. The GFP expressing cells observed did not have a bipolar morphology like oligodendrocyte progenitor (OPC) cells and therefore, together with the scRNA-Seq data, these results indicated that at day 28 post AAV treatment, Olig2+ astrocytes in the midbrain are being transduced by AAV particles comprising a TTM-002 capsid, in a cell type specific tropism.

Table 22. Quantification of payload positive cells and GFP positive cells

Cluster Identity	% Payload Cells/Cluster	Cluster Identity	% GFP Cells/Cluster
Endothelial-2	6.58	Endothelial-2	6.58
Astrocyte	4.50	Endothelial-1	3.45
Pericytes	4.23	Vascular and leptomeningeal Cells (VLM)	2.38
Mature Oligos	3.85	Astrocyte	2.37
Endothelial-1	3.09	Vascular smooth muscle cells (VSC)	1.03
Committed Oligos	1.90	Pericytes	0.77
Vascular smooth muscle cells (VSC)	1.72	Microglia	0.00
Microglia	0.40	Committed Oligos	0.00
Macrophages	0.00	Macrophages	0.00
Vascular and leptomeningeal cells (VLM)	0.00	Oligodendrocytes	0.00
Oligodendrocytes	0.00	Committed Oligos-2	0.00
Committed Oligos-2	0.00	Mature Oligos	0.00

Example 8. Identification of a receptor for TTM-001 and TTM-002 capsid variants

[0539] This example investigates the tropism and receptor of the TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2) capsid variants for crossing the blood brain barrier. Without wishing to be bound by theory, it is believed that identification of a receptor of these AAV capsid variants provides a better understand of translatability of these variants to different species, as well as the mechanism used for crossing the blood brain barrier that results in an increase in CNS transduction relative to AAV9.

A. Binding of TTM-001 and TTM-002 capsid variants to N-linked galactose

[0540] Primary glycan receptors have been identified for various AAV serotypes, including AAV9 which binds N-linked galactose. In order to investigate the ability of TTM-001 and TTM-002 AAV9 variants to retain this natural glycan binding, HeLa cells were treated with increasing concentrations of Neuraminidase (0, 5, 50, 500, and 100 mU/mL), which cleaves N-sialic acid and exposes N-galactose. The treated cells were then transduced with AAV particles comprising the TTM-001 capsid variant (AAV_TTM-001), the TTM-002 capsid variant (AAV_TTM-002), or an AAV9 control (AAV_AAV9cntl), transduction was measured by quantification of Luc2 activity (RLU), and data was normalized relative to the no neuraminidase control. As shown in **Table 25**, enzymatic removal of N-sialic acid and exposure of N-galactose on HeLa cells, resulted in a dose dependent increase, more specifically a 9- to 14-fold increase, in transduction by AAV particles comprising the TTM-001 capsid variant and AAV particles comprising the TTM-002 capsid variant. This was analogous to what was observed with the AAV9 control (**Table 25**). These data demonstrate that the TTM-001 and TTM-002 AAV9 capsid variants retained the natural binding affinity to terminal N-linked galactose observed with AAV9 wild-type.

Table 25. Quantification of HeLa cell transduction post-neuraminidase treatment and transduction with AAV_TTM-001 particles, AAV_TTM-002 particles, or AAV_AAV9cntrl particles. Data measured as fold change in Luc2 activity (RLU) relative to the no neuraminidase control

Capsid	Concentration of Neuraminidase			
	5 mU/mL	50 mU/mL	500 mU/mL	1000 mU/mL
AAV9 (control)	2.7	6.2	11.1	10.4
TTM-001	5.2	6.7	8.2	9.1
TTM-002	9.6	12.5	13.9	13.7

B. Receptor Identification

[0541] A cell binding array assay was then used to identify a receptor for the TTM-001 and TTM-002 capsid variants. Briefly, a library of over 5,500 cDNAs was overexpressed in human cells. Cells were contacted with a test ligand, e.g., AAV viral particles comprising a TTM-001 capsid variant or an AAV9 control capsid, which was applied to the array. Binding of the TTM-001 capsid variant or

the AAV9 control capsid to the cells was detected using an anti-AAV9 antibody followed by a labeled anti-IgG detection antibody. A comparison of the proteins contacted using AAV particles comprising a wild type AAV9 control capsid and AAV particles comprising a TTM-001 capsid variant revealed a unique interaction with the TTM-001 capsid variant but not the AAV9 wild-type control capsid. This interacting protein was identified as the GPI-anchored protein, alkaline phosphatase issue-nonspecific isozyme (NM_000478.4, which is incorporated by reference herein) (ALPL). ALPL is part of a family of membrane-bound glycoproteins that hydrolyze monophosphate esters at a high pH (see, e.g., Weiss et al., Isolation and characterization of a cDNA encoding a human liver/bone/kidney-type alkaline phosphatase. Proc. Natl. Acad. Sci., 83: 7182-7186 (1986), the contents of which are hereby incorporated by reference in their entirety).

[0542] ALPL is highly conserved across humans, mice, and cynomolgus macaques (*Macaca fascicularis*) when compared by sequence alignment (**Table 26**). Additionally, in humans ALPL is expressed on endothelial cells and neurons, and at a low level on astrocytes. The highest level of ALPL expression in human is on endothelial cells. In mice, ALPL is more highly expressed on astrocytes, oligodendrocyte progenitor cells (OPCs), and to a lesser extent on endothelial cells.

Table 26. Identity and similarity of ALPL receptor between different, species

	Percent Identity	Percent Similarity
Human to cynomolgus macaque	97.33%	98.5%
Human to mouse	89.89%	94.47%
Cynomolgus macaques to mouse	90.46%	95.23%

[0543] Furthermore, as shown in **Example 7** and **Table 22**, when mice were treated intravenously with AAV particles comprising the TTM-002 capsid variant expressing a payload, payload expression as measured by RNA-seq was the highest in a subset of endothelial cells (**FIG. 1A**). This same subset of endothelial cells also showed high expression of ALPL by RNA-seq (**FIG. 1B**). These data indicated a correlation between expression of ALPL and the TTM-002 tropism in mice.

[0544] Taken together, these data indicate that the TTM-001 and TTM-002 capsid variants are capable of binding ALPL, which could serve as a receptor for crossing the blood brain barrier and CNS transduction.

C. Characterization of interaction of the TTM-001 and/or TTM-002 with ALPL

[0545] In order to further characterize the interaction between the TTM-001 and TTM-002 capsid variants and the ALPL protein, it was investigated whether increased expression of the ALPL protein resulted in increased transduction of AAV particles comprising the TTM-001 or TTM-002 capsid variant. Briefly, a transduction assay was performed in that HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing ALPL, an AAVR positive control, or a pCMV6 negative control (250 ng or 500 ng of plasmid). AAVR is a universal AAV entry factor involved in AAV transduction. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL protein or other controls were transduced with an AAV particle comprising the TTM-001

capsid variant, the TTM-002 capsid variant, another AAV capsid variant (TTD-001), or an AAV9 control capsid protein, expressing a GFP payload. At 24-hours post-transduction, GFP expression and luciferase activity were measured to quantify and observe AAV cellular transduction. By immunofluorescence microscopy, expression of the ALPL protein resulted in a significant increase in the transduction of AAV particles comprising the TTM-002 capsid variant compared to particles comprising the AAV9 wild-type control capsid. Additionally, the increase in transduction of the AAV particles comprising the TTM-002 capsid variant was specific to ALPL expression, as expression of the AAVR control did not result in the same increase in transduction of AAV particles comprising the TTM-002 capsid variant. As summarized in **Table 27**, expression of ALPL led to a 35 and 45-fold increase in transduction the TTM-001 and TTM-002 AAV9 capsid variants, respectively, when measured by a luciferase assay. Transduction of AAV9 wild-type control as well as the AAV9 capsid variant TTD-001 was not affected by expression of ALPL, indicating the specific role of ALPL in transduction of TTM-001 and TTM-002. TTD-001 is an AAV9 capsid variant comprising a loop VIII modification, and the sequence and capsid characterized can be found in WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety). **Table 28** provides the results of a second experiment performed as described above, where HEK 293T cells expressing the ALPL protein or the other controls were transduced with AAV particles comprising the TTM-002 capsid variant or one of three AAV9 capsid variants also comprising a modification in loop IV: TTM-006 (SEQ ID NO: 39), TTM-018 (SEQ ID NO: 51), and TTM-019 (SEQ ID NO: 52). The TTM-002, TTM-006, TTM-018, and TTM-019 capsid variants all comprised the SPH motif immediately subsequent to position 455, numbered relative to SEQ ID NO: 138 and a positive residue in one of the next three residues subsequent to the SPH motif. The TTM-002, TTM-006, TTM-018, and TTM-019 capsid variants all led to an increase in transduction in cells expressing ALPL, which was not observed with the AAV9 control (**Table 28**).

Table 27. Transduction of TTM-001 and TTM-002 capsid variants as measured by luciferase assay relative to the AAV9 control and AAV variant TTD-001 with loop VIII modification (data shown as fold change relative to the pCMV6 transfected negative control cells)

	Luciferase Activity (fold change relative to pCMV6 transfected cells)		
Capsid	**pCMV6 transfected cells**	**AAVR transfected cells**	**ALPL transfected cells**
AAV9	1	4.3	1.5
TTD-001	1	22	2
TTM-001	1	2.4	35
TTM-002	1	4.1	44

Table 28. Transduction of TTM-002, TTM-006, TTM-018, and TTM-019 capsid variants as measured by luciferase assay relative to the AAV9 control (data shown as fold change relative to the pCMV6 transfected negative control cells)

	Luciferase Activity (fold change relative to pCMV6 transfected cells)		
Capsid	**pCMV6 transfected cells**	**AAVR transfected cells**	**ALPL transfected cells**
AAV9	1	14	0.9

TTM-002	1	24	166
TTM-006	1	25	91
TTM-018	1	25	90
TTM-019	1	28	88

[0546] Binding and internalization of AAV capsid variants comprising the TTM-001 capsid variant, the TTM-002 capsid variant, or the AAV9 control capsid was also investigated in cells engineered to express ALPL. HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing ALPL, an AAVR positive control, or a pCMV6 negative control. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL receptor were incubated with an AAV particle comprising the TTM-001 capsid variant, the TTM-002 capsid variant, or an AAV9 control capsid protein, expressing a GFP payload. At 2 or 3-hours post-incubation, cells were washed to remove unbound AAV particles and DNA was extracted to quantify viral genomes. As shown in **Table 29**, expression of ALPL led to a 3-fold and 6-fold increase in binding/internalization by TTM-001 and TTM-002, respectively. This effect was specific to TTM-001 and TTM-002, as the binding/internalization of the wild-type AAV9 control was unaffected by ALPL expression.

Table 29. Relative viral gene expression ($2^{-\Delta\Delta CT}$) of cells transfected with a plasmid expressing pCMV6 control, an AAVR control, or ALPL and subsequently transduced with AAV particles comprising the TTM-001 capsid variant, TTM-002 capsid variant, or AAV9 control

	Fold change in viral gene expression ($2^{-\Delta\Delta CT}$) relative to pCMV6 transfected cells		
Capsid	**pCMV6 transfected cells**	**AAVR transfected cells**	**ALPL transfected cells**
AAV9	1	7.04	1.06
TTM-001	1	6.64	3.16
TTM-002	1	3.2	5.92

[0547] Three isoforms of ALPL exist, isoform 1 (Alkaline phosphatase, placental-like 2 (ALPPL2), NM_031313, which is incorporated by reference herein), isoform 2 (alkaline phosphatase, placental (ALPP), NM_001632, which is incorporated by reference herein), and isoform 3 (alkaline phosphatase, intestinal (ALPLI), NM_001631, which is incorporated by reference herein), that can also be expressed on cell surfaces via a GPI-anchor. Isoform 1 is 56.25% identical and 72.54% similar (gaps: 4.17%) to ALPL, isoform 2 is 54.96% identical and 71.37% similar (gaps: 2.29%) to ALPL, and isoform 3 is 55.98% identical and 72.11% similar (gaps: 3.04%) to ALPL. The transduction assay described above was repeated with the three isoforms. HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing ALPL, isoform 1 of ALPL, isoform 2 of ALPL, isoform 3 of ALPL, an AAVR positive control, or a pCMV6 negative control. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL receptor were transduced with an AAV particle comprising the TTM-001 capsid variant or the TTM-002 capsid variant, expressing a Luc2-GFP payload. At 24-hours post-transduction, luciferase activity (RLU) was measured to quantify AAV cellular transduction. As shown in **Table 30**, the increase in transduction observed for TTM-001 and TTM-002 when cells express ALPL did not occur in cells expressing isoform 1, 2, or 3. This

demonstrates that the significant increase in TTM-001 and TTM-002 transduction is a specific function of ALPL.

Table 30. Transduction of TTM-001 and TTM-002 capsid variants as measured by luciferase assay (RLU) in cells expressing ALPL or isoforms thereof

	Luciferase Activity (RLU)					
Capsid	pCMV6 transfected cells	AAVR transfected cells	ALPL transfected cells	Isoform 1 transfected cells	Isoform 2 transfected cells	Isoform 3 transfected cells
TTM-001	5,212	13,981	93,268	4,072	7,456	3,300
TTM-002	2,894	11,261	46,182	1,114	3,465	2,074

[0548] Endogenous ALPL was also removed from the surface of HeLa cells by treatment with increasing concentrations phosphatidylinositol-specific phospholipase C (PI/PLC), which cleaves GPI anchored proteins (0, 1, 3, 6, or 10 U/mL), for 1.5 hours at 37°C. Following PI/PLC treatment, cells were incubated with 1E4 VG/cell for three hours of AAV particles comprising the TTM-002 capsid variant or AAV particles comprising an AAV9 control capsid, cells were then washed to remove free virus, and luciferase activity was measured 24 hours post-transduction (RLU). As shown in **Table 31**, treatment with PI/PLC and removal of the GPI-anchored proteins, significantly decreased transduction by the TTM-002 capsid variant, indicating that increased transduction by TTM-002 in HeLa cells is dependent on a GPI-anchored protein.

Table 31. Transduction of TTM-002 capsid variant or AAV9 control in HeLa as measured by luciferase assay (RLU) following treatment with PI/PLC

	Concentration of PI/PLC				
Capsid	0 U/mL	1 U/mL	3 U/mL	6 U/mL	10 U/mL
AAV9	239.3	193.7	222	207.7	212
TTM-002	1731	474	441	396.7	239.3

[0549] To determine if deletion of the endoplasmic reticulum (ER) localization signal of ALPL affected transduction of the TTM-001 and TTM-002 capsid variants, HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing ALPL, ALP with a deletion of the ER localization signal (ALPL transcript variant 2 that lacks the ER signal (NM_001127501, which is incorporated by reference herein)), or a pCMV6 negative control. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL receptor were transduced with an AAV particle comprising the TTM-001 capsid variant, the TTM-002 capsid variant, or an AAV9 capsid control, expressing a GFP payload. At 24-hours post-transduction, luciferase activity (RLU) was measured to quantify AAV cellular transduction. Data was normalized to fold change in luciferase activity (RLU) compared to the pCMV6 control. As shown in **Table 33**, the increase in transduction observed for TTM-001 and TTM-002 when cells express ALPL did not occur in cells expressing ALPL comprising a deleted ER localization signal, and therefore did not express ALPL on the surface of the cells. Similar results were observed by immunofluorescence microscopy staining for GFP

expression, as no GFP staining was observed in cells transfected with the ALPL mutant comprising a deletion of the ER localization signal that were transduced with AAV particles comprising the TTM-001 and TTM-002 capsid variants. These data demonstrate that the ER localization signal may play an important role in effect ALPL has on the transduction of the TTM-001 and TTM-002 capsid variants.

Table 33. Transduction of TTM-001 and TTM-002 capsid variants as measured by luciferase assay (RLU) in cells expressing ALPL

Capsid	Fold change in luciferase activity relative to the pCMV6 control cells		
	Cells transfected with pCMV6	Cells transfected with ALPL	Cells transfected with ALPL comprising a deletion of the ER localization signal
AAV9	1	1	1
TTM-001	1	22	2
TTM-002	1	24	2

[0550] To determine if the TTM-001 and TTM-002 capsid variants could bind to both the human ALPL protein (NM_000478.6, which is incorporated by reference herein) and the mouse ALPL ortholog (NM_001287172.1, which is incorporated by reference herein), HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing human ALPL, the murine ortholog of ALPL, or a pCMV6 negative control. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL receptor were transduced with an AAV particle comprising the TTM-001 capsid variant, the TTM-002 capsid variant, or an AAV9 control capsid protein, expressing a Luc2-GFP payload. At 24-hours post-transduction, luciferase activity (RLU) was measured to quantify AAV cellular transduction. As shown in **Table 34**, the increase in transduction observed for TTM-001 and TTM-002 when cells express human ALPL was also observed in cells expressing the murine ALPL ortholog. These luciferase results were also confirmed by immunofluorescence microscopy staining for GFP. These data indicate that the murine ALPL protein is also a receptor for the TTM-001 and TTM-002 capsid variants.

Table 34. Transduction of TTM-001 and TTM-002 capsid variants as measured by luciferase assay (RLU) in cells expressing ALPL

Capsid	Fold change in luciferase activity relative to the pCMV6 control cells		
	Cells transfected with pCMV6	Cells transfected with human ALPL protein	Cells transfected with murine ALPL protein
AAV9	1	1	2
TTM-001	1	22	18
TTM-002	1	24	24

[0551] To determine if the TTM-002 capsid variants could bind also bind to the cynomolgus ALPL protein (XM_005544525, which is incorporated by reference herein), HEK 293T cells were transfected via calcium phosphate transfection with a plasmid expressing human ALPL, the ortholog of ALPL in cynomolgus macaques (*Macaca fascicularis*), the murine ortholog of ALPL (NM_001287172.1, which is incorporated by reference herein), an AAVR positive control (universal

AAV entry factor involved in AAV transduction), or a pCMV6 negative control. At 24 hours post-transfection, the HEK 293T cells expressing the ALPL receptor were transduced with an AAV particle comprising the TTM-002 capsid variant or an AAV9 control capsid protein, expressing a Luc2-GFP payload. At 24-hours post-transduction, luciferase activity (RLU) was measured to quantify AAV cellular transduction. As shown in **Table 35**, the increase in transduction observed for TTM-002 when cells express human ALPL and murine ortholog was also observed in cells expressing the ALPL ortholog in cynomolgus macaques. These luciferase results were also confirmed by immunofluorescence microscopy staining for GFP. These data indicate that the ALPL protein in cynomolgus macaques is also a receptor for the TTM-002 capsid variant.

Table 35. Transduction of the TTM-002 capsid variant as measured by luciferase assay (RLU) in cells expressing human ALPL, murine ALP, and cynomolgus ALPL

Receptor	Fold change in luciferase activity relative to the pCMV6 control cells	
	AAV9	TTM-002
pCMV6	1.0	1.0
AAVR	4.9	7.3
Human ALPL	0.5	60.3
Cynomolgus ALPL	0.4	81.5
Murine ALPL	1.0	57.2

[0552] Direct binding and specific interaction of the TTM-002 capsid variant and the AAV9 capsid control to ALPL was measured by Surface Plasmon Resonance (SPR) on a Biacore 8K instrument. His-tagged ALPL was first captured on a CM5 sensor chip pre-immobilized with anti-His antibody by passing 5 µg/ml of ALPL for 240 seconds (**FIGs. 2A-2B**). AAV9 or TTM-002 and buffer were then passed over ALPL to monitor rates of association and dissociation, respectively. The concentration of AAV used was from 0.0625 to 1 nM (e.g., 0.0625 nM, 0.125 nM, 0.25 nM, 0.5 nM, and 1 nM; **FIGs. 2A-2B**) and association/dissociation rates were monitored for 120 seconds. The surface was regenerated using two pulses of 10 mM glycine pH 1.7 for 30 seconds. A flow rate of 30 µl/min was used for all steps and the running buffer used was PBS-P+. In a second experiment, the AAV9 control or TTM-002 capsid variant were immobilized on the CM5 sensor chip (**FIGs. 2C-2D**). His-tagged ALPL and buffer were then passed over the AAV9 control or TTM-002 capsid variant to monitor rates of association and dissociation. The concentration of ALPL used was from 15.625 to 250 nM (e.g., 15.625 nM, 32.25 nM, 62.5 nM, 125 nM, and 250 nM; **FIGs. 2C-2D**).

[0553] As shown in **FIG 2A** and **FIG. 2C**, TTM-002 was capable of directly and specifically binding to ALPL in a dose dependent manner, where as AAV9 showed no binding (**FIG. 2B** and **FIG. 2D**). The dissociation constant (K_D) was quantified for the TTM-002 capsid variant binding to ALPL, and was determined to be approximately 32 nM (k_{on}: 3.2e4 1/Ms; k_{off}: 1.27e-3 1/s) (**Table 41**). For this experiment, the density of TTM-002 on the chip was approximately 6300RU.

[0554] Additional experiments were performed varying the density of the TTM-002 capsid variant on the chip. His-tagged ALPL and buffer were then passed over the AAV9 control or TTM-

002 capsid variant to monitor rates of association and dissociation. The concentration of ALPL used was from 15.625 to 250 nM (e.g., 15.625 nM, 32.25 nM, 62.5 nM, 125 nM, and 250 nM, As shown in **Table 41**, despite varying the density of the TTM-002 capsid variant, the affinity values of TTM-002 to ALPL were similar.

Table 41. Binding Affinities of varying densities of TTM-002 to ALPL

Density of TTM-002	k_{on} (1/MS)	k_{off} (1/s)	K_D (nM)
6300 RU	3.98e4	1.27e-3	32
6680 RU	4.36e4	1.16e-3	26.8
2400 RU	3.94e4	1.85e-3	47.0
127 RU	9.8e4	5.1e-3	52.3

[0555] The dissociation of the TTM-002 capsid variant from the ALPL receptor was also investigated at a low pH, as lower affinity has been observed at more acidic pH values for other receptors which can promote efficient transcytosis. pH dependence on interaction of the TTM-002 capsid variant and the AAV9 capsid control to ALPL was measured by Surface Plasmon Resonance (SPR) on Biacore 8K instrument at a pH of 7.4 during association phase and a pH of 7.4 or 5.5 during dissociation phase. The TTM-002 capsid variant was immobilized on the CM5 sensor chip. His-tagged ALPL and buffer was then passed over the AAV9 control or TTM-002 capsid variant to monitor the rate of association at pH 7.4 and dissociation, respectively, at a pH of 7.4 (**FIG. 3A**) or a pH of 5.5 (**FIG. 3B**). The concentration of ALPL used was from 0 to 250 nM (e.g., 0, 7.8 nM, 15.6 nM, 32.25 nM, 62.5 nM, 125 nM, and 250 nM; **FIGs. 3A-3B**). As shown in **FIGs. 3A-3B**, the rate of dissociation between the TTM-002 capsid variant and ALPL increased when the pH was decreased from 7.4 to 5.5, indicating a pH dependent dissociation between the TTM-002 capsid variant and the ALPL receptor.

[0556] Additionally, siRNAs were used to knockdown endogenous levels of ALPL in HeLa cells. HeLa cells were transfected with one of two siRNAs targeting ALPL, both siRNAs targeting ALPL, or a non-ALPL targeting siRNA control using lipofectamine 2000 (5pmol of the siRNA per well of 96-well plate). At 48-hours post transfection, the cells were transduced with 1E4 VG/cell of AAV particles comprising a TTD-002 capsid variant or an AAV9 control capsid and a viral genome encoding a Luc2-GFP payload. At 24-hours post transduction, luciferase activity (RLU) was measured to quantify AAV cellular transduction (**FIG 4**). siRNA mediated knockdown of ALPL led to a 60% reduction in TTM-002 transduction, indicating that knockdown of endogenous ALPL expression inhibits TTM-002 transduction.

[0557] To determine if blocking the ALPL receptor via an anti-ALPL antibody would reduce cellular transduction of the TTM-002 capsid variant, HeLa cells were incubated with 0, 3.125, 6.25, or 12.5 µg/mL of an antibody against ALPL or an IgG isotype-control antibody for 1 hour at 4°C. Following incubation, the cells were transduced with 1E4 VG/cell of AAV particles comprising the TTM-002 capsid variant or an AAV9 control expressing a GFP-luciferase payload. After a 4 hour incubation period the AAV and media were removed and replaced with fresh media and luciferase

activity was measured (RLU) 24 hours post transduction. As shown in **Table 36**, increasing concentrations of the anti-ALPL antibody resulted in decreased TTM-002 transduction in a dose dependent manner, as compared to the isotype control. Similar results were seen by immunohistochemistry. The AAV9 control showed similar levels of transduction in the presence of the anti-ALPL antibody or the isotype control. These data indicated that blocking access to surface expressed ALPL with this antibody reduces TTM-002 transduction.

Table 36. Transduction of TTM-002 capsid variant in HeLa cells as measured by luciferase assay (RLU) following treatment with an anti-ALPL antibody or IgG isotype control

Capsid	Antibody	Antibody Concentration			
		0 µg/mL	3.125 µg/mL	6.25 µg/mL	12.5 µg/mL
TTM-002	Anti-ALPL antibody	4701	5042.67	5220	5556.67
	IgG isotype control	4113.33	527.33	186.67	136
AAV9	Anti-ALPL antibody	530.67	527.33	451.67	417.33
	IgG isotype control	433.33	457.67	516.33	551.33

[0558] An inhibitor of ALPL was also investigated to determine if blocking the ALPL receptor would reduce cellular transduction of the TTM-002 capsid variant. The small molecule Tissue-Nonspecific Alkaline Phosphatase Inhibitor (TNAPi) (CAS 496014-13-2; 2,5-Dimethoxy-N-(quinolin-3-yl)benzenesulfonamide) was selected for use, as kinetic studies such as Dahl et al. ("Discovery and Validation of a Series of Aryl Sulfonamides as Selective Inhibitors of Tissue-Nonspecific Alkaline Phosphatase (TNAP)," *J Med Chem*, 2009; 52)21):6919-6925), which is incorporated by reference in its entirety, indicate that this inhibitor demonstrated an allosteric inhibition mechanism, and inhibition is uncompetitive with respect to the phosphate donor substrate and non-competitive with respect to the acceptor substrate. The IC_{50} was measured to be 190 nM. The inhibitor or a vehicle control (equivalent amount of DMSO) was added to HeLa cells expressing ALPL 1 hour prior to transduction with an AAV viral particle comprising the TTM-002 capsid variant or an AAV9 control capsid and expressing a GFP-luciferase transgene under the control of a CAG promoter. The cells were then transduced with 1E4 VG/cell of an AAV particle comprising the TTM-002 capsid variant or an AAV9 control capsid. Media and virus were removed 4 hours post transduction and luciferase activity was measured (RLU) 24 hours post transduction. As shown in **FIGs. 5A** and **5B**, increasing concentrations of the TNAPi inhibitor resulted in significantly decreased TTM-002 transduction compared to the no inhibitor control (**FIG. 5A**) and the vehicle control (**FIG. 5B**). The IC_{50} of the TNAPi inhibitor in this assay was calculated to be 0.34 nM (**FIG 5C**). The AAV9 control showed similar levels of transduction in the presence of this inhibitor as compared to the no inhibitor control (**FIG. 5A**). These experiments were repeated using a second inhibitor, SBI-425 (5-((5-chloro-2-methoxyphenyl)sulfonamido)nicotinamide; e.g., as described in Pinkerton et al., "Discovery of 5-((5-chloro-2-methoxyphenyl)sulfonamido)nicotinamide (SBI-425), a potent and orally bioavailable tissue-nonspecific alkaline phosphatase (TNAP) inhibitor," *Bioorg Med Chem Lett.*, 2018; 28(1):31-34, the contents of which are hereby incorporated by reference in their entirety),

which is the drug product of TNAPi. Similar to TNAPi, increasing concentrations of SBI-425 significantly inhibited TTM-002 transduction compared to the vehicle control (**FIG. 6A**), but had no effect on the transduction of AAV9 at the same concentrations (**FIG. 6B**). The IC_{50} of the SBI-425 inhibitor in this assay was also calculated to be 0.34 nM (**FIG 6C**). Similar results were obtained by immunofluorescence microscopy for both inhibitors tested. These data with both the TNAPi and SBI-425 inhibitors, indicated that blocking access to surface expressed ALPL with a small molecule inhibitor significantly inhibited TTM-002 transduction.

[0559] The ability of ALPL to transport the TTM-002 capsid variant across the cell membrane (transcytosis) was also investigated using a transcytosis assay and Madin-Darby Canine Kidney (MDCK) cells engineered to overexpress ALPL. MDCK cells were used as they demonstrate clear apico-basolateral polarity and well defined tight junctions. The MDCK cells were plated and resistance was measured. The MDCK control cells that did not express ALPL demonstrated resistance levels above 1000 Ohm and the MDCK cells expressing ALPL demonstrated resistance levels between 5000 to 7000 Ohm. AAV particles with the TTM-002 capsid variant were then added to the top of the cells and the ability of these particles to move from the top to the bottom chamber was measured by qPCR. The percent of the AAV particles comprising the TTM-002 capsid variant detected in the bottom chamber relative to the input of the particles to the top was then calculated. No transcytosis was observed in the MDCK cells that did not express ALPL. However, the MDCK cells overexpressing ALPL demonstrated highly effective transcytosis of TTM-002, such that MDCK ALPL overexpressing cells demonstrated 149-fold greater virus detected in the bottom chamber (percent to the initial virus) as compared to the MDCK cells that did not express ALPL. Additionally, MDCK ALPL overexpressing cells demonstrated 252-fold greater virus detected in the bottom chamber (percent to the initial virus) of the TTM-002 capsid variant as compared to AAV9. A single MDCK cell clone engineered to express ALPL was selected from the pool of MDCK cells that were engineered to overexpress ALPL. The transcytosis assay was done using MDCK cells expressing ALPL from this single clone. MDCK cells expressing ALPL from this single clone demonstrated 7478-fold greater virus detected in the bottom chamber (percent to the initial virus) of the TTM-002 capsid variant as compared to AAV9.

D. Conclusions

[0560] Taken together, these data demonstrate the ALPL is a likely surface receptor for the TTM-001 and TTM-002 capsid variants, as overexpression led to an increase in TTM-001 and TTM-002 transduction as well as cell binding/internalization, which was specific for ALPL. Enzymatic removal of ALPL from the cell surface, mutating the ER localization signal of ALPL, or knockdown of the ALPL receptor by siRNA, also reduced TTM-002 transduction. Without wishing to be bound by theory, it is believed in some embodiments, that the binding of TTM-001 and TTM-002 to ALPL is part of the mechanism leading to increased crossing of the blood brain barrier relative to the AAV9

control. The highly conserved nature of the ALPL receptor protein across species is predictive of cross-species compatibility of the TTM-001 and TTM-002 capsid variants.

Example 9. Identification of Minimal ligand and alternative ALPL-binding moieties

[0561] Biotinylated peptides corresponding to 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acid sequences from loop IV of the TTM-001 and TTM-002 AAV9 capsid variants, including tandem/multimers thereof, as well as from the loop IV domain of the control AAV9 capsid, are generated to investigate the minimal peptide sequence needed for binding ALPL. In some embodiments, loop IV comprises positions 449-475 (e.g., amino acids KTINGSGQNQQTLKFSVAGPSNMAVQG (SEQ ID NO: 6404)), numbered according to SEQ ID NO: 138. In some embodiments loop IV comprises positions 449-460 (e.g., amino acids KTINGSGQNQQT (SEQ ID NO: 6405)), numbered according to SEQ ID NO: 138. The AAV9 capsid variants, TTM-001 (SEQ ID NO: 981 (amino acid) and 983 (DNA), comprising SEQ ID NO: 941) and TTM-002 (SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2), are outlined in Table 3 above. The amino acid and DNA sequences of TTM-001 and TTM-002 are provided, e.g., in Tables 4 and 5, respectively.

[0562] These biotinylated peptides are first captured on an SA sensor chip pre-immobilized with streptavidin by passing 5 µg/ml of peptide for 240 seconds. Recombinant ALPL and buffer are then passed over these peptides to monitor rates of association and dissociation, respectively. The concentration of ALPL to be used ranges from 0.0625 to 1 nM and association/dissociation rates are monitored for 120 seconds. The surface is regenerated using two pulses of 10 mM glycine pH 1.7 for 30 seconds. A flow rate of 30 µl/min is used for all steps and the running buffer used was PBS-P+. This will identify minimum sequences from loop IV of the TTM-001 and TTM-002 capsid variants needed for binding the ALPL receptor.

[0563] Alternative constrained conformations of peptides isolated from loop IV of the TTM-001, and TTM-002 capsid variants or the AAV9 control capsid are tested for binding to ALPL. Biotinylated cyclic peptides corresponding to 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 amino acid sequences from loop IV of the TTM-001 and TTM-002 capsid variants, as well as from the loop IV domain of AAV9 control, are generated. Biotinylated peptides are first captured on an SA sensor chip pre-immobilized with streptavidin by passing 5 µg/ml of peptide for 240 seconds. Recombinant ALPL and buffer are then passed over these peptides to monitor rates of association and dissociation, respectively. The concentration of ALPL used ranges from 0.0625 to 1 nM and association/dissociation rates are monitored for 120 seconds. The surface is regenerated using two pulses of 10 mM glycine pH 1.7 for 30 seconds. A flow rate of 30 µl/min is used for all steps and the running buffer used is PBS-P+. This will further identify the minimum sequences from loop IV of the TTM-001 and TTM-002 capsid variants needed for binding the ALPL receptor.

[0564] Direct binding of alternative ALPL binding ligands, including antibody molecules and other protein-based aptamers, to ALPL is measured by Surface Plasmon Resonance (SPR) on Biacore 8K instrument. His-tagged ALPL is first captured on a CM5 sensor chip pre-immobilized with anti-His antibody by passing 5 µg/ml of ALPL for 240 seconds. The AAV9 control or the TTM-001 or TTM-002 capsid variant and buffer are then passed over ALPL to monitor rates of association and dissociation, respectively. The concentration of AAV used ranges from 0.0625 to 1 nM and association/dissociation rates are monitored for 120 seconds. The surface is regenerated using two pulses of 10 mM glycine pH 1.7 for 30 seconds. A flow rate of 30 µl/min is used for all steps and the running buffer used was PBS-P+. This will aid in identification of additional binding moieties capable of binding ALPL and the minimum sequences/components needed for binding.

Example 10. Functionalization of antibody molecules through recombinant approaches at or near C-terminus

[0565] Once the minimum sequences isolated from loop IV of the TTM-001 and TTM-002 capsid variants needed to bind ALPL are identified using the techniques described in Example 1, DNA sequences encoding fusion antibody constructs encoding a therapeutic antibody molecule fused to an ALPL-binding ligand in the CH3 domain of the Fc region are generated. Specifically, the coding sequence for an antibody molecule to a therapeutic protein, is cloned as a full-length antibody, that will be modified to comprise an antigen binding domain (e.g., VH and VL) specific for a therapeutic protein, and an Fc region, wherein peptides comprising sequences of varying lengths isolated from loop IV of the TTM-001 and TTM-002 capsid variants (e.g., ALPL-binding ligands) are fused at or near the C-terminus of the CH3 domain of the Fc region of the antibody molecule. The recombinant antibody molecules fused to the sequences of varying lengths isolated from loop IV of the TTM-001 and TTM-002 capsid variants (e.g., ALPL-binding ligands) or unmodified antibody molecules against the same therapeutic protein are injected into a relevant mouse model. These results will indicate the ability of a recombinant fusion protein, e.g., an antibody molecule, containing an ALPL-binding ligand comprising a portion of the loop IV sequence of the TTM-001 or TTM-002 capsid variants to show improved crossing of the blood brain barrier. Without wishing to be bound by theory, it is believed in some embodiments, that the recombinant antibody molecules fused to the ALPL-binding ligands may exhibit greater biodistribution in the brain and increased efficacy in the brain relative to a control non-modified antibody molecule alone.

[0566] A multispecific antibody molecule comprising a first binding domain for a therapeutic protein and a second binding domain that binds to ALPL (e.g., an anti-ALPL binding domain) is generated. These multispecific antibody molecules are designed to comprise two Fabs for binding a therapeutic protein and ALPL, two scFvs for binding a therapeutic protein and ALPL, or combinations of an Fab and scFv. The multispecific antibody molecule comprising a first binding domain to a therapeutic protein and an anti-ALPL binding domain, a monospecific antibody molecule

that binds to the therapeutic protein, or a multispecific antibody comprising a binding domain to the therapeutic protein and IgG control binding domain is intravenously administered to a relevant mouse model. These results will indicate the ability of a recombinant fusion protein, e.g., a multispecific antibody molecule, containing an ALPL-binding domain or an ALPL-binding ligand, in addition to a therapeutic protein or a second binding domain to a therapeutic protein, to show improved crossing of the blood brain barrier. Without wishing to be bound by theory, it is believed in some embodiments that the recombinant multispecific antibody molecule comprising the anti-ALPL-binding domain and the binding domain to a therapeutic protein will exhibit increased biodistribution and efficacy in the brain relative to a monospecific antibody that binds the therapeutic protein or a multispecific antibody comprising binding domain specific for the therapeutic protein and an IgG control binding domain.

[0567] Without wishing to be bound by theory, it is believed in some embodiments that these examples will demonstrate that multiple ALPL-binding ligands, including those comprising portions of the amino acid sequence of loop IV of the TTM-001 and TTM-002 capsid variants, can be genetically encoded as fusion proteins to therapeutic proteins, to convey an improved ability to cross the blood brain barrier. Similar methodologies as described in this Example can be used to functionalize other therapeutic proteins and enzymes of interest.

Example 11. Functionalization of therapeutic proteins, antibodies, or enzymes of interest through post-translational linker strategies of ALPL-binding moieties to the protein, antibody, or enzyme of interest.

[0568] Methods for stochastic and site-specific methods of labeling amino acids on protein sequences are disclosed by Shadish JA and DeForest CA, Site-Selective Protein Modification: From Functionalized Proteins to Functional Biomaterials. Matter 2020 2:50-70 (the contents of which are hereby incorporated by reference in their entirety). Similarly, numerous methods have also been employed specifically for conjugation of various payloads to antibodies, including linkage chemistries used for antibody-drug conjugates (e.g., as described in Fu et al. Antibody drug conjugate: the "biological missile" for targeted cancer therapy. Signal Transduction and Targeted Therapy 2022 7:93; and Drago et al. Unlocking the potential of antibody-drug conjugates for cancer therapy. Nat Rev Clin Oncol 2021 18:327-344; the contents of which are hereby incorporated by reference in their entirety).

[0569] Numerous methods for chemical- or enzyme-mediated site-specific or site-agnostic introduction of ALPL-binding ligands for functionalization of various active agents including but not limited to therapeutic proteins, antibody molecules that bind therapeutic proteins, and enzymes are tested. Peptides of varying lengths of amino acids derived from the loop IV region of the TTM-001 or TTM-002 capsid variants are chemically modified with NHS-esters and are allowed to react with free amines on the surface of an active agent, such as an antibody molecule that binds a therapeutic protein, a therapeutic protein, or an enzyme. The compositions comprising the modified antibody

molecules, therapeutic proteins, or enzymes that are chemically linked to the peptides derived from the TTM-001 or TTM-002 capsid variants are injected into a relevant mouse model and the ability of these modified compositions to cross the blood-brain barrier is evaluated.

[0570] Additionally, peptides of varying lengths of amino acids derived from the loop IV region of the TTM-001 or TTM-002 capsid variants are linked by a cleavable linker, e.g., a pH-sensitive hydrazone linker, to an active agent such as an antibody molecule that binds a therapeutic protein, a therapeutic protein, or enzyme. The compositions comprising the modified antibody molecules, therapeutic proteins, or enzymes that are linked to the peptides derived from the TTM-001 or TTM-002 capsid variants are injected into a relevant mouse model and the ability of these modified antibodies molecule to cross the blood-brain barrier is evaluated. Without wishing to be bound by theory, it is believed in some embodiments, that the peptide derived from the TTM-001 or TTM-002 capsid variants (the ALPL-binding ligand) will be cleaved from the active agent (e.g., an antibody molecule that binds a therapeutic protein, a therapeutic protein, or enzyme) during transcytosis of the composition across the blood brain barrier, resulting in increased biodistribution of the active agent in the brain upon release to the parenchymal side of the blood brain barrier.

[0571] An alternative post-translational method that can be used to link ALPL-binding ligands to an active agent involves the use of click chemistry. Synthesis of a clickable RGD peptide is obtained by reacting Lys side chain with azido acetic acid on an ALPL-binding ligand, which is then linked to another peptide fragment on an active agent, such as a therapeutic protein, antibody molecule that binds a therapeutic protein, or enzyme. The compositions comprising the modified antibody molecules, therapeutic proteins, or enzymes that are linked to the peptides derived from the TTM-001 or TTM-002 capsid variants are injected into a relevant mouse model and the ability of these modified antibodies molecule to cross the blood-brain barrier is evaluated.

[0572] The ability of active agents such as antibody molecules, therapeutic proteins, and enzymes to cross the blood brain barrier is also assessed through chemically-induced dimerization of synthetic proteins containing ALPL-binding interfaces. The coding sequence of the FRB domain of mTOR1 (mammalian target of Rapamycin complex 1) is cloned as a C-terminal fusion to the Fc domain of an antibody molecule that binds a therapeutic protein, the C-terminus of a therapeutic protein, or the C-terminus of an enzyme. The coding sequence of FKBP12 is cloned as a C-terminal fusion to an antibody molecule that binds ALPL. The modified proteins encoded by these sequences are produced, and allowed to dimerize in the presence of rapamycin or AP20187. These modified proteins are then administered to a relevant mouse model. The ability of the modified proteins to cross the blood brain barrier and accumulate on the parenchymal side of the blood brain barrier is evaluated.

[0573] Collectively, this example provides multiple methods that can be used to chemically link an ALPL-binding ligand to an active agent, such that this conjugated composition can be evaluated for improved ability to cross the blood brain barrier relative to the non-conjugated control.

Example 12: Conjugation of siRNA molecules to increase blood brain barrier crossing

[0574] Peptides of varying lengths of amino acids isolated from loop IV of the TTM-001 or TTM-002 capsid variants, including tandem or multimer orientations thereof, are synthesized. Similar to the method described, e.g., by Eyford et al. A Nanomule Peptide Carrier Delivers siRNA Across the Intact Blood Brain Barrier to Attenuate Ischemic Stroke. Front Mol Biosci 2021 8:611367 (the contents of which are hereby incorporated by reference in their entirety), siRNA molecules against a therapeutic target and conjugates thereof with peptides derived from loop IV of the TTM-001 or TTM-002 capsid variants or control peptides are synthesized. More specifically, the peptides derived from TTM-001 and TTM-002 and the siRNA molecules are produced independently and then are chemically conjugated using the crosslinker succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate. The siRNA-TTM001 or TTM-002 peptide conjugates and the siRNA-control peptide conjugates are administered intravenously to a relevant mouse model. At 24 hours post-intravenous administration, mice are sacrificed and total RNA is isolated from the brains of each animal. qPCR analysis is performed to measure expression of the target gene in the mice treated with the siRNA-TTM001 or TTM-002 peptide conjugates relative to the mice treated with the siRNA-control peptide conjugates, as a measure of the ability of theses siRNA conjugates to cross the blood brain barrier. Without wishing to be bound by theory, it is believed in some embodiments, that siRNA molecules conjugated to peptides isolated from loop IV of the TTM-001 and TTM-002 capsid variants may demonstrate an increased ability to cross the blood brain barrier and increased efficacy relative to non-conjugated siRNA molecules.

[0575] Additionally, as described, e.g., in Yang et al. A microfluidic method for synthesis of transferrin-lipid nanoparticle loaded with siRNA LOR-1284 for therapy of acute myeloid leukemia. Nanoscale 2014 6(16):9742-9751 (the contents of which are hereby incorporated by reference in their entirety), cationic lipid nanoparticles are prepared by ethanol injection method and comprise an siRNA molecule that binds to a therapeutic target. Peptides of varying lengths of amino acids derived from loop IV of the TTM-001 or TTM-002 capsid variants, including tandem or multimer orientations thereof, are synthesized and conjugated to the surface of the nanoparticles. Mice are injected intravenously with the LNPs comprising the siRNA molecule coated with the peptides isolated from the TTM-001 and TTM-002 capsid variants, or a control non-coated LNP comprising the siRNA molecule. At 24 hours post-intravenous administration, mice are sacrificed and total RNA is isolated from the brains of each animal. qPCR analysis is performed to measure expression of the target gene in the mice treated with the LNPs comprising the siRNA molecule that are coated with the TTM001 or TTM-002 peptides relative to the mice treated with the control non-coated LNPs comprising the siRNA molecule, as a measure of the ability of theses siRNA conjugates to cross the blood brain barrier. Without wishing to be bound by theory, it is believed in some embodiments, that siRNA molecules comprise within LNPs coated with peptides isolated from loop IV of the TTM-001 and

TTM-002 capsid variants may demonstrate an increased ability to cross the blood brain barrier and increased efficacy relative to siRNA molecules comprised with in non-coated LNPs.

[0576] Taken together, this example provides methods that can be used to link an ALPL-binding ligand (e.g., a peptide isolated from loop IV of the TTM-001 and TTM-002 capsid variants) to siRNA agent or LNP comprising a siRNA molecule, such that this composition can be evaluated for improved ability to cross the blood brain barrier relative to non-linked controls.

Example 13: Binding of Peptide Ligands to ALPL

[0577] This example investigates the ability of peptides of varying lengths comprising the insert sequences from loop IV of the TTM-001 and TTM-001 capsid variants, which are SPHSKA (SEQ ID NO: 941) and HDSPHK (SEQ ID NO: 2), respectively, plus or minus amino acid residues present N-terminal and/or C-terminal relative to the insert sequences. The peptide sequences tested are provided in **Table 37**. Each peptide was modified at the N-terminus by biotin followed by a GSGS linker (SEQ ID NO: 6409) and at the C-terminus by amidation.

Table 37. Peptide Ligands (GSGS (SEQ ID NO: 6409) is the linker; bolded sequence is the peptide insert of TTM-002 (HDSPHK (SEQ ID NO: 2)))

Peptide Span	TTM-002 Derived Peptide Sequence	SEQ ID NO
n (insert)	GSGS**HDSPHK**	4498
n-1_n+1	GSGSG**HDSPHK**S	4499
n-2_n+2	GSGSNG**HDSPHK**SG	4500
n-3_n+3	GSGSING**HDSPHK**SGQ	4501
n-4_n+4	GSGSTING**HDSPHK**SGQN	4502
n-5_n+5	GSGSKTING**HDSPHK**SGQNQ	4503
Control: -5_+5 without insert	GSGSKTINGSGQNQ	4504

[0578] Binding and interaction between the peptides and the ALPL receptor were measured by Surface Plasmon Resonance (SPR) on a Biacore 8K instrument. The biotinylated peptides in **Table 37** were captured on the streptavidin biosensor chip and 1000 nM of ALPL and buffer were passed over the peptides to monitor binding. No binding was detected with these peptides. His-tagged ALPL was then captured on the chip pre-immobilized with anti-His antibody, and 50 µM of the peptides in **Table 37** or 250 nM of the AAV9 control capsid, the TTM-001 capsid variant (SEQ ID NO: 981), or the TTM-002 capsid variant (SEQ ID NO: 982) were flowed over the ALPL to check for binding. No detectable binding was observed with the peptides or the AAV9 control. Binding to ALPL was detected for the TTM-001 and TTM-002 capsid variants, confirming what was observed in **Example 8** above. The longer peptides derived from TTM-002, GSGSLYYLSKTINGHDSPHKSGQNQQTLKF (SEQ ID NO: 19) and GSGSRLMNPLIDQYLYYLSKTINGHDSPHKSGQNQQTLKFSVAGPSNMAV (SEQ ID NO: 20), were also flowed over the his-tagged ALPL captured on the CM5 sensor chip pre-immobilized with anti-His antibody. As was observed with the peptides in **Table 37**, no binding to ALPL was observed with the longer peptides.

[0579] By LC-MS, the overall serine phosphorylation was calculated to be 60% across the TTM-002 capsid variant (SEQ ID NO: 982), which was much higher than the overall serine phosphorylation levels measured in other AAV9 capsid variants tested comprising modifications in loop IV or loop VIII. The overall serine phosphorylation was calculated to be 4.7% across the TTM-019 capsid variant (SEQ ID NO: 52), 1.9% across the TTM-018 capsid variant (SEQ ID NO: 51), and 1.5% across the TTD-001 AAV9 capsid variant comprising a loop VIII modification, the sequence and characterization of which can be found in WO 2021/230987 (the contents of which are hereby incorporated by reference in their entirety). Additionally, only TTM-002 demonstrated an 80-90% phosphorylation level on the serine in the SPHK (SEQ ID NO: 6398) motif that is present in loop IV of the TTM-002 capsid variant (SEQ ID NO: 982). This serine is present at position 456 in the TTM-002 capsid variant, numbered according to SEQ ID NO: 982. This increased phosphorylation was detected in the TTM-002 capsid variant encapsulating two different payloads. Without wishing to be bound by theory, it is believed in some embodiments, that this SPHK (SEQ ID NO: 6398) motif present in the TTM-002 capsid variant is a consensus motif for the CDK5 kinase and peptides comprising this motif and a modification, e.g., a phosphate group, on this serine, may demonstrate binding to ALPL, as compared to their non-phosphorylated counterparts.

[0580] Additional peptides derived from the TTM-002 capsid variant and peptide insert of HDSPHK (SEQ ID NO: 2) that is present in loop IV, were then generated that comprised a phosphorylated serine to test binding with the ALPL receptor. These phosphorylated peptides are provided in **Table 38**. Each peptide was also modified at the N-terminus by biotin followed by a GSGS linker (SEQ ID NO: 6409) and at the C-terminus by amidation.

Table 38. Phosphorylated Peptide Ligands (GSGS (SEQ ID NO: 6409) is the linker)

Peptide span	TTM-002 Derived Phospho Peptide Sequence	SEQ ID NO	Peptide length
n-2_n+2	GSGSNGHD**pS**PHKSG	4512	10
n-5_n+5	GSGSKTINGHD**pS**PHKSGQNQ	4513	16
n-8_n+11	GSGSYLSKTINGHD**pS**PHKSGQNQQTLKFS	4514	25

[0581] Binding and interaction between the phosho-peptides and the ALPL receptor were also measured by SPR. The biotinylated and phosphorylated peptides in **Table 38** were captured on the streptavidin biosensor chip along with their non-phosphorylated counterparts, GSGSNGHDSPHKSG (SEQ ID NO: 4500) and GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503). ALPL and buffer were passed over the peptides to monitor binding. The concentration of ALPL used ranged from 0 to 500 nM (0 nM, 125 nM, 250 nM, and 500 nM). As shown in **FIGs. 7A-7B**, both phospho-peptides demonstrated dose dependent binding to ALPL, whereas no binding was observed with their non-phosphorylated counterparts (**FIGs. 7A-7B**).

[0582] In a second experiment, his-tagged ALPL was first captured on a CM5 sensor chip pre-immobilized with anti-His antibody. The phosphorylated peptides in **Table 38** and their non-phosphorylated counterparts, GSGSNGHDSPHKSG (SEQ ID NO: 4500) and

GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) and buffer were flowed over the ALPL protein to monitor binding. The concentration of the peptides ranged from 0 to 50 µM (0 µM, 1.56 µM, 3.125 µM, 6.25 µM, 12.5 µM, 25 µM, or 50 µM). Both phospho-peptides showed low signal with dose dependency, but a higher signal was observed for SEQ ID NO: 4513 relative to SEQ ID NO: 4512 (**FIGs 8A-8B**).

[0583] Binding of the phospho-peptides of SEQ ID NO: 4512 and SEQ ID NO: 4513 to ALPL was also investigated using Bio Layer Interferometry (BLI)/Octet. Biotinylated peptides were first loaded on streptavidin biosensor tips followed by placing tips in tips in varying concentration of his-tagged ALPL ranging from 1.56 to 100 nM to measure binding kinetics. Similar to what was observed by SPR, both of the phospho-peptides of SEQ ID NO: 4512 and SEQ ID NO: 4513 demonstrated binding to ALPL, but no binding was observed with their non-phosphorylated counterparts, GSGSNGHDSPHKSG (SEQ ID NO: 4500) and GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503) (**FIGs. 9A-9B**). The loading level of the phosphorylated peptides and the non-phosphorylated peptides was comparable which confirmed that the lack of binding observed for the non-phosphorylated peptides is not due to a lower loading level. Also, as shown in **Table 39**, the dissociation constant (K_D) was quantified for both of the phospho-peptides of SEQ ID NO: 4512 and SEQ ID NO: 4513. The K_D for SEQ ID NO: 4512 was 112 nM and the K_D for SEQ ID NO: 4513 was of 20.7 nM. These binding affinities of the phospho-peptides derived from the TTM-002 capsid variant to ALPL were similar to the binding affinity of the TTM-002 capsid variant to ALPL as provided in **Table 41** of Example 8. Additionally, no binding to ALPL was observed with a phosphorylated tau peptide, indicating that the interaction is specific to ALPL and the phospho-peptides derived from TTM-002. The binding of the phospho-peptide of SEQ ID NO: 4513 to ALPL was also tested at a pH of 5.5, and no binding was observed.

Table 39. Binding affinities for SEQ ID NO: 4512 and SEQ ID NO: 4513 to ALPL

GSGSNGHD**pS**PHKSG **(SEQ ID NO: 4512)**		GSGSKTINGHD**pS**PHKSGQNQ **(SEQ ID NO: 4513)**	
k_{on}	2.3e4 1/Ms	k_{on}	3.2e4 1/Ms
k_{off}	2.6e-3 1/s	k_{off}	6.7e-4 1/s
K_D	112 nM	K_D	20.7 nM

[0584] Binding of the phospho-peptides of SEQ ID NO: 4512 and 4513 along with their non-phosphorylated counterparts, GSGSNGHDSPHKSG (SEQ ID NO: 4500) and GSGSKTINGHDSPHKSGQNQ (SEQ ID NO: 4503), respectively, to ALPL was also confirmed by ELISA with ALPL coated on the wells of microplate. (**FIGs. 10A-10B**). Both of the phospho-peptides of SEQ ID NO: 4512 and 4513 but not their non-phosphorylated counterparts demonstrated dose-dependent binding to ALPL (**FIGs. 10A-10B**). The EC_{50} value calculated for the phospho-peptide of SEQ ID NO: 4512 was 1.243 µg/mL and the EC_{50} value calculated for the phospho-peptide of SEQ ID NO: 4513 was 10.05 µg/mL.

[0585] Taken together, these data indicate that phosphorylated peptides derived from the TTM-002 capsid variant are capable of binding ALPL, as measured by at least three independent methods (Biacore, Octet, and ELISA); and that this binding appeared to be sequence specific as other phosphorylated control peptides showed no binding.

Example 14: Binding and Transcytosis of Antibodies that bind to ALPL

[0586] This example investigates the ability of exemplary antibodies to bind ALPL and to also be transported across the cell membrane (transcytosis).

[0587] Several anti-ALPL antibodies provided in **Table 40** were tested to determine if they were able to bind ALPL by ELISA and Surface Plasmon Resonance (SPR) on a Biacore. For the SPR/Biacore assay, the antibodies were captured on the chip and 1.6 to 1000 nM of ALPL was flowed over the chip. Anti-ALPL antibodies #3, #8, #9, #15, #16, #18, #20, and #22 were capable of binding ALPL as measured by ELISA and the anti-ALPL antibodies #5, #8, #9, #15, #16, #19, #22, and #30 were capable of binding ALPL as measured by SPR. Also, as shown in **Table 42**, the dissociation constant (K_D) for binding to human ALPL/binding affinity was quantified for anti-ALPL antibody #9 (Ab 9) and antibody # 22 (Ab 22).

Table 40. Exemplary Anti-ALPL antibodies

Antibody Number	Clone	Cat #	Name
3	2F4	Ab126820	Anti-Alkaline Phosphatase, Tissue Non-Specific antibody [2F4]
5	BGN/03/662	NB100-66385-0.025mg	Alkaline Phosphatase, Liver/Bone/Kidney Antibody (BGN/03/662) - BSA Free
8	4H1	H00000249-M01	Alkaline Phosphatase, Tissue Non-Specific Antibody (4H1)
9	B4-78	MAB1448-SP	Human Alkaline Phosphatase/ALPL Antibody
14	Clone 928929	MAB29092-SP	Human Alkaline Phosphatase/ALPL Antibody
15	TRA-2-49	NB600-540	Alkaline Phosphatase, Tissue Non-Specific Antibody (TRA-2-49) - BSA Free
16	OTI31A	TA809117	Alkaline Phosphatase (ALPL) Mouse Monoclonal Antibody [Clone ID: OTI3A1]
18	13B28	O29950	029950 Mouse Anti-ALPL (Alkaline phosphatase liver/bone/kidney)
19	4H1	LS-C196650-100	Monoclonal Mouse anti-Human ALPL / Alkaline Phosphatase Antibody (clone 4H1, WB) LS-C196650
20	1.H.3	DCABH-5801	Anti-ALPL monoclonal antibody
22	3H414	P4071-18	P4071-18 Mouse Anti-Phosphatase, Alkaline, Liver, Bone, Kidney (Akp2, ALPI, ALPL)
29	TRA-2-49/6E	MAB4349	Anti-TRA-2-49 Antibody, Liver/Bone/Kidney Alkaline Phosphatase, clone TRA-2-49/6E
30	USB-3-65/3K	DMABT-H21767	Anti-TRA-2-54 monoclonal antibody
31	077	MA5-41041	ALPL Recombinant Rabbit Monoclonal Antibody
32	SA40-00	MA5-42414	ALPL Recombinant Rabbit Monoclonal Antibody
33	034	100356-R034	Recombinant Anti-Alkaline Phosphatase/ALPL Antibody, Rabbit Monoclonal
34	071	100356-R071	Recombinant Anti-Alkaline Phosphatase/ALPL Antibody, Rabbit Monoclonal
35	1G9	Orb612236	Alkaline Phosphatase antibody

36	3G2	MA5-24845	ALPL Recombinant Rabbit Monoclonal Antibody (3G2)
37	7H11L3	702454	ALPL Recombinant Rabbit Monoclonal Antibody (7H11L3)
38	DBF-1	MAB19541	ALPL monoclonal antibody
Poly #1	N/A	215232	Rabbit Anti-Alkaline Phosphatase (Tissue-nonspecific isozyme, AP-TNAP, TNSALP, ALPL)
Poly #2	N/A	305750	AP-TNAP, TNSALP, Alkaline phosphatase liver/bone/kidney isozyme, ALPL
Poly #3	N/A	350543	Rabbit Anti-Alkaline Phosphatase, NT (ALPL, Alkaline phosphatase, liver/bone/kidney)

Table 42. Binding affinities for anti-ALPL antibody #9 and #22 to human ALPL

Ab 9		**Ab 22**	
k_{on}	6.8e3 1/Ms	k_{on}	3.1e4 1/Ms
k_{off}	1.5e-3 1/s	k_{off}	8.3e-4 1/s
K_D	215 nM	K_D	26.4 nM

[0588] Several anti-ALPL antibodies provided in **Table 40** were also tested for their ability to compete for binding with an AAV particle comprising the TTM-002 capsid variant SEQ ID NO: 982 (amino acid) and 984 (DNA), comprising SEQ ID NO: 2). hCMEC/D3 cells overexpressing ALPL were incubated with murine and rabbit monoclonal and rabbit polyclonal anti-ALPL antibodies for 1 hour, and then transduced with AAV particles comprising the TTM-002 capsid variant and comprising a viral genome encoding a GFP-luciferase transgene. Luciferase activity was measured (RLU) to determine if the AAV particle was capable of transducing the cells. No luciferase signal was indicative of an antibody that competes with the TTM-002 capsid, suggestive of an identical binding site or pocket on ALPL. As shown in **FIG. 12A**, anti-ALPL antibody #9 (Ab 9), #22 (Ab 22), and #29 had little to no detectable luciferase activity, indicating competition with the TTM-002 capsid variant. In fact, pre-incubation with Ab 9 led to no luciferase activity.

[0589] The anti-ALPL antibodies Ab 9 and Ab 22 were then added to wild-type hCMEC/D3 that do not overexpress ALPL and hCMEC/D3 cells that were engineered to express ALPL and allowed to internalize for five hours. The cells were then washed to remove any unbound antibodies, fixed, and incubated overnight with an anti-mouse FITC antibody to measure antibody internalization by fluorescence microscopy. Ab 9 and Ab 22 were internalized only in the hCMEC/D3 cells that were engineered to express ALPL. No internalization was observed in the wild-type hCMEC/D3 that did not overexpress ALPL.

[0590] Anti-ALPL antibodies Ab 9 and Ab 22 were then tested to determine if they were capable of binding ALPL and being subsequently transported across the cell membrane (transcytosis). The MDCK ALPL-expressing cells from the single clone generated in Example 8 were plated at a density of 200,000 cells in 250 µl in complete growth media at the apical side of a Transwell® insert (12-well, 0.4 µm pore size Transwell®-65 mm). The cells were incubated for 2-3 days to allow for polarization and electrical resistance was measured to calculate the integrity of tight junctions. The antibodies were then added to the top of the chamber at a concentration of 12 µg in 250 µl of media. The two

anti-ALPL antibodies, Ab 9 and Ab 22, were tested along with a PT3 control antibody that does not bind ALPL and an anti-mouse IgG1 isotype control (MOPC). The cells and the antibodies were incubated together overnight. Media was then collected from both chambers, and the concentrations of the antibodies in the top and bottom chambers were measured using a mouse IgG alphaLISA, in order to determine the ability of the antibodies to move from the top to the bottom chamber (transcytosis). **FIG. 11A** shows that the control antibodies as well as the ALPL binding antibodies were loaded at similar levels to the top of the cells. As shown in **FIG. 11B**, the MDCK cells overexpressing ALPL demonstrated highly effective transcytosis of Ab 9 that binds ALPL as evidenced by the high levels of the antibody quantified in the bottom chamber. Very little of the PT3 antibody, which does not bind ALPL, Ab 22 which is capable of binding ALPL, or the MOPC isotype control antibody, were detected in the bottom chamber. The percentage of the antibody detected in the bottom chamber to the load was also quantified (**FIG. 11C**) and the percentage of the Ab 9 that binds to ALPL in the bottom chamber relative to the load was greatly increased relative to the PT3 and MOPC controls and Ab 22 (**FIG. 11C**). Also as shown in **FIG. 11C**, MDCK cells that were not engineered to express ALPL, demonstrated little to no transcytosis of Ab 9 that binds ALPL or the PT3 control antibody.

[0591] Taken together, these data demonstrate that certain anti-ALPL antibodies are capable of binding ALPL and being transported across the cell membrane *in vitro*.

We claim:

1. A composition, e.g., a fusion molecule or a conjugate molecule, comprising:

(i) a ligand that binds to alkaline phosphatase (ALPL); and

(ii) an active agent, e.g., a therapeutic agent or a diagnostic agent,

wherein the ligand is fused or coupled to the active agent,

wherein the ligand is capable of binding ALPL:

(a) at a K_D of at least about 10-250 nM, e.g., when measured by an SPR assay, e.g., as described in Example 8; and/or

(b) in a pH dependent manner, wherein the ligand binds to ALPL at physiological pH and/or does not substantially bind ALPL at an acidic pH, e.g., as measured by an assay e.g., an SPR or Biacore assay, e.g., as described in Example 8 or 13.

2. The composition of claim 1, wherein the ligand binds human, cynomolgus, or murine ALPL.

3. The composition of claim 1 or 2, wherein the ligand is or comprises a peptide, a protein, an antibody molecule, a nucleic acid molecule (e.g., an aptamer), or a small molecule.

4. The composition of any one of claims 1-3, wherein the ligand is conjugated to the active agent via a linker.

5. The composition of any one of claims 1-3, wherein the ligand is fused directly or indirectly via a linker to the active agent, e.g., as part of a fusion peptide or protein.

6. The composition of any one of claims 1-5, wherein the ligand is not a component of a viral particle, e.g., an adeno-associated viral (AAV) particle.

7. The composition of any one of claims 3-6, wherein the linker is a cleavable linker or a non-cleavable linker.

8. The composition of embodiment 7, wherein the cleavable linker is a pH sensitive linker or an enzyme sensitive linker, optionally wherein:

(i) the pH sensitive linker comprises a hydrazine/hydrazone linker or a disulfide linker; or

(ii) the enzyme sensitive linker comprises a peptide-based linker, e.g., a peptide linker sensitive to a protease (e.g., a lysosomal protease); or a beta-glucuronide linker.

9. The composition of claim 7, wherein the non-cleavable linker is a linker comprising a thioether group or a maleimidocaproyl group.

10. The composition of any one of claims 1-9, wherein the ligand is or comprises a protein or a peptide comprising an amino acid sequence having the following formula: [N1]-[N2]-[N3], wherein:

(i) optionally [N1] comprises X1, X2, and X3, wherein at least one of X1, X2, or X3 is G;

(ii) [N2] comprises the amino acid sequence of SPH, optionally wherein S comprises a modification, e.g., comprises a phosphate group;

(ii) [N3] comprises X4, X5, and X6, wherein at least one of X4, X5, or X6 is a basic amino acid, e.g., a K or R.

11. The composition of claim 10, wherein:

(a) position X4 of [N3] is: K, S, A, V, T, G, F, W, V, N, or R;

(b) position X5 of [N3] is: S, K, T, F, I, L, Y, H, M, or R; and/or

(c) position X6 of [N3] is: G, A, R, M, I, N, T, Y, D, P, V, L, E, W, N, Q, K, or S.

12. The composition of claim 10 or 11, wherein:

(i) [N3] comprises KSG, SKA, ARM, VKS, ASR, VKI, KKN, VRM, RKA, KTS, KFG, KIG, KLG, KTT, KTY, KYG, SKD, SKP, TRG, VRG, KRG, GAR, KSA, KSR, SKL, SRA, SKR, SLR, SRG, SSR, FLR, SKW, SKS, WKA, VRR, SKV, SKT, SKG, GKA, TKA, NKA, SKL, SKN, AKA, KTG, KSL, KSE, KSV, KSW, KSN, KHG, KSQ, KSK, KLW, WKG, KMG, KMA, or RSG; and/or

(ii) [N2]-[N3] comprises SPHKSG (SEQ ID NO: 946), SPHSKA (SEQ ID NO: 941), SPHARM (SEQ ID NO: 947), SPHVKS (SEQ ID NO: 948), SPHASR (SEQ ID NO: 949), SPHVKI (SEQ ID NO: 950), SPHKKN (SEQ ID NO: 954), SPHVRM (SEQ ID NO: 955), SPHRKA (SEQ ID NO: 956), SPHKFG (SEQ ID NO: 957), SPHKIG (SEQ ID NO: 958), SPHKLG (SEQ ID NO: 959), SPHKTS (SEQ ID NO: 963), SPHKTT (SEQ ID NO: 964), SPHKTY (SEQ ID NO: 965), SPHKYG (SEQ ID NO: 966), SPHSKD (SEQ ID NO: 967), SPHSKP (SEQ ID NO: 968), SPHTRG (SEQ ID NO: 972), SPHVRG (SEQ ID NO: 973), SPHKRG (SEQ ID NO: 974), SPHGAR (SEQ ID NO: 975), SPHKSA (SEQ ID NO: 977), SPHKSR (SEQ ID NO: 951), SPHSKL (SEQ ID NO: 960), SPHSRA (SEQ ID NO: 969), SPHSKR (SEQ ID NO: 978), SPHSLR (SEQ ID NO: 952), SPHSRG (SEQ ID NO: 961), SPHSSR (SEQ ID NO: 970), SPHFLR (SEQ ID NO: 979), SPHSKW (SEQ ID NO: 953), SPHSKS (SEQ ID NO: 962), SPHWKA (SEQ ID NO: 971), SPHVRR (SEQ ID NO: 980), SPHSKT (SEQ ID NO: 4731), SPHSKG (SEQ ID NO: 4732), SPHGKA (SEQ ID NO: 4733), SPHNKA (SEQ ID NO: 4734), SPHSKN (SEQ ID NO: 4735), SPHAKA (SEQ ID NO: 4736), SPHSKV (SEQ ID NO: 4737), SPHKTG (SEQ ID NO: 4738), SPHTKA (SEQ ID NO: 4739), SPHKSL (SEQ ID NO: 4740), SPHKSE (SEQ ID NO: 4741), SPHKSV (SEQ ID NO: 4742), SPHKSW (SEQ ID NO: 4743), SPHKSN (SEQ ID NO: 4744), SPHKHG (SEQ ID NO: 4745), SPHKSQ (SEQ ID NO: 4746), SPHKSK (SEQ ID NO: 4747), SPHKLW (SEQ ID NO: 4748), SPHWKG (SEQ ID NO: 4749), SPHKMG (SEQ ID NO: 4750), SPHKMA (SEQ ID NO: 4751), or SPHRSG (SEQ ID NO: 976);

13. The compositions of any one of claims 10-12, wherein:

(a) position X1 of [N1] is: G, V, R, D, E, M, T, I, S, A, N, L, K, H, P, W, or C;

(b) position X2 of [N1] is: H, S, V, L, N, D, R, P, G, T, I, A, E, Y, M, or Q; and/or

(c) position X3 of [N1] is: D, G, C, L, E, Y, H, V, A, N, P, or S.

14. The composition of any one of claims 10-13, wherein:

(i) [N1] comprises GHD, GSG, GQD, VSG, CSG, GRG, CSH, GQS, GSH, RVG, GSC, GLL, GDD, GHE, GNY, MSG, RNG, TSG, ISG, GPG, ESG, SSG, GNG, ASG, NSG, LSG, GGG, KSG, HSG, GTG, PSG, GSV, RSG, GIG, WSG, DSG, IDG, GLG, DAG, DGG, MEG, ENG, GSA, KNG, KEG, AIG, GYD, GHG, GRD, GND, GPD, GMG, GQV, GHN, GHP, or GHS;

(ii) [N1]-[N2] comprises GHDSPH (SEQ ID NO: 4784), GSGSPH (SEQ ID NO: 4695), GQDSPH (SEQ ID NO: 4785), VSGSPH (SEQ ID NO: 4786), CSGSPH (SEQ ID NO: 4787), GRGSPH (SEQ ID NO: 4788), CSHSPH (SEQ ID NO: 4789), GQSSPH (SEQ ID NO: 4790), GSHSPH (SEQ ID NO: 4791), GDDSPH (SEQ ID NO: 4792), GHESPH (SEQ ID NO: 4793), GNYSPH (SEQ ID NO: 4794), RVGSPH (SEQ ID NO: 4795), GSCSPH (SEQ ID NO: 4796), GLLSPH (SEQ ID NO: 4797), MSGSPH (SEQ ID NO: 4798), RNGSPH (SEQ ID NO: 4799), TSGSPH (SEQ ID NO: 4800), ISGSPH (SEQ ID NO: 4801), GPGSPH (SEQ ID NO: 4802), ESGSPH (SEQ ID NO: 4803), SSGSPH (SEQ ID NO: 4804), GNGSPH (SEQ ID NO: 4805), ASGSPH (SEQ ID NO: 4806), NSGSPH (SEQ ID NO: 4807), LSGSPH (SEQ ID NO: 4808), GGGSPH (SEQ ID NO: 4809), KSGSPH (SEQ ID NO: 4810), HSGSPH (SEQ ID NO: 4811), GTGSPH (SEQ ID NO: 4812), PSGSPH (SEQ ID NO: 4813), GSVSPH (SEQ ID NO: 4814), RSGSPH (SEQ ID NO: 4815), GIGSPH (SEQ ID NO: 4816), WSGSPH (SEQ ID NO: 4817), DSGSPH (SEQ ID NO: 4818), IDGSPH (SEQ ID NO: 4819), GLGSPH (SEQ ID NO: 4820), DAGSPH (SEQ ID NO: 4821), DGGSPH (SEQ ID NO: 4822), MEGSPH (SEQ ID NO: 4823), ENGSPH (SEQ ID NO: 4824), GSASPH (SEQ ID NO: 4825), KNGSPH (SEQ ID NO: 4826), KEGSPH (SEQ ID NO: 4827), AIGSPH (SEQ ID NO: 4828), GYDSPH (SEQ ID NO: 4829), GHGSPH (SEQ ID NO: 4830), GRDSPH (SEQ ID NO: 4831), GNDSPH (SEQ ID NO: 4832), GPDSPH (SEQ ID NO: 4833), GMGSPH (SEQ ID NO: 4834), GQVSPH (SEQ ID NO: 4835), GHNSPH (SEQ ID NO: 4836), GHPSPH (SEQ ID NO: 4837), or GHSSPH (SEQ ID NO: 4838); and/or

(iii) [N1]-[N2]-[N3] comprises GHDSPHKSG (SEQ ID NO: 4698), GSGSPHSKA (SEQ ID NO: 4697), GSGSPHARM (SEQ ID NO: 4906), GSGSPHVKS (SEQ ID NO: 4907), GQDSPHKSG (SEQ ID NO: 4908), GSGSPHASR (SEQ ID NO: 4909), GSGSPHVKI (SEQ ID NO: 4910), GSGSPHKKN (SEQ ID NO: 4911), GSGSPHVRM (SEQ ID NO: 4912), VSGSPHSKA (SEQ ID NO: 4913), CSGSPHSKA (SEQ ID NO: 4914), GSGSPHRKA (SEQ ID NO: 4915), CSGSPHKTS (SEQ ID NO: 4916), CSHSPHKSG (SEQ ID NO: 4917), GQSSPHRSG (SEQ ID NO: 4918), GRGSPHASR (SEQ ID NO: 4919), GRGSPHSKA (SEQ ID NO: 4920), GSGSPHKFG (SEQ ID

NO: 4921), GSGSPHKIG (SEQ ID NO: 4922), GSGSPHKLG (SEQ ID NO: 4923), GSGSPHKTS (SEQ ID NO: 4924), GSGSPHKTT (SEQ ID NO: 4925), GSGSPHKTY (SEQ ID NO: 4926), GSGSPHKYG (SEQ ID NO: 4927), GSGSPHSKD (SEQ ID NO: 4928), GSGSPHSKP (SEQ ID NO: 4929), GSGSPHTRG (SEQ ID NO: 4930), GSGSPHVRG (SEQ ID NO: 4931), GSHSPHKRG (SEQ ID NO: 4932), GSHSPHKSG (SEQ ID NO: 4933), VSGSPHASR (SEQ ID NO: 4934), VSGSPHGAR (SEQ ID NO: 4935), VSGSPHKFG (SEQ ID NO: 4936), GHDSPHKRG (SEQ ID NO: 4937), GDDSPHKSG (SEQ ID NO: 4938), GHESPHKSA (SEQ ID NO: 4939), GHDSPHKSA (SEQ ID NO: 4940), GNYSPHKIG (SEQ ID NO: 4941), GHDSPHKSR (SEQ ID NO: 4942), GSGSPHSKL (SEQ ID NO: 4943), GSGSPHSRA (SEQ ID NO: 4944), GSGSPHSKR (SEQ ID NO: 4945), GSGSPHSLR (SEQ ID NO: 4946), GSGSPHSRG (SEQ ID NO: 4947), GSGSPHSSR (SEQ ID NO: 4948), RVGSPHSKA (SEQ ID NO: 4949), GSCSPHRKA (SEQ ID NO: 4950), GSGSPHFLR (SEQ ID NO: 4951), GSGSPHSKW (SEQ ID NO: 4952), GSGSPHSKS (SEQ ID NO: 4953), GLLSPHWKA (SEQ ID NO: 4954), GSGSPHVRR (SEQ ID NO: 4955), GSGSPHSKV (SEQ ID NO: 4956), MSGSPHSKA (SEQ ID NO: 4957), RNGSPHSKA (SEQ ID NO: 4958), TSGSPHSKA (SEQ ID NO: 4959), ISGSPHSKA (SEQ ID NO: 4960), GPGSPHSKA (SEQ ID NO: 4961), GSGSPHSKT (SEQ ID NO: 4962), ESGSPHSKA (SEQ ID NO: 4963), SSGSPHSKA (SEQ ID NO: 4964), GNGSPHSKA (SEQ ID NO: 4965), ASGSPHSKA (SEQ ID NO: 4966), NSGSPHSKA (SEQ ID NO: 4967), LSGSPHSKA (SEQ ID NO: 4968), GGGSPHSKA (SEQ ID NO: 4969), KSGSPHSKA (SEQ ID NO: 4970), GGGSPHSKS (SEQ ID NO: 4971), GSGSPHSKG (SEQ ID NO: 4972), HSGSPHSKA (SEQ ID NO: 4973), GTGSPHSKA (SEQ ID NO: 4974), PSGSPHSKA (SEQ ID NO: 4975), GSVSPHGKA (SEQ ID NO: 4976), RSGSPHSKA (SEQ ID NO: 4977), GSGSPHTKA (SEQ ID NO: 4978), GIGSPHSKA (SEQ ID NO: 4979), WSGSPHSKA (SEQ ID NO: 4980), DSGSPHSKA (SEQ ID NO: 4981), IDGSPHSKA (SEQ ID NO: 4982), GSGSPHNKA (SEQ ID NO: 4983), GLGSPHSKS (SEQ ID NO: 4984), DAGSPHSKA (SEQ ID NO: 4985), DGGSPHSKA (SEQ ID NO: 4986), MEGSPHSKA (SEQ ID NO: 4987), ENGSPHSKA (SEQ ID NO: 4988), GSASPHSKA (SEQ ID NO: 4989), GNGSPHSKS (SEQ ID NO: 4990), KNGSPHSKA (SEQ ID NO: 4991), KEGSPHSKA (SEQ ID NO: 4992), AIGSPHSKA (SEQ ID NO: 4993), GSGSPHSKN (SEQ ID NO: 4994), GSGSPHAKA (SEQ ID NO: 4995), GHDSPHKIG (SEQ ID NO: 4996), GYDSPHKSG (SEQ ID NO: 4997), GHESPHKSG (SEQ ID NO: 4998), GHDSPHKTG (SEQ ID NO: 4999), GRGSPHKRG (SEQ ID NO: 5000), GQDSPHKSG (SEQ ID NO: 4908), GHDSPHKSL (SEQ ID NO: 5001), GHGSPHSKA (SEQ ID NO: 5002), GHDSPHKSE (SEQ ID NO: 5003), VSGSPHSKA (SEQ ID NO: 4913), GRDSPHKSG (SEQ ID NO: 5004), GNDSPHKSV (SEQ ID NO: 5005), GQDSPHKIG (SEQ ID NO: 5006), GHDSPHKSV (SEQ ID NO: 5007), GPDSPHKIG (SEQ ID NO: 5008), GPDSPHKSG (SEQ ID NO: 5009), GHDSPHKSW (SEQ ID NO: 5010), GHDSPHKSN (SEQ ID NO: 5011), GMGSPHSKT (SEQ ID NO: 5012), GHDSPHKHG (SEQ ID NO: 5013), GQVSPHKSG (SEQ ID NO: 5014), GDDSPHKSV (SEQ ID NO: 5015), GHNSPHKSG (SEQ ID NO: 5016), GNGSPHKRG (SEQ ID NO: 5017), GHDSPHKYG

(SEQ ID NO: 5018), GHDSPHKSQ (SEQ ID NO: 5019), GNDSPHKIG (SEQ ID NO: 5020), GHDSPHKSK (SEQ ID NO: 5021), GHDSPHKLW (SEQ ID NO: 5022), GHPSPHWKG (SEQ ID NO: 5023), GHDSPHKMG (SEQ ID NO: 5024), GHDSPHKMA (SEQ ID NO: 5025), or GHSSPHRSG (SEQ ID NO: 5026).

15. The composition of any one of claims 10-14, wherein:

(a) [N1]-[N2]-[N3] comprises GHDSPHKSG (SEQ ID NO: 4698); or

(b) [N1]-[N2]-[N3] comprises GSGSPHSKA (SEQ ID NO: 4697).

16. The composition of any one of claims 10-15, which further comprises:

(i) [N4], wherein [N4] comprises QNQQ (SEQ ID NO: 5028), WNQQ (SEQ ID NO: 5029), QYYV (SEQ ID NO: 5030), RRQQ (SEQ ID NO: 5031), GCGQ (SEQ ID NO: 5032), LRQQ (SEQ ID NO: 5033), RNQQ (SEQ ID NO: 5034), VNQQ (SEQ ID NO: 5035), FRLQ (SEQ ID NO: 5036), FNQQ (SEQ ID NO: 5037), LLQQ (SEQ ID NO: 5038), SNQQ (SEQ ID NO: 5039), RLQQ (SEQ ID NO: 5040), LNQQ (SEQ ID NO: 5041), QRKL (SEQ ID NO: 5042), LRRQ (SEQ ID NO: 5043), QRLR (SEQ ID NO: 5044), QRRL (SEQ ID NO: 5045), RRLQ (SEQ ID NO: 5046), RLRQ (SEQ ID NO: 5047), SKRQ (SEQ ID NO: 5048), QLYR (SEQ ID NO: 5049), QLTV (SEQ ID NO: 5050), QNKQ (SEQ ID NO: 5051), KNQQ (SEQ ID NO: 5052), QKQQ (SEQ ID NO: 5053), QTQQ (SEQ ID NO: 5054), QNHQ (SEQ ID NO: 5055), QHQQ (SEQ ID NO: 5056), QNQH (SEQ ID NO: 5057), QHRQ (SEQ ID NO: 5058), LTQQ (SEQ ID NO: 5059), QNQW (SEQ ID NO: 5060), QNTH (SEQ ID NO: 5061), RRRQ (SEQ ID NO: 5062), QYQQ (SEQ ID NO: 5063), QNDQ (SEQ ID NO: 5064), QNRH (SEQ ID NO: 5065), RDQQ (SEQ ID NO: 5066), PNLQ (SEQ ID NO: 5067), HVRQ (SEQ ID NO: 5068), PNQH (SEQ ID NO: 5069), HNQQ (SEQ ID NO: 5070), QSQQ (SEQ ID NO: 5071), QPAK (SEQ ID NO: 5072), QNLA (SEQ ID NO: 5073), QNQL (SEQ ID NO: 5074), QGQQ (SEQ ID NO: 5075), LNRQ (SEQ ID NO: 5076), QNPP (SEQ ID NO: 5077), QNLQ (SEQ ID NO: 5078), QDQE (SEQ ID NO: 5079), QDQQ (SEQ ID NO: 5080), HWQQ (SEQ ID NO: 5081), PNQQ (SEQ ID NO: 5082), PEQQ (SEQ ID NO: 5083), QRTM (SEQ ID NO: 5084), LHQH (SEQ ID NO: 5085), QHRI (SEQ ID NO: 5086), QYIH (SEQ ID NO: 5087), QKFE (SEQ ID NO: 5088), QFPS (SEQ ID NO: 5089), QNPL (SEQ ID NO: 5090), QAIK (SEQ ID NO: 5091), QNRQ (SEQ ID NO: 5092), QYQH (SEQ ID NO: 5093), QNPQ (SEQ ID NO: 5094), QHQL (SEQ ID NO: 5095), QSPP (SEQ ID NO: 5096), QAKL (SEQ ID NO: 5097), KSQQ (SEQ ID NO: 5098), QDRP (SEQ ID NO: 5099), QNLG (SEQ ID NO: 5100), QAFH (SEQ ID NO: 5101), QNAQ (SEQ ID NO: 5102), HNQL (SEQ ID NO: 5103), QKLN (SEQ ID NO: 5104), QNVQ (SEQ ID NO: 5105), QAQQ (SEQ ID NO: 5106), QTPP (SEQ ID NO: 5107), QPPA (SEQ ID NO: 5108), QERP (SEQ ID NO: 5109), QDLQ (SEQ ID NO: 5110), QAMH (SEQ ID NO: 5111), QHPS (SEQ ID NO: 5112), PGLQ (SEQ ID NO: 5113), QGIR (SEQ ID NO: 5114), QAPA (SEQ ID NO: 5115), QIPP (SEQ ID NO: 5116), QTQL (SEQ ID NO: 5117), QAPS (SEQ ID NO: 5118), QNTY (SEQ ID NO: 5119), QDKQ (SEQ ID NO: 5120),

QNHL (SEQ ID NO: 5121), QIGM (SEQ ID NO: 5122), LNKQ (SEQ ID NO: 5123), PNQL (SEQ ID NO: 5124), QLQQ (SEQ ID NO: 5125), QRMS (SEQ ID NO: 5126), QGIL (SEQ ID NO: 5127), QDRQ (SEQ ID NO: 5128), RDWQ (SEQ ID NO: 5129), QERS (SEQ ID NO: 5130), QNYQ (SEQ ID NO: 5131), QRTC (SEQ ID NO: 5132), QIGH (SEQ ID NO: 5133), QGAI (SEQ ID NO: 5134), QVPP (SEQ ID NO: 5135), QVQQ (SEQ ID NO: 5136), LMRQ (SEQ ID NO: 5137), QYSV (SEQ ID NO: 5138), QAIT (SEQ ID NO: 5139), QKTL (SEQ ID NO: 5140), QLHH (SEQ ID NO: 5141), QNII (SEQ ID NO: 5142), QGHH (SEQ ID NO: 5143), QSKV (SEQ ID NO: 5144), QLPS (SEQ ID NO: 5145), IGKQ (SEQ ID NO: 5146), QAIH (SEQ ID NO: 5147), QHGL (SEQ ID NO: 5148), QFMC (SEQ ID NO: 5149), QNQM (SEQ ID NO: 5150), QHLQ (SEQ ID NO: 5151), QPAR (SEQ ID NO: 5152), QSLQ (SEQ ID NO: 5153), QSQL (SEQ ID NO: 5154), HSQQ (SEQ ID NO: 5155), QMPS (SEQ ID NO: 5156), QGSL (SEQ ID NO: 5157), QVPA (SEQ ID NO: 5158), HYQQ (SEQ ID NO: 5159), QVPS (SEQ ID NO: 5160), RGEQ (SEQ ID NO: 5161), PGQQ (SEQ ID NO: 5162), LEQQ (SEQ ID NO: 5163), QNQS (SEQ ID NO: 5164), QKVI (SEQ ID NO: 5165), QNND (SEQ ID NO: 5166), QSVH (SEQ ID NO: 5167), QPLG (SEQ ID NO: 5168), HNQE (SEQ ID NO: 5169), QIQQ (SEQ ID NO: 5170), QVRN (SEQ ID NO: 5171), PSNQ (SEQ ID NO: 5172), QVGH (SEQ ID NO: 5173), QRDI (SEQ ID NO: 5174), QMPN (SEQ ID NO: 5175), RGLQ (SEQ ID NO: 5176), PSLQ (SEQ ID NO: 5177), QRDQ (SEQ ID NO: 5178), QAKG (SEQ ID NO: 5179), QSAH (SEQ ID NO: 5180), QSTM (SEQ ID NO: 5181), QREM (SEQ ID NO: 5182), QYRA (SEQ ID NO: 5183), QRQQ (SEQ ID NO: 5184), QWQQ (SEQ ID NO: 5185), QRMN (SEQ ID NO: 5186), GDSQ (SEQ ID NO: 5187), QKIS (SEQ ID NO: 5188), PSMQ (SEQ ID NO: 5189), SPRQ (SEQ ID NO: 5190), MEQQ (SEQ ID NO: 5191), QYQN (SEQ ID NO: 5192), QIRQ (SEQ ID NO: 5193), QSVQ (SEQ ID NO: 5194), RSQQ (SEQ ID NO: 5195), QNKL (SEQ ID NO: 5196), QIQH (SEQ ID NO: 5197), PRQQ (SEQ ID NO: 5198), HTQQ (SEQ ID NO: 5199), QRQH (SEQ ID NO: 5200), RNQE (SEQ ID NO: 5201), QSKQ (SEQ ID NO: 5202), QNQP (SEQ ID NO: 5203), QSPQ (SEQ ID NO: 5204), QTRQ (SEQ ID NO: 5205), QNLH (SEQ ID NO: 5206), QNQE (SEQ ID NO: 5207), LNQP (SEQ ID NO: 5208), QNQD (SEQ ID NO: 5209), QNLL (SEQ ID NO: 5210), QLVI (SEQ ID NO: 5211), RTQE (SEQ ID NO: 5212), QTHQ (SEQ ID NO: 5213), QDQH (SEQ ID NO: 5214), QSQH (SEQ ID NO: 5215), VRQQ (SEQ ID NO: 5216), AWQQ (SEQ ID NO: 5217), QSVP (SEQ ID NO: 5218), QNIQ (SEQ ID NO: 5219), LDQQ (SEQ ID NO: 5220), PDQQ (SEQ ID NO: 5221), ESQQ (SEQ ID NO: 5222), QRQL (SEQ ID NO: 5223), QIIV (SEQ ID NO: 5224), QKQS (SEQ ID NO: 5225), QSHQ (SEQ ID NO: 5226), QFVV (SEQ ID NO: 5227), QSQP (SEQ ID NO: 5228), QNEQ (SEQ ID NO: 5229), INQQ (SEQ ID NO: 5230), RNRQ (SEQ ID NO: 5231), RDQK (SEQ ID NO: 5232), QWKR (SEQ ID NO: 5233), ENRQ (SEQ ID NO: 5234), QTQP (SEQ ID NO: 5235), QKQL (SEQ ID NO: 5236), RNQL (SEQ ID NO: 5237), ISIQ (SEQ ID NO: 5238), QTVC (SEQ ID NO: 5239), QQIM (SEQ ID NO: 5240), LNHQ (SEQ ID NO: 5241), QNQA (SEQ ID NO: 5242), QMIH (SEQ ID NO: 5243), RNHQ (SEQ ID NO: 5244), or QKMN (SEQ ID NO: 5245); and/or

(ii) [N0], wherein [N0] comprises TIN, SMN, TIM, YLS, GLS, MPE, MEG, MEY, AEW, CEW, ANN, IPE, ADM, IEY, ADY, IET, MEW, CEY, RIN, MEI, LEY, ADW, IEI, DIM, FEQ, MEF, CDQ, LPE, IEN, MES, AEI, VEY, IIN, TSN, IEV, MEM, AEV, MDA, VEW, AEQ, LEW, MEL, MET, MEA, IES, MEV, CEI, ATN, MDG, QEV, ADQ, NMN, IEM, ISN, TGN, QQQ, HDW, IEG, TII, TFP, TEK, EIN, TVN, TFN, SIN, TER, TSY, ELH, AIN, SVN, TDN, TFH, TVH, TEN, TSS, TID, TCN, NIN, TEH, AEM, AIK, TDK, TFK, SDQ, TEI, NTN, TET, SIK, TEL, TEA, TAN, TIY, TFS, TES, TTN, TED, TNN, EVH, TIS, TVR, TDR, TIK, NHI, TIP, ESD, TDL, TVP, TVI, AEH, NCL, TVK, NAD, TIT, NCV, TIR, NAL, VIN, TIQ, TEF, TRE, QGE, SEK, NVN, GGE, EFV, SDK, TEQ, EVQ, TEY, NCW, TDV, SDI, NSI, NSL, EVV, TEP, SEL, TWQ, TEV, AVN, GVL, TLN, TEG, TRD, NAI, AEN, AET, ETA, NNL.

17. The composition of claims 16, wherein [N0]-[N1]-[N2]-[N3]-[N4] comprises the amino acid sequence of any one of SEQ ID NOs: 2243, 2242, or 2242-2886.

18. The composition of claim 16 or 17, which comprises from N-terminus to C-terminus, [N0]-[N1]-[N2]-[N3]-[N4].

19. The composition of any one of claims 2-18, wherein:

(i) the peptide comprises the amino acid sequence of SPH, wherein the S comprises a modification, e.g., comprises a phosphate group;

(ii) the peptide comprises the amino acid sequence of SPHSKA (SEQ ID NO: 941), wherein the S at position 1, numbered according to SEQ ID NO: 941, comprises a modification, e.g., comprises a phosphate group;

(iii) the peptide comprises the amino acid sequence of SPHK (SEQ ID NO: 6398), wherein the S comprises a modification, e.g., comprises a phosphate group; or

(iv) the peptide comprises the amino acid sequence of HDSPHK (SEQ ID NO: 2), wherein the S comprises a modification, e.g., comprises a phosphate group.

20. The composition of any one of claims 2-20, wherein the peptide comprises the amino acid sequence of:

(i) GHDSPHKS (SEQ ID NO: 4487), optionally wherein the S at position 4 of SEQ ID NO: 4487, comprises a modification, e.g., comprises a phosphate group;

(ii) NGHDSPHKSG (SEQ ID NO: 4489), optionally wherein the S at position 5 of SEQ ID NO: 4489, comprises a modification, e.g., comprises a phosphate group;

(iii) INGHDSPHKSGQ (SEQ ID NO: 4490), optionally wherein the S at position 6 of SEQ ID NO: 4490, comprises a modification, e.g., comprises a phosphate group;

(iv) TINGHDSPHKSGQN (SEQ ID NO: 4491), optionally wherein the S at position 7 of SEQ ID NO: 4491, comprises a modification, e.g., comprises a phosphate group;

(v) KTINGHDSPHKSGQNQ (SEQ ID NO: 4492), optionally wherein the S at position 8 of SEQ ID NO: 4492, comprises a modification, e.g., comprises a phosphate group;

(vi) LYYLSKTINGHDSPHKSGQNQQTLKF (SEQ ID NO: 4518), optionally wherein the S at position 13 of SEQ ID NO: 4518, comprises a modification, e.g., comprises a phosphate group;

(vii) RLMNPLIDQYLYYLSKTINGHDSPHKSGQNQQTLKFSVAGPSNMAV (SEQ ID NO: 4519), optionally wherein the S at position 23 of SEQ ID NO: 4519, comprises a modification, e.g., comprises a phosphate group;

(viii) GSPHSKAQ (SEQ ID NO: 4493), optionally wherein the S at position 2 of SEQ ID NO: 4493, comprises a modification, e.g., comprises a phosphate group;

(ix) SGSPHSKAQN (SEQ ID NO: 4494), optionally wherein the S at position 3 of SEQ ID NO: 4494, comprises a modification, e.g., comprises a phosphate group;

(x) GSGSPHSKAQNQ (SEQ ID NO: 4495), optionally wherein the S at position 4 of SEQ ID NO: 4495, comprises a modification, e.g., comprises a phosphate group;

(xi) NGSGSPHSKAQNQQ (SEQ ID NO: 4496), optionally wherein the S at position 5 of SEQ ID NO: 4496, comprises a modification, e.g., comprises a phosphate group; or

(xii) INGSGSPHSKAQNQQT (SEQ ID NO: 4497), optionally wherein the S at position 6 of SEQ ID NO: 4497, comprises a modification, e.g., comprises a phosphate group.

21. The composition of any one of claims 9-20, wherein the modification comprises a phosphate group.

22. The composition of any one of claims 2-21, wherein the peptide comprises the amino acid sequence of NGHDpSPHKSG (SEQ ID NO: 4515); KTINGHDpSPHKSGQNQ (SEQ ID NO: 4516); or YLSKTINGHDpSPHKSGQNQQTLKFS (SEQ ID NO: 4517).

23. The composition of any one of claims 1-9, wherein the ligand is or comprises an antibody molecule, optionally wherein the variable domain of the antibody molecule binds to ALPL, e.g., human ALPL.

24. The composition of any one of claims 3-9 or 23, wherein the antibody molecule:

(i) is an antibody as provided in Table 40 (e.g. Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, 2F4, or a variant thereof;

(ii) binds the same or substantially the same epitope as any one of an antibody as provided in Table 40 (e.g., Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, 2F4, or a variant thereof; and/or

(iii) competes for binding with any one of an antibody as provided in Table 40 (e.g., Ab 9), AF2910-SP, AF2909, NBP2-67295, LS-B3666, MA524845, 2F4, or a variant thereof.

25. The composition of any one of claims 1-3, 4-6, 23, or 24, wherein the ligand is a first variable domain of a multispecific antibody molecule and the active agent is a second variable of the multispecific antibody molecule, optionally wherein the second variable domain binds a therapeutic target comprising:

(i) a CNS related target, e.g., an antigen associated with a neurological or neurodegenerative disorder, e.g., β-amyloid or tau;

(ii) a muscular or neuromuscular related target, e.g., an antigen associated with a muscular or neuromuscular disorder; or

(iii) a neuro-oncology related target, e.g., an antigen associated with a neuro-oncological disorder, e.g., HER2, or EGFR (e.g., EGFRvIII).

26. The composition of any one of claims 1-9, wherein the ligand is a small molecule, wherein the small molecule is an inhibitor of ALPL.

27. The composition of any one of claims 1-9 or 26, wherein the small molecule is:

(i) an aryl sulfonamide, a phosphonate derivative, a pyrazole, a triazole, or an imidazole; or

(ii) 2,5-Dimethoxy-N-(quinolin-3-yl)benzenesulfonamide (Tissue-Nonspecific Alkaline Phosphatase Inhibitor (TNAPi)) or 5-((5-chloro-2-methoxyphenyl)sulfonamido)nicotinamide (SBI-425).

28. The composition of any one of claims 1-27, wherein the ligand is present or coupled to a carrier, e.g., an exosome, a microvesicle, or a lipid nanoparticle (LNP), optionally wherein the carrier comprises the active agent, e.g., a therapeutic agent.

29. The composition of any one of claims 1-28, wherein the active agent comprises a therapeutic agent chosen from a protein (e.g., an enzyme), an antibody molecule, a nucleic acid molecule (e.g., an RNAi agent), or a small molecule.

30. The composition of any one of claims 3-9, 23, 24, or 29, wherein the antibody molecule comprises a full antibody or an antigen binding fragment, optionally wherein the antigen binding fragment is a Fab or a Fab fragment, a F(ab)2 fragment, an Fv fragment, dAb fragment, a single chain antibody (scFv) or a scFv fragment, an antibody variable region, a diabody, a VHH, a camelid antibody, a single domain antibody or a nanobody.

31. The composition of any one of claims 3-9, 23, 24, 29, or 30, wherein the antibody molecule is a monospecific antibody, a multispecific antibody, e.g., a bispecific or biparatopic antibody.

32. The composition of any one of claims 28-31, wherein the therapeutic agent is an antibody molecule that binds:

(i) a CNS related target, e.g., an antigen associated with a neurological or neurodegenerative disorder, e.g., β-amyloid or tau;

(ii) a muscular or neuromuscular related target, e.g., an antigen associated with a muscular or neuromuscular disorder; or

(iii) a neuro-oncology related target, e.g., an antigen associated with a neuro-oncological disorder, e.g., HER2, or EGFR (e.g., EGFRvIII).

33. The composition of claim 28 or 29, wherein the therapeutic agent is an RNAi agent.

34. The composition of claim 33, wherein the RNAi agent is a dsRNA, a siRNA, a shRNA, a pre-miRNA, a pri-miRNA, a miRNA, a stRNA, a lncRNA, a piRNA, an antisense oligonucleotide agent (ASO), or a snoRNA (e.g., an siRNA or an ASO).

35. The composition of claim 33 or 34, wherein ligand is conjugated to the RNAi agent via a crosslinker, optionally wherein the crosslinker comprises succinimidyl-4-(N-maleimidomethyl) and/or a saturated or unsaturated hydrocarbon chain (e.g., cyclohexane-1-carboxylate).

36. The composition of claim 28 or 29, wherein the therapeutic agent is a therapeutic protein or functional variant thereof, wherein the therapeutic protein or variant thereof is associated with (e.g., aberrantly expressed in) a neurological or neurodegenerative disorder, a muscular or neuromuscular disorder, or a neuro-oncological disorder.

37. The composition of claim 36, wherein the therapeutic protein or functional variant thereof is chosen from apolipoprotein E (APOE) (e.g., ApoE2, ApoE3 and/or ApoE4); human survival of motor neuron (SMN) 1 or SMN2; glucocerebrosidase (GBA1); aromatic L-amino acid decarboxylase (AADC); aspartoacylase (ASPA); tripeptidyl peptidase I (CLN2); beta-galactosidase (GLB1); N-sulphoglucosamine sulphohydrolase (SGSH); N-acetyl-alpha-glucosaminidase (NAGLU); iduronate 2-sulfatase (IDS); intracellular cholesterol transporter (NPC1); or gigaxonin (GAN).

38. The composition of any one of claims 1-28, wherein the active agent is a diagnostic agent, optionally wherein the diagnostic agent is or comprises an imaging agent (e.g., a protein or small molecule compound coupled to a detectable moiety).

39. A cell comprising the composition of any one of claims 1-38, optionally wherein the cell is a mammalian cell, a cell of the central nervous system, or and/or a cell present in the blood brain barrier.

40. A method of making the composition of any one of claims 1-38, comprising:

(i) providing the ligand that binds to the GPI anchored protein, e.g., ALPL, and the active agent; and

(ii) incubating the ligand and active agent under conditions suitable to fuse or couple the ligand to the active agent,

thereby generating the composition.

41. A pharmaceutical composition comprising the composition of any one of claims 1-38, and a pharmaceutically acceptable excipient.

42. A method of delivering an active agent, e.g., a therapeutic agent or a diagnostic agent, to a cell or tissue (e.g., a CNS cell or a CNS tissue), comprising administering the composition of any one of claims 1-38 or the pharmaceutical composition of claim 41.

43. The method of claim 42, wherein the cell is:

(i) a cell of a brain region or a spinal cord region, optionally a cell of the frontal cortex, sensory cortex, motor cortex, caudate, cerebellar cortex, cerebral cortex, brain stem, hippocampus, or thalamus; and/or

(ii) in a subject.

44. A method of increasing central nervous system transduction (e.g., increased crossing of the blood brain barrier) in a subject, comprising administering the composition of any one of embodiments 1-38, or the pharmaceutical composition of claim 41.

45. The method of claim 43 or 44, wherein the subject has, has been diagnosed with having, or is at risk of having a genetic disorder (e.g., a monogenic disorder or a polygenic disorder), a neurological disorder, a neurodegenerative disorder, a neuro-oncological disorder, a muscular disorder, or a neuromuscular disorder.

46. A method of treating a subject having or diagnosed with having a genetic disorder (e.g., a monogenic disorder or a polygenic disorder), a neurological disorder, a neurodegenerative disorder, a neuro-oncological disorder, a muscular disorder, or a neuromuscular disorder, comprising

administering the composition of any one of embodiments 1-38, or the pharmaceutical composition of claim 41.

47. The method of claim 46 or 47, wherein the genetic disorder, neurological disorder, neurodegenerative disorder, muscular disorder, neuromuscular disorder, or neuro-oncological disorder is Huntington's Disease, Amyotrophic Lateral Sclerosis (ALS), Gaucher Disease, Dementia with Lewy Bodies, Parkinson's disease, Spinal Muscular Atrophy, Alzheimer's Disease, a leukodystrophy (e.g., Alexander disease, autosomal dominant leukodystrophy with autonomic diseases (ADLD), Canavan disease, cerebrotendinous xanthomatosis (CTX), metachromatic leukodystrophy (MLD), Pelizaeus-Merzbacher disease, or Refsum disease), or a cancer (e.g., a HER2/neu positive cancer or a glioblastoma).

48. The method of any one of claims 42-47, wherein the composition or the pharmaceutical composition is administered to the subject intravenously, via intra-cisterna magna injection (ICM), intracerebrally, intrathecally, intracerebroventricularly, via intraparenchymal administration, intraarterially, or intramuscularly.

49. The composition of any one of claims 1-38 or the pharmaceutical composition of claim 41 for use in method of delivering a payload to a cell or tissue.

50. The composition of any one of claims 1-38 or the pharmaceutical composition of claim 41 for use in method of treating a genetic disorder (e.g., a monogenic disorder or a polygenic disorder), a neurological disorder, a neurodegenerative disorder, a neuro-oncological disorder, a muscular disorder, or a neuromuscular disorder.

51. Use of the composition of any one of claims 1-38 or the pharmaceutical composition of claim 41, in the manufacture of a medicament.

52. Use of the composition of any one of claims 1-38 or the pharmaceutical composition of claim 40, in the manufacture of a medicament for increasing CNS transduction (e.g., increased crossing of the blood brain barrier); and/or treating a genetic disorder (e.g., a monogenic disorder or a polygenic disorder), a neurological disorder, a neurodegenerative disorder, a neuro-oncological disorder, a muscular disorder, or a neuromuscular disorder.

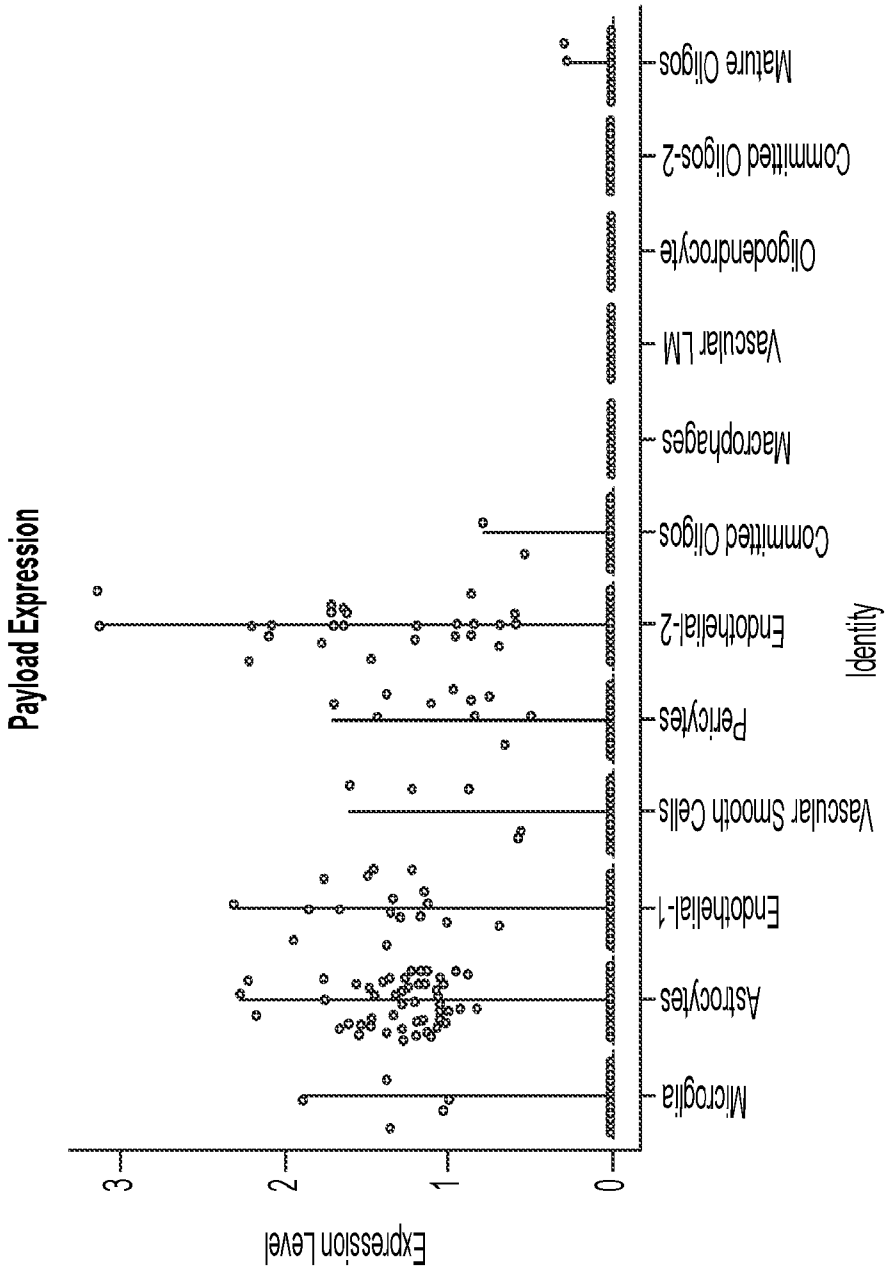

Payload Expression
Expression Level
0
1
2
3
Microglia
Astrocytes
Endothelial-1
Vascular Smooth Cells
Pericytes
Endothelial-2
Committed Oligos
Macrophages
Vascular LM
Oligodendrocyte
Committed Oligos-2
Mature Oligos
Identity

FIG. 1A

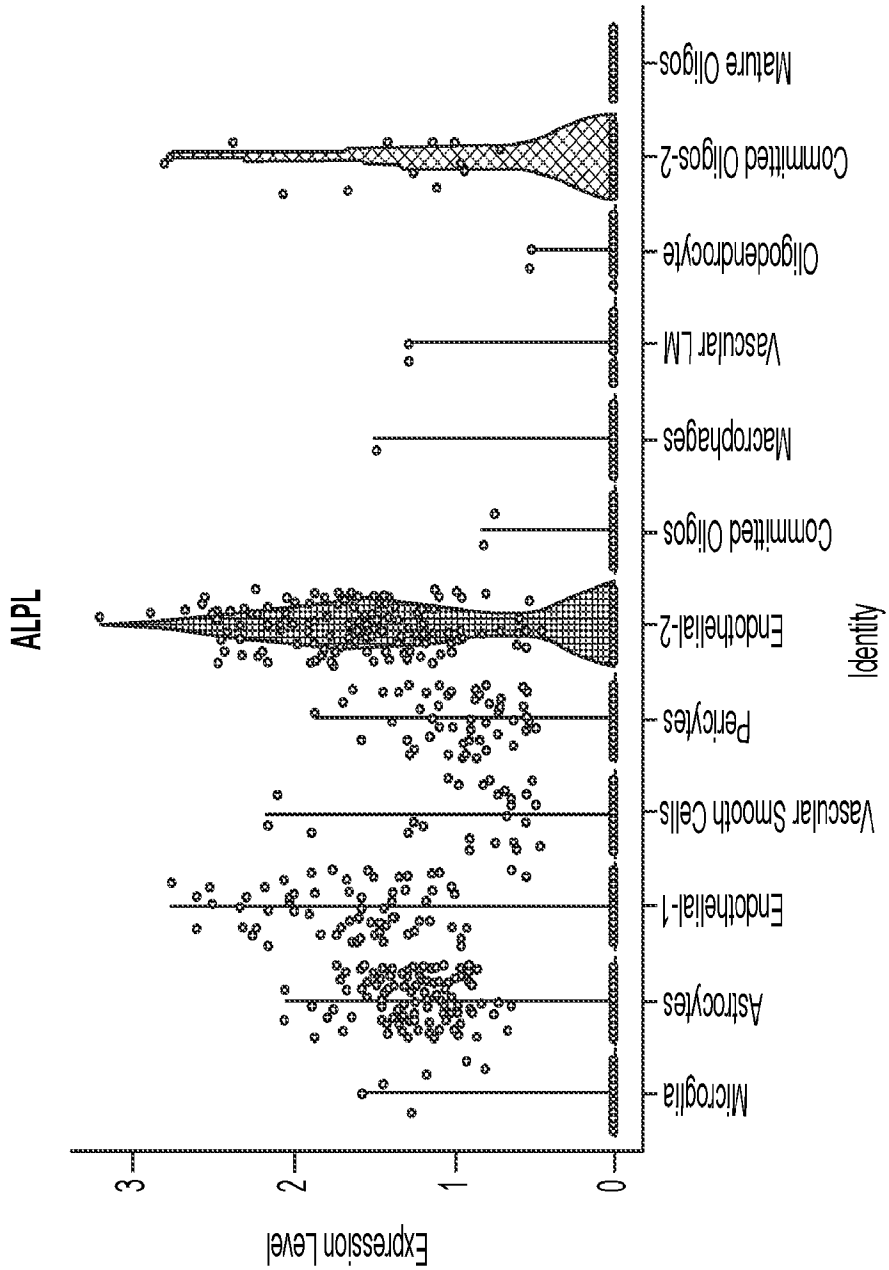

ALPL
Expression Level
0
1
2
3
Identity
Microglia
Astrocytes
Endothelial-1
Vascular Smooth Cells
Pericytes
Endothelial-2
Committed Oligos
Macrophages
Vascular LM
Oligodendrocyte
Committed Oligos-2
Mature Oligos

FIG. 1B

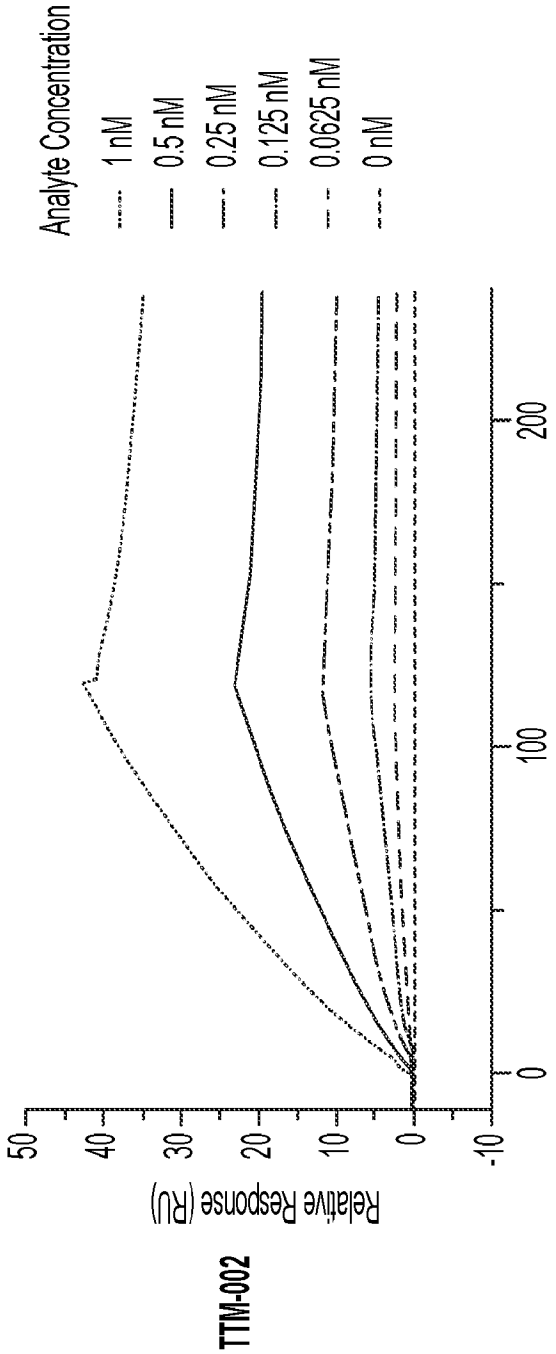

Analyte Concentration
1 nM
0.5 nM
0.25 nM
0.125 nM
0.0625 nM
0 nM
Relative Response (RU)
50
40
30
20
10
0
-10
0
100
200
Time (s)
TTM-002

FIG. 2A

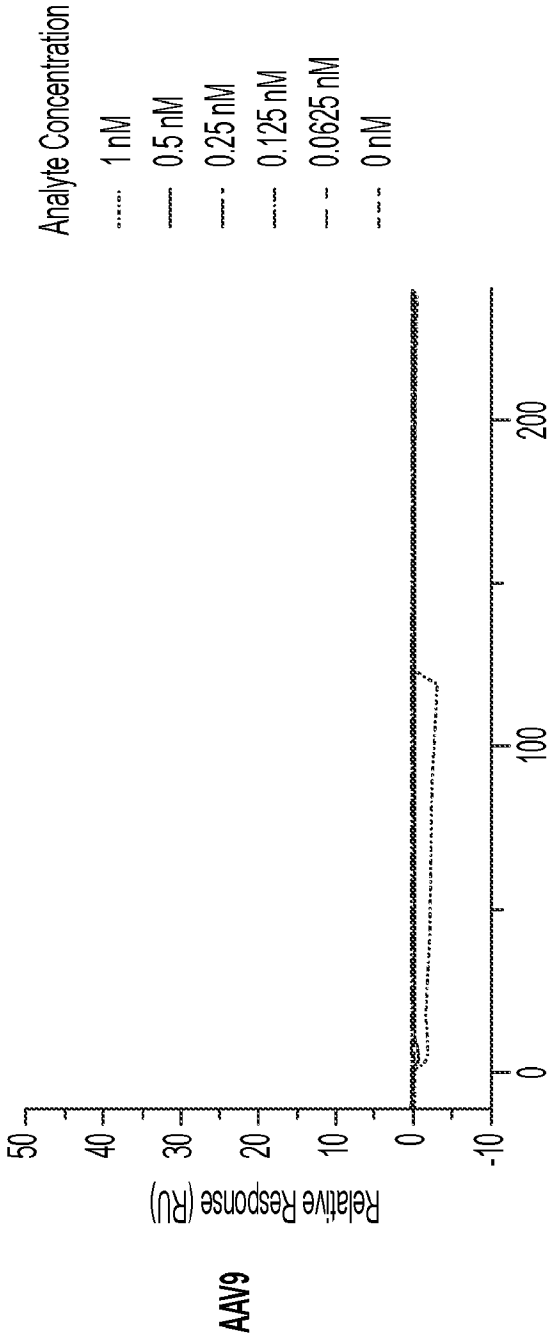

AAV9
Relative Response (RU)
50
40
30
20
10
0
-10
0
100
200
Time (s)
Analyte Concentration
1 nM
0.5 nM
0.25 nM
0.125 nM
0.0625 nM
0 nM

FIG. 2B

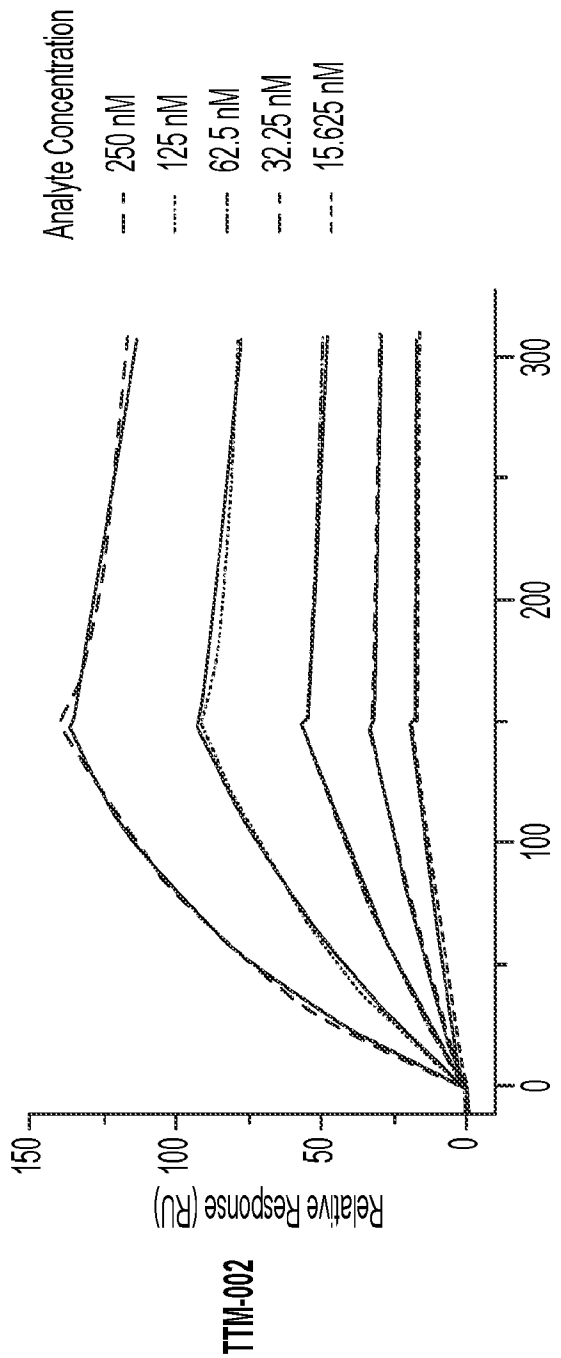

Analyte Concentration
250 nM
125 nM
62.5 nM
32.25 nM
15.625 nM
Relative Response (RU)
150
100
50
0
0
100
200
300
Time (s)
TTM-002

FIG. 2C

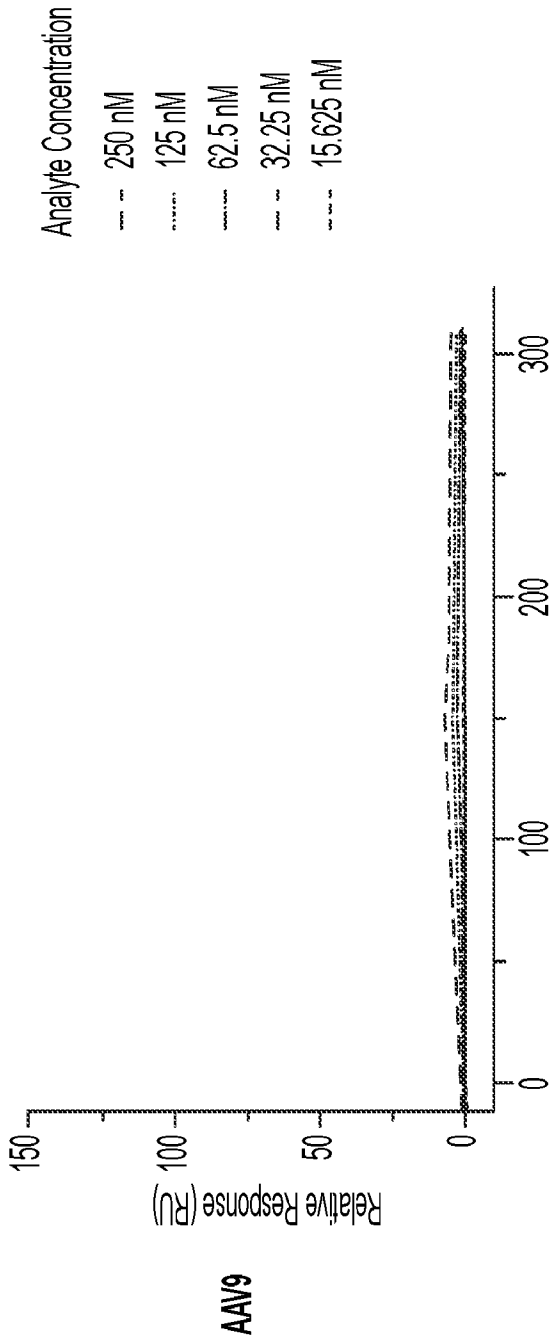

Analyte Concentration
250 nM
125 nM
62.5 nM
32.25 nM
15.625 nM
Relative Response (RU)
150
100
50
0
0
100
200
300
Time (s)
AAV9

FIG. 2D

FIG. 3A

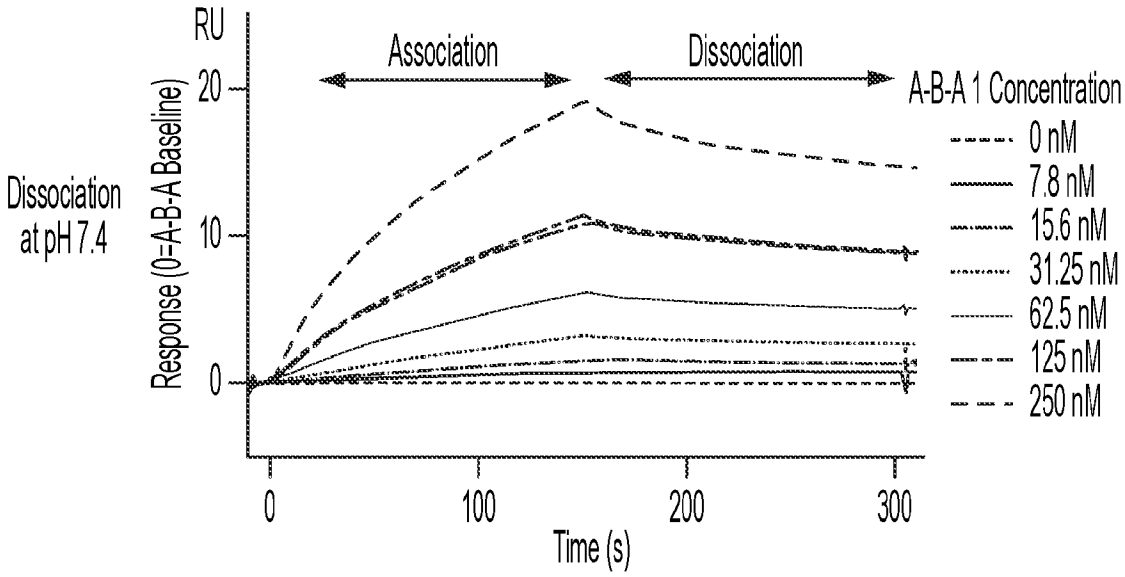
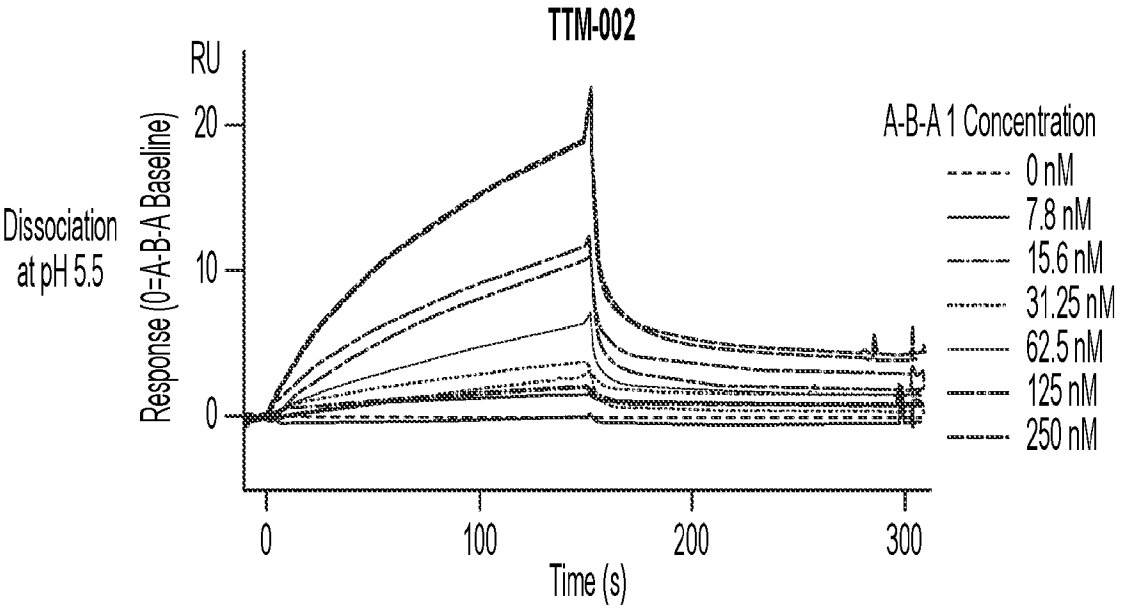

TTM-002
RU
Association
Dissociation
A-B-A 1 Concentration
0 nM
7.8 nM
15.6 nM
31.25 nM
62.5 nM
125 nM
250 nM
Dissociation at pH 7.4
Response (0=A-B-A Baseline)
20
10
0
0
100
200
300
Time (s)

FIG. 3B


TTM-002
RU
A-B-A 1 Concentration
0 nM
7.8 nM
15.6 nM
31.25 nM
62.5 nM
125 nM
250 nM
Dissociation at pH 5.5
Response (0=A-B-A Baseline)
20
10
0
0
100
200
300
Time (s)

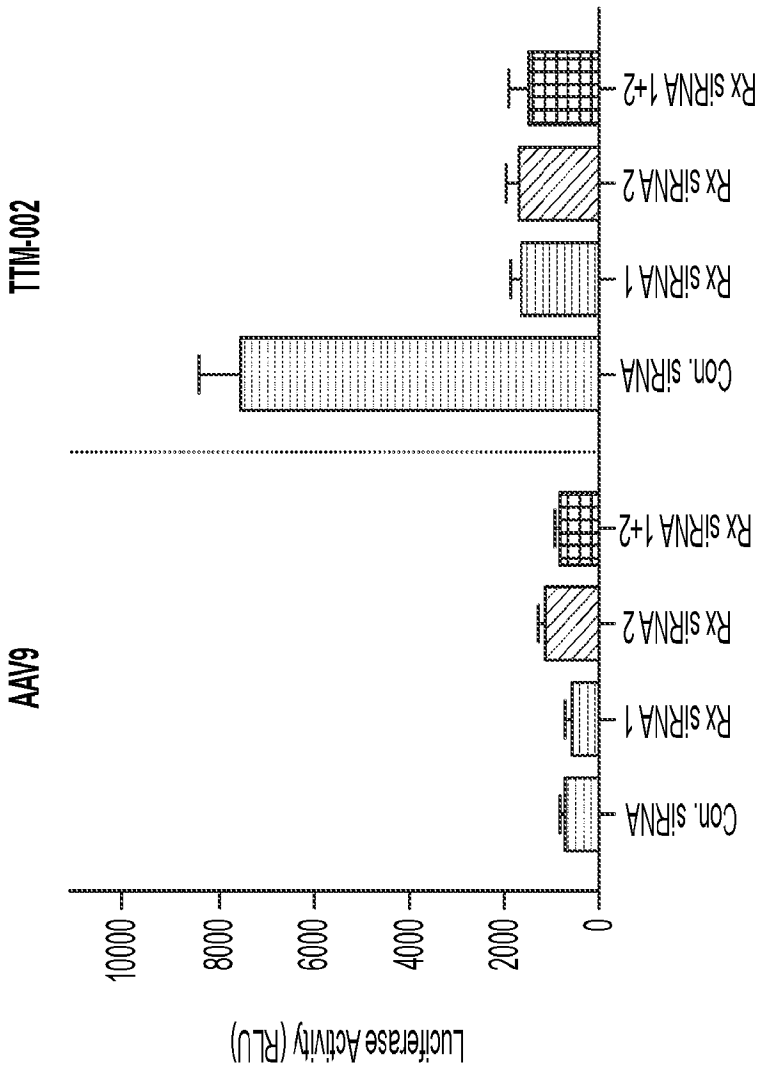

Luciferase Activity (RLU)
10000
8000
6000
4000
2000
0
AAV9
TTM-002
Con. siRNA
Rx siRNA 1
Rx siRNA 2
Rx siRNA 1+2
Con. siRNA
Rx siRNA 1
Rx siRNA 2
Rx siRNA 1+2

FIG. 4

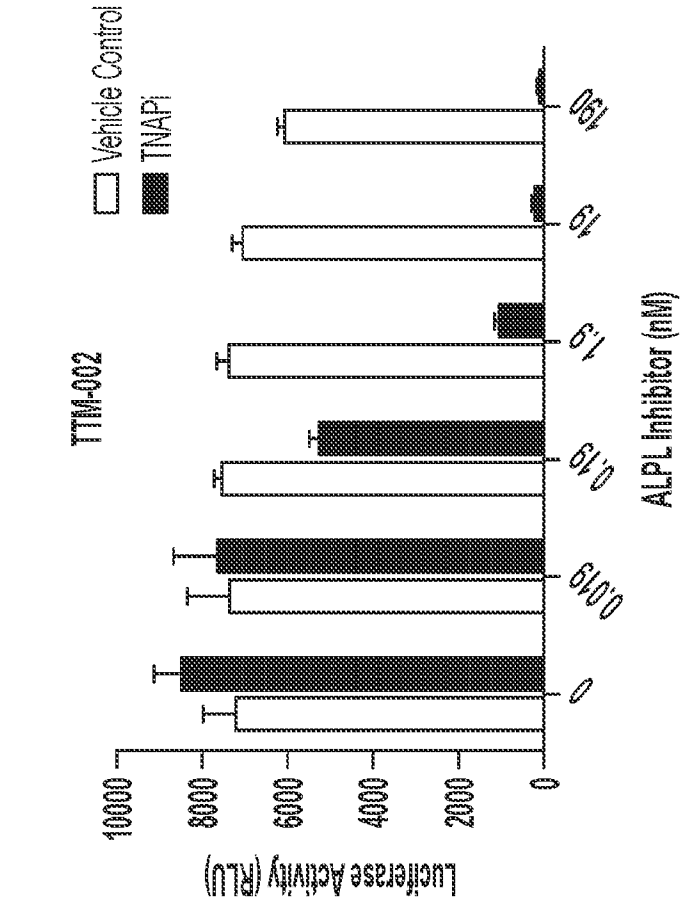

Vehicle Control
TNAPi
TTM-002
Luciferase Activity (RLU)
10000
8000
6000
4000
2000
0
0
0.0019
0.019
0.19
1.9
19
190
ALPL Inhibitor (nM)

FIG. 5B

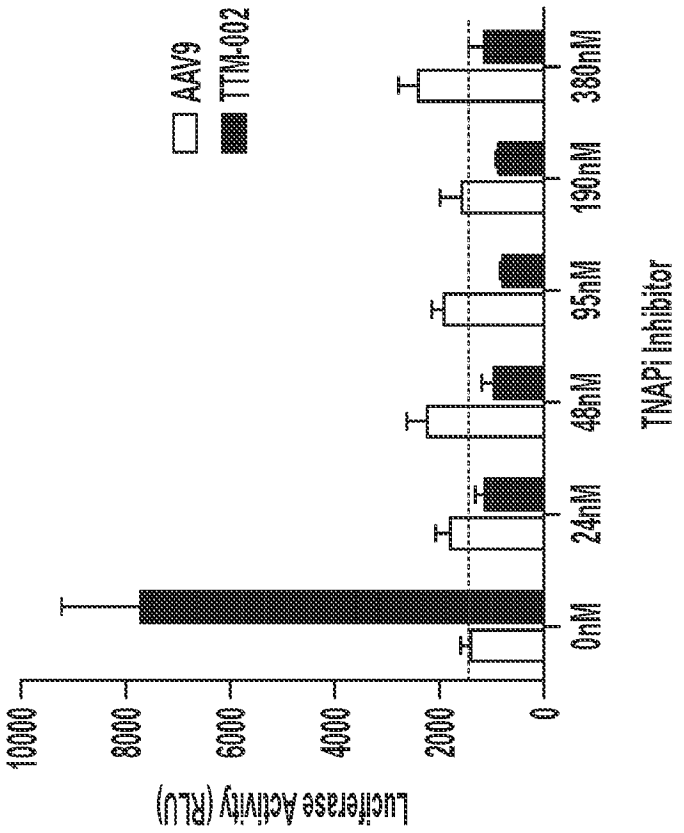

AAV9
TTM-002
Luciferase Activity (RLU)
10000
8000
6000
4000
2000
0
0nM
24nM
48nM
95nM
190nM
380nM
TNAPi Inhibitor

FIG. 5A

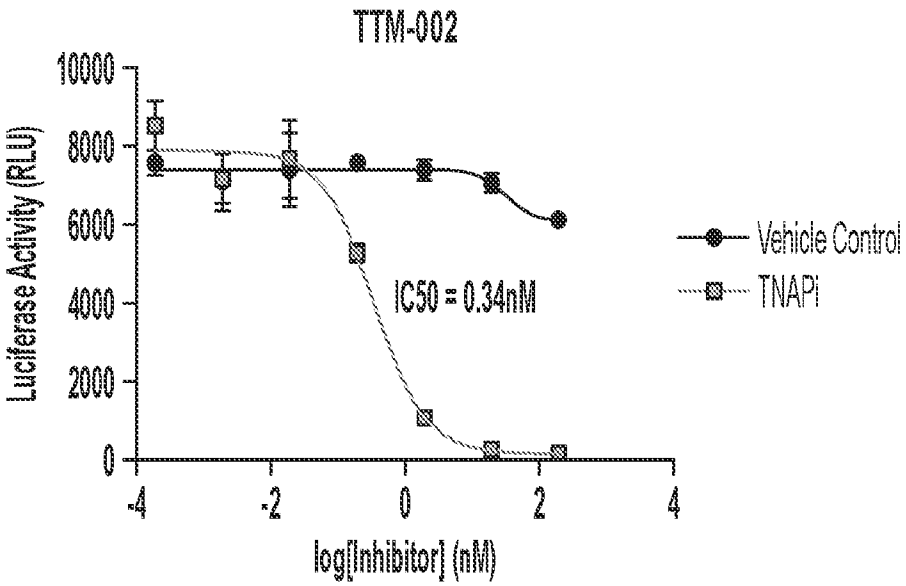

TTM-002
Luciferase Activity (RLU)
10000
8000
6000
4000
2000
0
-4
-2
0
2
4
log[Inhibitor] (nM)
IC50 = 0.34nM
Vehicle Control
TNAPi

FIG. 5C

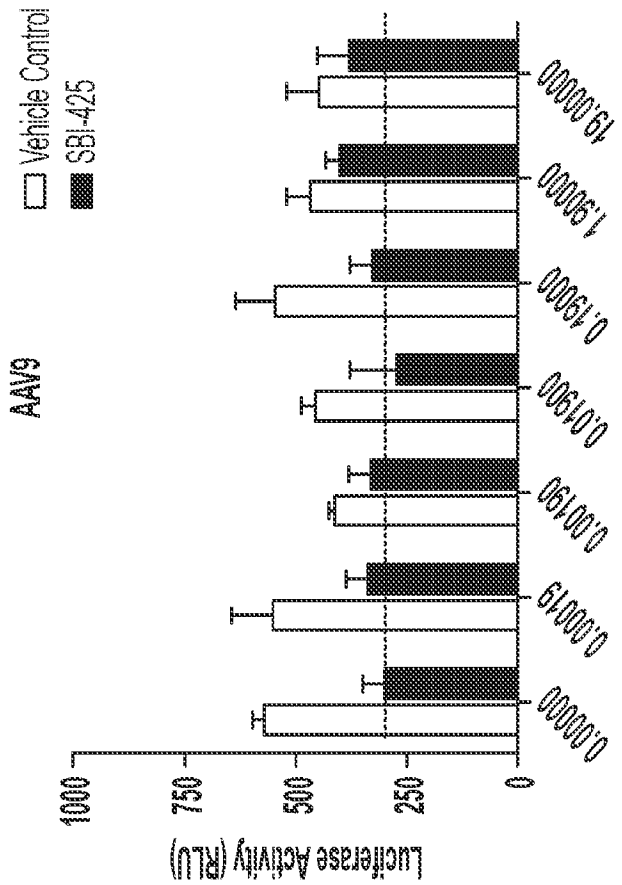

AAV9
Vehicle Control
SBI-425
Luciferase Activity (RLU)
1000
750
500
250
0
0.00000
0.00019
0.00061
0.00190
0.00610
0.01900
0.06100
0.19000
1.90000
19.00000
ALPL Inhibitor (nM)

FIG. 6B

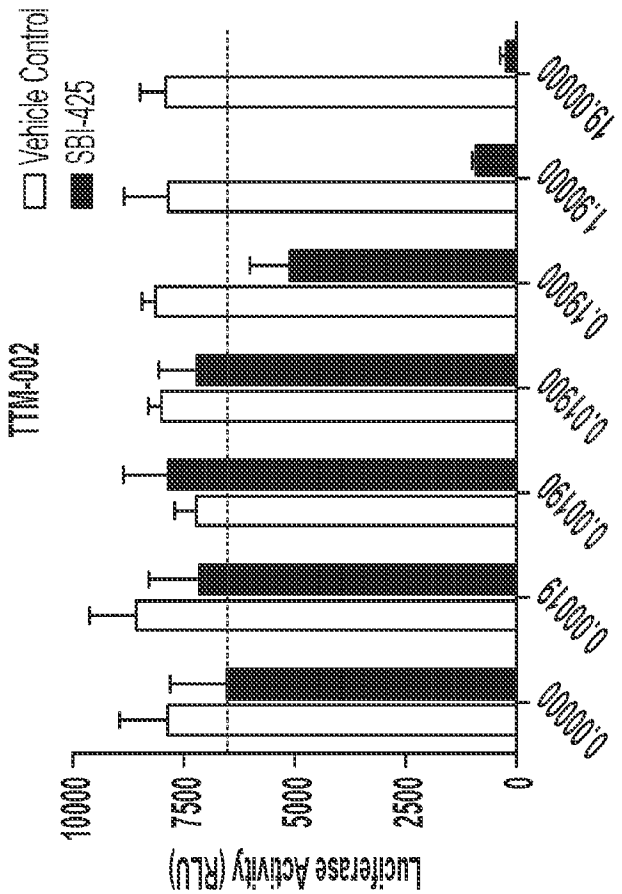

TTM-002
Vehicle Control
SBI-425
Luciferase Activity (RLU)
10000
7500
5000
2500
0
0.00000
0.00019
0.00061
0.00190
0.00610
0.01900
0.06100
0.19000
1.90000
19.00000
ALPL Inhibitor (nM)

FIG. 6A

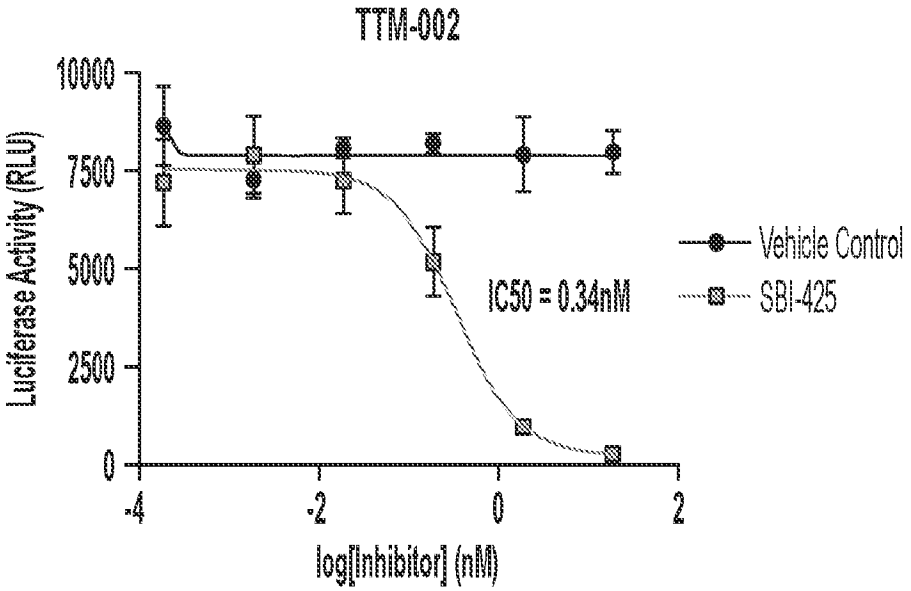

TTM-002
Luciferase Activity (RLU)
10000
7500
5000
2500
0
-4
-2
0
2
log[Inhibitor] (nM)
IC50 = 0.34nM
Vehicle Control
SBI-425

FIG. 6C

FIG. 7A

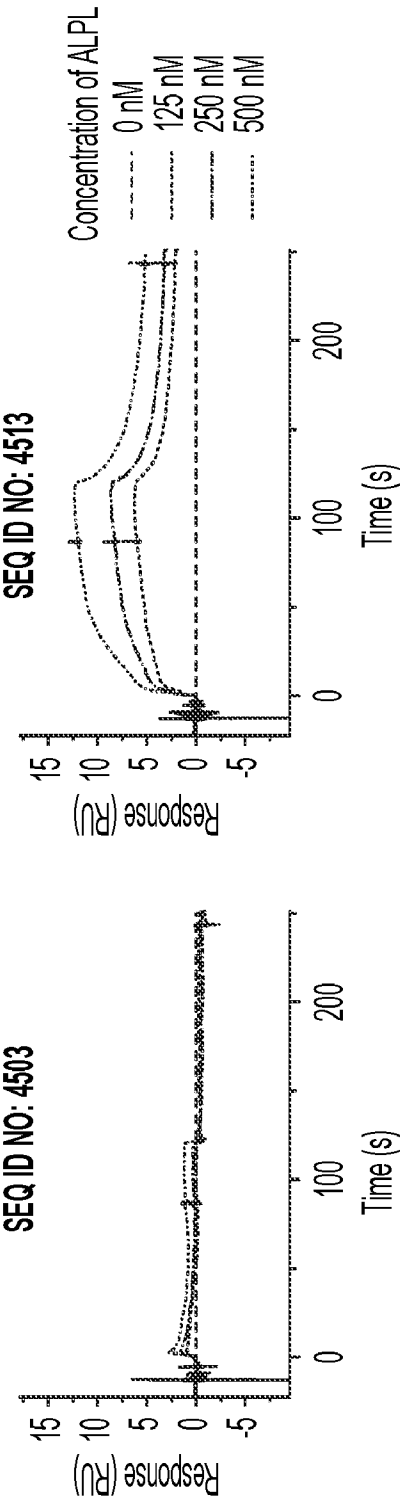

SEQ ID NO: 4513
SEQ ID NO: 4503
Concentration of ALPL
0 nM
125 nM
250 nM
500 nM
Response (RU)
Time (s)
15
10
5
0
-5
0
100
200

FIG. 7B

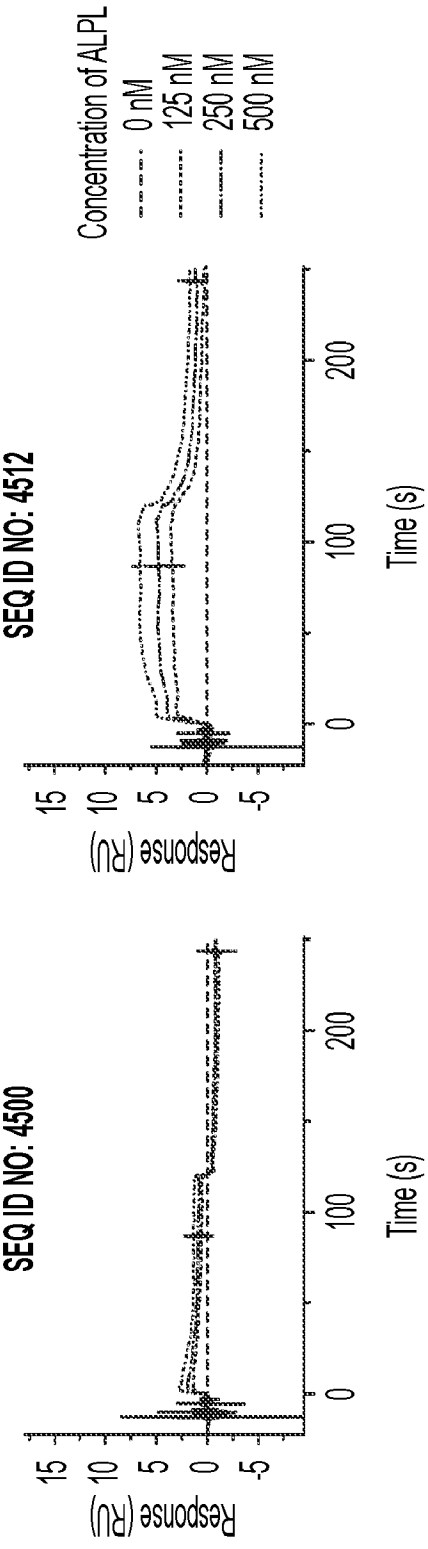

SEQ ID NO: 4512
SEQ ID NO: 4500
Concentration of ALPL
0 nM
125 nM
250 nM
500 nM
Response (RU)
Time (s)
15
10
5
0
-5
0
100
200

FIG. 8A

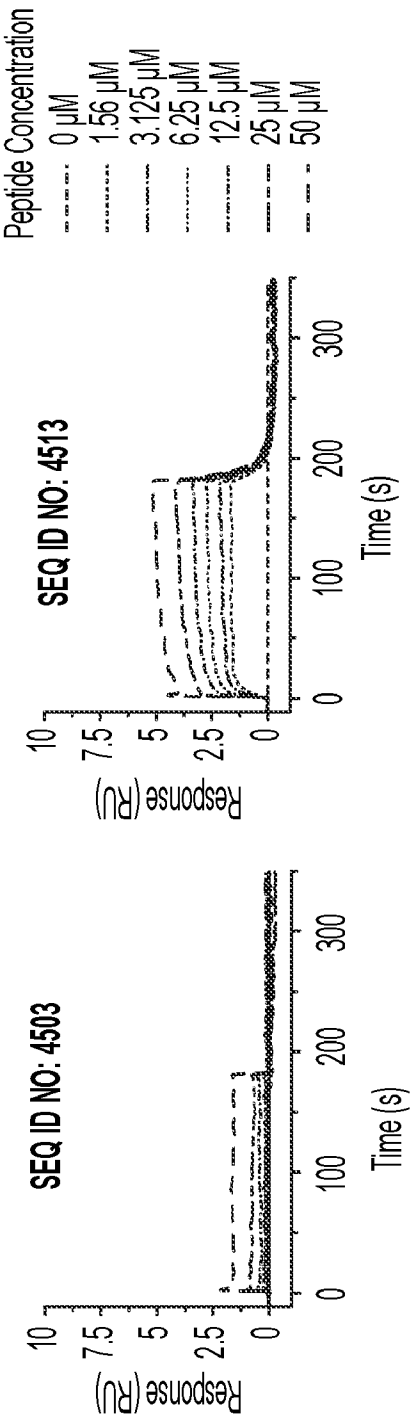

Peptide Concentration
0 µM
1.56 µM
3.125 µM
6.25 µM
12.5 µM
25 µM
50 µM
SEQ ID NO: 4513
SEQ ID NO: 4503
Response (RU)
Time (s)
0
2.5
5
7.5
10
100
200
300

FIG. 8B

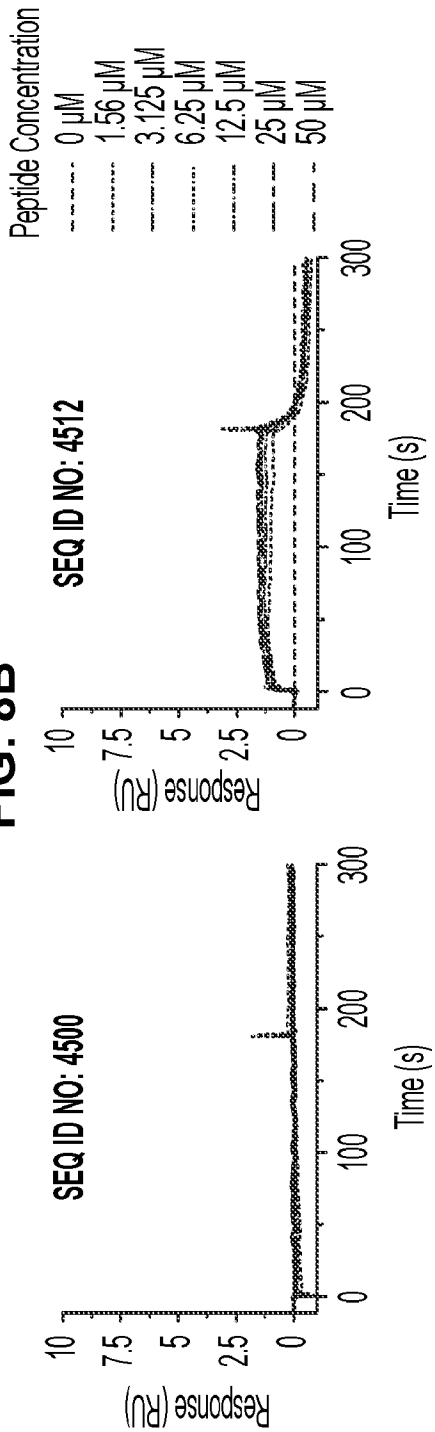

Peptide Concentration
0 µM
1.56 µM
3.125 µM
6.25 µM
12.5 µM
25 µM
50 µM
SEQ ID NO: 4512
SEQ ID NO: 4500
Response (RU)
Time (s)
0
2.5
5
7.5
10
100
200
300

FIG. 9A

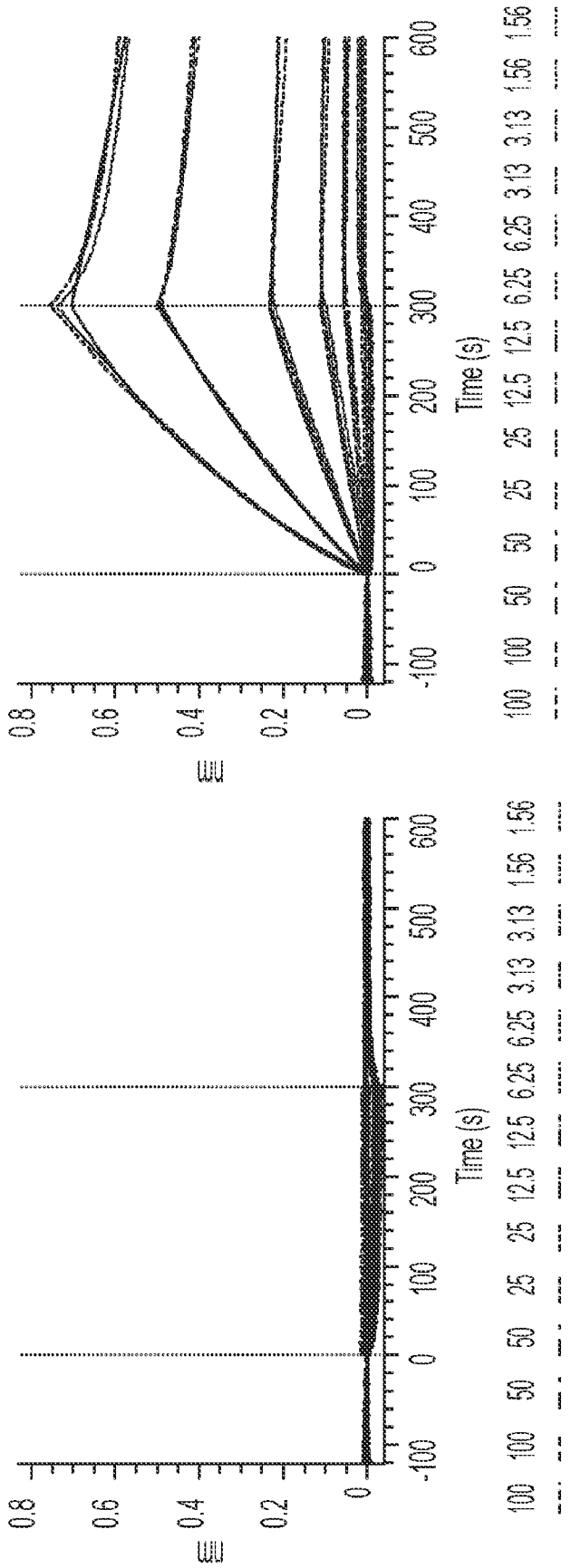

SEQ ID NO: 4503
SEQ ID NO: 4513
RU
Time (s)
No Binding
kon : 3.2e4 1/Ms
koff : 6.7e-4 1/s
KD : 20.7 nM

FIG. 9B

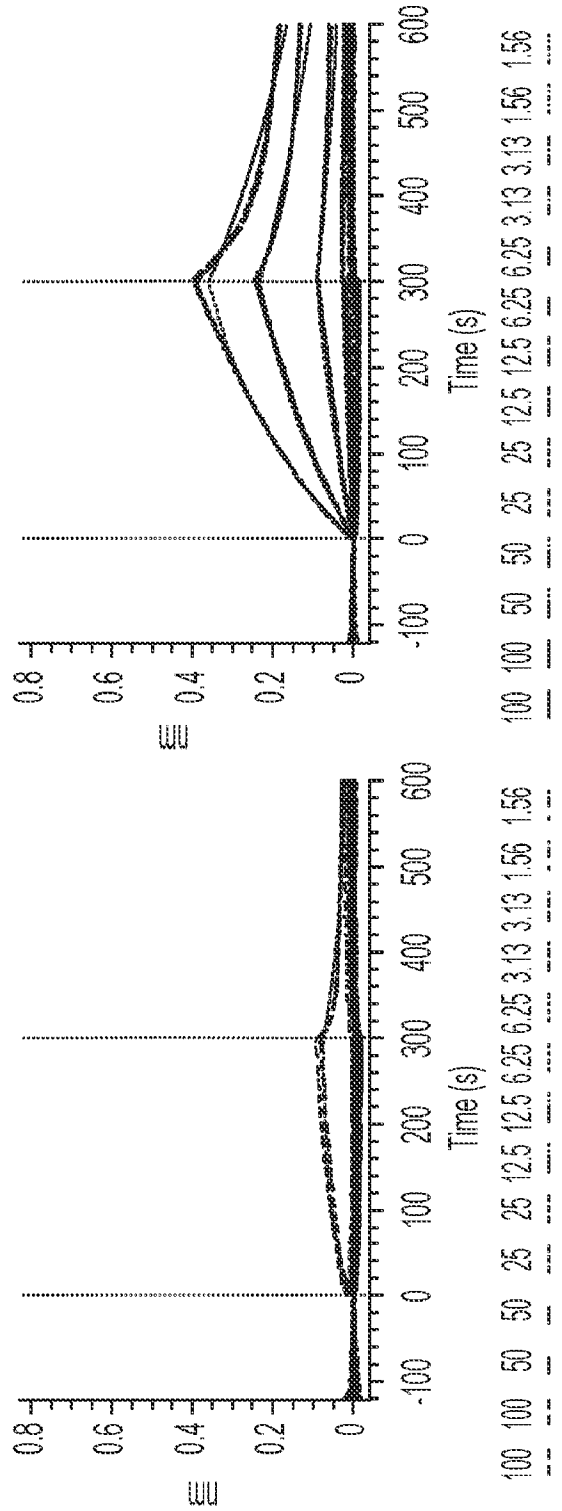

SEQ ID NO: 4500
SEQ ID NO: 4512
RU
Time (s)
No Binding
k_{on} : 2.3e4 1/Ms
k_{off} : 2.6e-3 1/s
K_D : 112 nM

FIG. 10A

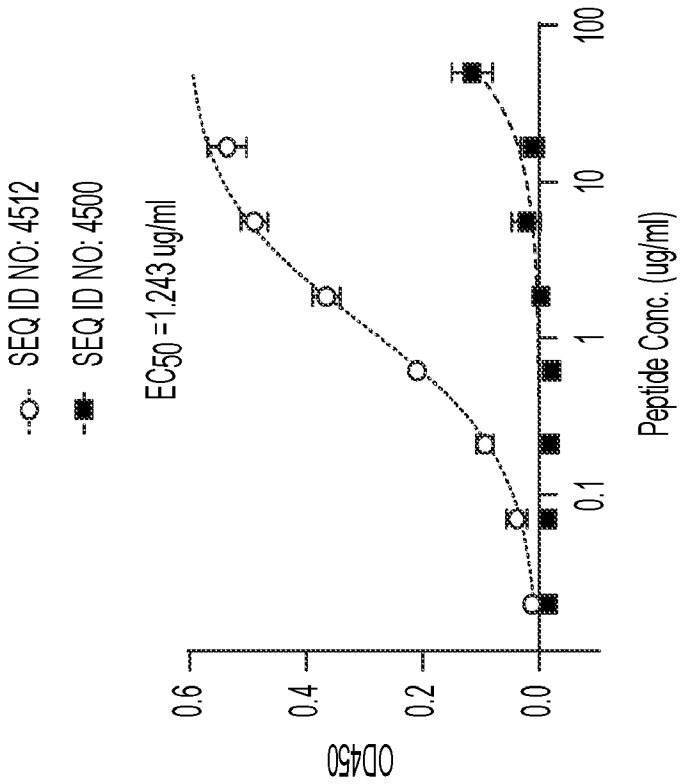
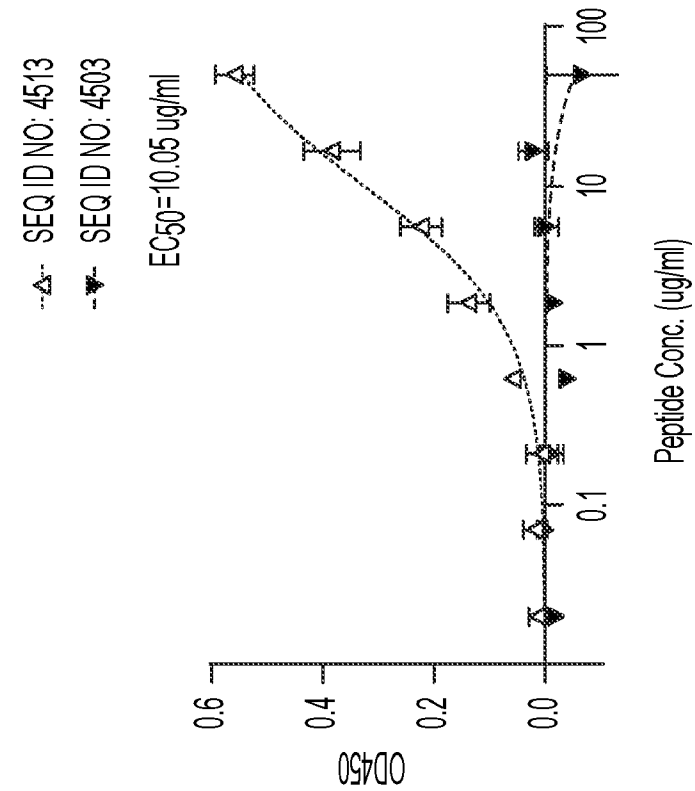


FIG. 10B



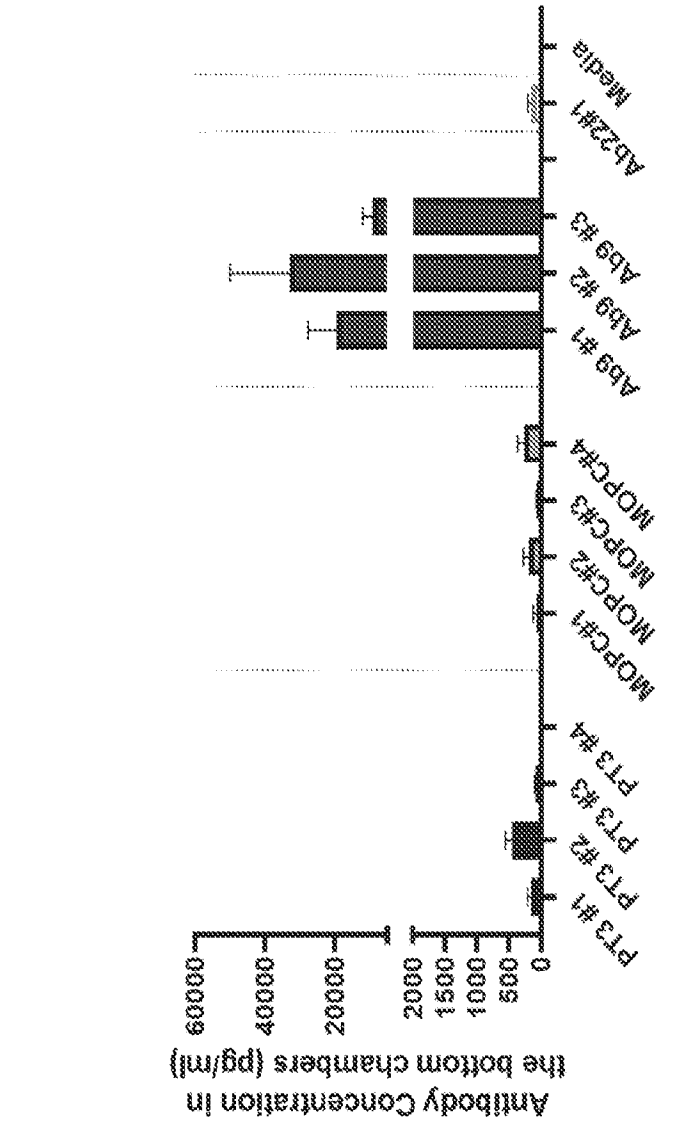


FIG. 11B

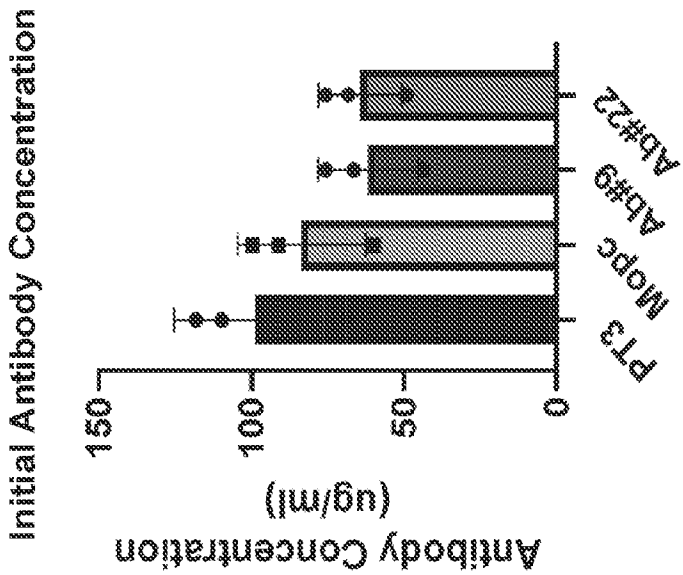


FIG. 11A

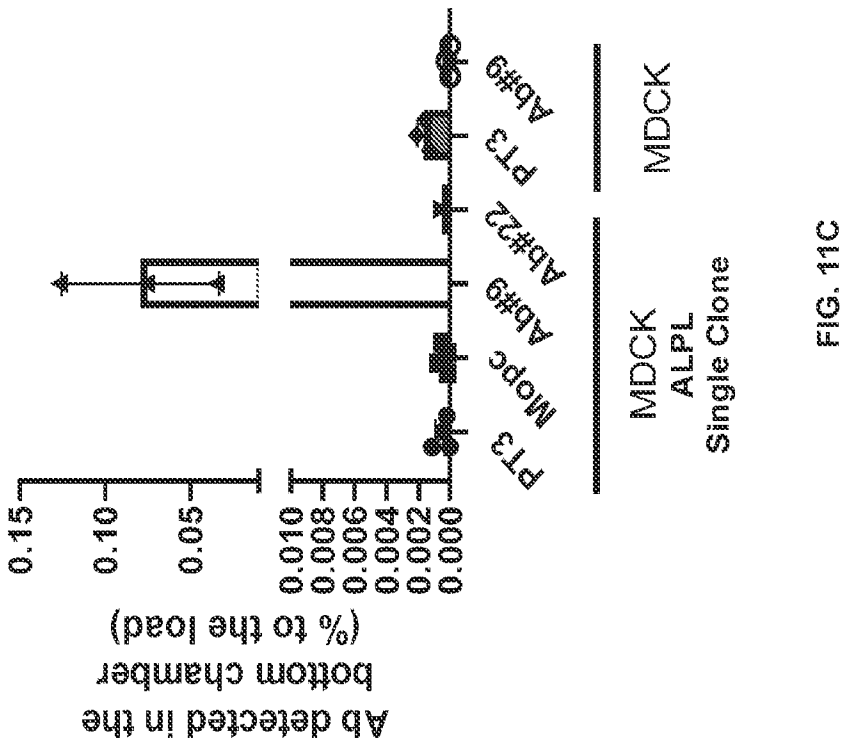

Ab detected in the bottom chamber (% to the load)
0.15
0.10
0.05
0.010
0.008
0.006
0.004
0.002
0.000
PT3
MOPC
Ab#9
Ab#22
PT3
Ab#9
MDCK ALPL Single Clone
MDCK

FIG. 11C

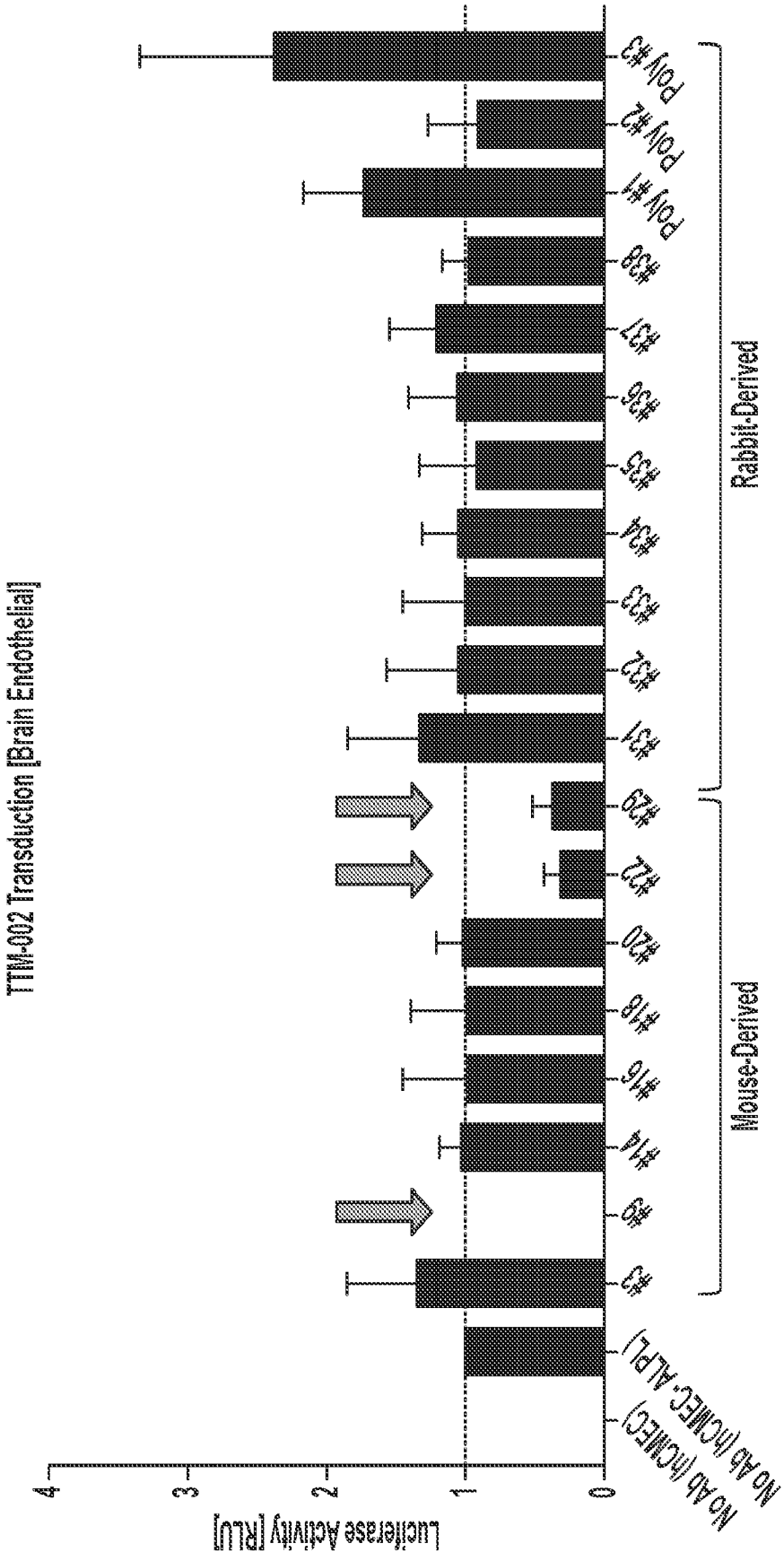

TTM-002 Transduction [Brain Endothelial]
Luciferase Activity [RLU]
0
1
2
3
4
No AB (hCMEC)
No AB (hCMEC; ALPL)
#3
#9
#14
#16
#18
#20
#22
#29
#31
#32
#33
#34
#35
#36
#37
#38
Poly #1
Poly #2
Poly #3
Mouse-Derived
Rabbit-Derived

FIG. 12