



(51) International Patent Classification:

C07C 229/48 (2006.01) *C07K 19/00* (2006.01)
C12N 15/113 (2010.01) *A61K 47/48* (2006.01)
C07K 7/04 (2006.01)

(21) International Application Number:

PCT/US2015/032143

(22) International Filing Date:

22 May 2015 (22.05.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/002,298 23 May 2014 (23.05.2014) US

(71) Applicant: GENZYME CORPORATION [US/US]; 500 Kendall Street, Cambridge, Massachusetts 02142 (US).

(72) Inventor: WEEDEN, Timothy E.; c/o SANOFI, 55 Corporate Drive, Mail Code: 55A-505A, Bridgewater, New Jersey 08807 (US).

(74) Agents: TOUSIGNANT, Jennifer D. et al.; SANOFI U.S., 55 Corporate Drive, Mail Code: 55A - 505A, Bridgewater, New Jersey 08807 (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, CZ, DE, DK, DM,

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))
- of inventorship (Rule 4.17(iv))

Published:

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

(54) Title: LIBRARY AND SCREENING APPROACH FOR IMPROVED CELL PENETRATING PEPTIDES

(57) Abstract: The invention provides compounds and related compositions that include peptide sequences for enabling transduction of a cargo into cells. The invention further includes methods for screening a library of such compounds.



LIBRARY AND SCREENING APPROACH FOR IMPROVED CELL PENETRATING PEPTIDES

CROSS-REFERENCE TO RELATED APPLICATION

[001] This application claims the benefit of priority to U.S. Provisional Patent Application No 62/002,298, filed May 23, 2014, which is incorporated by reference in its entirety.

DESCRIPTION OF THE INVENTION

[002] The present invention relates to a library of compounds including cell penetrating peptides, and methods of screening such a library for improved transduction, toxicity, and localization in cells.

[003] Cell penetrating peptides (CPP) are peptides that are capable of crossing cell membranes. CPPs are typically short, amphipathic or cationic alpha helices. Naturally occurring examples of CPP include HIV-1 TAT (Transactivator of Transcription), associated with the sequence GRKKRRQRRRPP (SEQ ID NO.: 1), and Drosophila ANTP (Antennapedia), associated with the sequence RQIKIWFQNRRMKWKK (SEQ ID NO.: 2). See Jones, S. W., R. Christison, et al. (2005) *Br J Pharmacol* 145(8): 1093-1102. CPPs may be associated with a molecular cargo, such as a peptide sequence covalently linked to the C-terminus of the CPP. Some CPPs are capable of transducing a cargo more than ten times the weight of the CPP into cells.

[004] The use of CPPs to import molecules into cells has great therapeutic potential. For example, some diseases might be effectively treated by introducing a therapeutic compound into the cells of a patient. Some diseases might be effectively treated by introducing a therapeutic compound into just a specific cell type.

Accordingly, there is a great need for the identification of CPPs that can efficiently transduce a cargo into cells. Such CPPs would be even more useful if they are associated with low toxicity and specific targeting to particular subcellular compartments.

[005] Accordingly, it is a primary object of the invention to provide peptide molecular scaffolds including peptide sequences that are associated with improved transduction, improved toxicity, improved protease stability, and improved targeting, as well as methods of screening to identify additional compounds having these qualities.

[006] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl.

[007] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the variable sequence comprises an α -, β -, γ -, or δ -amino acid, or a cycloalkane structure.

[008] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the variable sequence causes the compound to be targeted to the nucleus.

[009] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein n ≥ 7; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the variable sequence causes the compound to be targeted to the cytosol.

[010] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein n ≥ 7; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the variable sequence causes the compound to be targeted to the mitochondria.

[011] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein n ≥ 7; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the spacer comprises an aminohexanoic acid (Ahx).

[012] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein n ≥ 7; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the spacer comprises (Ahx)B (SEQ ID NO.: 4), wherein B is selected from β -alanine or β -glycine.

[013] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises a cleavage motif.

[014] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; wherein the linker comprises a cleavage motif; and wherein the cleavage motif is cleavable by a hydrolytic enzyme.

[015] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FS (SEQ ID NO.: 5).

[016] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FSQ (SEQ ID NO.: 6).

[017] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FSQK (SEQ ID NO.: 7).

[018] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine.

[019] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of glutamic acid.

[020] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is

Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of aspartic acid.

[021] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of lysine.

[022] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of serine.

[023] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues

selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of threonine.

[024] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the spacer comprises an aminohexanoic acid (Ahx); and wherein the linker comprises FSQG-OH (SEQ ID NO.: 9).

[025] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl; and wherein the linker comprises FxyB (SEQ ID NO.: 8), x is any amino acid, y is selected from glutamic acid (E), aspartic acid (D), lysine (K), serine (S), and threonine (T), and B is β -alanine; and wherein y is a non-natural analog of threonine; and wherein n is 7, and the spacer is (Ahx) (SEQ ID NO.: 10).

[026] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl, wherein the variable sequence, the spacer, and the linker comprise a sequence selected from: SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11.

[027] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker .

[028] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the cargo comprises a chromophore.

[029] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the cargo comprises fluorescein.

[030] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the cargo comprises one or more therapeutic compounds.

[031] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the cargo comprises a phosphorodiamidate morpholino oligomer (PMO).

[032] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the cargo further comprises one or more additional PMOs.

[033] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the compound is capable of crossing the blood-brain barrier.

[034] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and

Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, and further comprising a carrier.

[035] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, and further comprising a carrier, wherein the carrier comprises saline.

[036] In some embodiments, the invention provides a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, further comprising a cargo conjugated to the linker, wherein the compound is lyophilized.

[037] In some embodiments, the invention provides a method of screening for compounds with improved transduction, comprising: incubating cells, a cell medium, a control compound, and one or more compounds comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O)_nR (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, wherein said compound may or may not further comprise a cargo conjugated to the linker; removing the cell medium; obtaining a measurement of the amount of transduced one or more compounds as measured by the level of fluorescence of the one or more compounds; obtaining a measurement of the amount of transduced control compound as measured by the level of

fluorescence of the control compound; and identifying the one or more compounds wherein the level of fluorescence of the one or more compounds is greater than the level of fluorescence of the control compound.

[038] In some embodiments, the invention provides a method of screening for cell penetrating peptide sequences with reduced toxicity, comprising: incubating a first sample of cells with one or more compounds one or more compounds comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O) n R (SEQ ID NO.: 3), wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, wherein said compound comprises a chromophore cargo conjugated to the linker; incubating a second sample of cells with a control compound; and identifying the one or more compounds in the first sample where the number of cells in the first sample of cells is greater than 75% of the number of cells in the second sample of cells.

[039] In some embodiments, the invention provides a method of screening for compounds with nuclear localization, comprising: incubating a sample of cells with a compound comprising the formula variable sequence – spacer - linker, wherein the variable sequence is Ac-R(O) n R (SEQ ID NO.: 3) wherein $n \geq 7$; wherein O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is cis-2-aminocyclopentane-1-carbonyl, wherein said compound comprise a chromophore cargo conjugated to the linker; incubating the sample of cells with a nuclear stain; imaging the sample of cells; and identifying the compound wherein the fluorescent signal of the compound substantially overlaps with the location of the nuclear stain.

[040] Additional objects and advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims.

[041] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

[042] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and together with the description, serve to explain the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[043] Figure 1 shows exemplary conjugation strategies.

[044] Figure 2 shows a variety of exemplary peptide scaffolds.

[045] Figure 3 shows the transduction efficiency of the various peptide scaffolds in Figure 6.

[046] Figure 4 shows exemplary beta-, gamma-, and delta-amino acids that can be integrated into a peptide scaffold to control or constrain the peptide structure.

[047] Figure 5 shows an exemplary peptide scaffold library.

[048] Figure 6 shows an exemplary assessment of the purity of a library of peptides.

[049] Figure 7 shows an exemplary approach for identifying improved peptides by evaluating a peptide library for transduction efficiency.

[050] Figure 8 shows exemplary peptides selected for a secondary screen based on a transduction assay.

[051] Figure 9 shows the results of a toxicity evaluation of certain exemplary peptides.

[052] Figure 10 shows exemplary images evaluating the nuclear localization of a variety of exemplary peptides.

[053] Figure 11 shows exemplary images evaluating the nuclear localization of a variety of exemplary peptides.

[054] Figure 12 shows exemplary peptides associated with nuclear localization (18a) and cytosolic localization (18b).

DETAILED DESCRIPTION OF THE INVENTION

[055] As used herein, the term “peptide” is used broadly to refer to a polymer of amino acids. A peptide may include unmodified, modified, or nonstandard amino acids and nonstandard conjugation between amino acids or subunits in the polymer. Nonstandard amino acids may include analogs of the amino acids found naturally in proteins.

[056] The term “nonstandard” used in conjunction with terms including but not limited to “residue” is to be interpreted in the same manner as if it were used in conjunction with “amino acid.”

[057] The term “PMO” refers to phosphorodiamidate morpholino oligomer.

[058] The term “PPMO” refers to peptide phosphorodiamidate morpholino oligomer, or a peptide conjugated to one or more PMO moiety.

[059] The term molecular or structural “scaffold” is used broadly to refer to a class of peptides or compounds that may share one or more aspect of their chemical structures.

[060] The compounds of the invention comprise a peptide library that may enable efficient transduction of a cargo into cells.

[061] In some embodiments, the peptide moiety of the peptide library can be structured according to the formula “variable sequence – spacer – linker.” See, e.g., Figure 5.

[062] In some embodiments, the variable sequence can be structured according to the formula $\text{Ac-R(O)}_n\text{R}$ (SEQ ID NO.: 3). O is a sequence of residues selected from R, X, and Z; wherein R is L-arginine or arginine; X is 3-cis-aminocyclohexane or 1,3 cis-aminocyclohexane carboxylic acid; and Z is cis-2-aminocyclopentane-1-carbonyl or cis-(1R,2S)-2-aminocyclopentane carboxylic acid. The sequence length of O can range from 1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-20, 21-25, 26-35, or more residues, where the length of O is n residues. In some embodiments, O can be 6, 7, 8, 9, or 10 residues long. O can contain any combination of 0, 1, or more residues that are R, X, and Z. In some embodiments, O can also include other types of residues, such as proline, glycine, or alanine, or additional modified or nonstandard amino acids. In some embodiments, the variable sequence can form an alpha helix, an H10/12 helix, a circular fold, or a Z-strand. In some embodiments, the variable sequence includes alpha, beta, gamma, or delta amino acids, or cycloalkane structures.

[063] In some embodiments, less than 50% of the amino acid residues in the compound are arginine. In some embodiments, 40-50%, 35-40%, 30-35%, 25-30%, 20-25%, 15-20%, 10-15%, 5-10%, or less than 5% of the amino acid residues in the compound or in the variable sequence are arginine.

[064] In some embodiments, the variable sequence can cause the compound containing the variable sequence to be targeted to a specific cellular compartment. In some embodiments, the compound is targeted to the nucleus. In some embodiments, the compound is targeted to the cytosol. In some

embodiments, the compound is targeted to the mitochondria. In some embodiments, the compound is targeted to two or more cellular compartments.

[065] In some embodiments, the spacer includes an aminohexanoic acid (Ahx). In some embodiments, the spacer includes alpha, beta, gamma, or delta amino acids, or cycloalkane structures. In some embodiments, the spacer comprises (Ahx)B, wherein B is β -alanine or β -glycine.

[066] In some embodiments, the linker is a sequence that contains a cleavage motif. In some embodiments, the cleavage motif can be cleaved by any hydrolytic enzyme. In some embodiments, the cleavage motif can be cleaved by a peptidase or protease such as cathepsin or trypsin. In some embodiments, the linker can be designed to include a cleavage motif recognized by a particular serine protease, threonine protease, cysteine protease, aspartate protease, glutamic acid protease, or metalloprotease, or a group of more than one peptidase. In some embodiments, the linker may include two or more cleavage motifs that are overlapping or nonoverlapping. In some embodiments, the linker may contain no cleavage motifs. In some embodiments, the linker may be designed so that less than 100%, less than 75%, less than 50%, or less than 10% of the linkers are cleaved. In some embodiments, the linker is designed so that greater than 90% or 99% of the linkers are cleaved.

[067] In some embodiments, the linker includes the sequence FS (SEQ ID NO.: 5). In some embodiments, the linker includes the sequence FSQ (SEQ ID NO.: 6) or FSQK (SEQ ID NO.: 7). In some embodiments, the linker includes the sequence FxyB (SEQ ID NO.: 8), where x is any amino acid, standard or nonstandard, y is glutamic acid (E), aspartic acid (D), and lysine (K), serine (S), or threonine (T), and B is β -alanine or β -glycine.

[068] In some embodiments, compound may include a variable sequence – spacer – linker according to any of the sequences of Table 1:

Table 1:

Hit	Sequence	Localization	SEQ ID NO:
2C4	Ac-RXXXXXXRRR(Ahx)FSQG-OH	Nucleus	11
4G9	Ac-RXXXXXXRR(Ahx)FSQG-OH	Nucleus	12
9F5	Ac-RXXXRXRXR(Ahx)FSQG-OH	Nucleus	13
12G4	Ac-RRXXZXXXR(Ahx)FSQG-OH	Nucleus	14
12D10	Ac-RRRXXXXXR(Ahx)FSQG-OH	Nucleus	15
12D11	Ac-RXRXXXXXR(Ahx)FSQG-OH	Nucleus	16
12E4	Ac-RRZXXXXXR(Ahx)FSQG-OH	Nucleus	17
21A5	Ac-RXXXXZXZR(Ahx)FSQG-OH	Nucleus	18
11G1	Ac-RXXZXRXRXR(Ahx)FSQG-OH	Cytosol	19
12D4	Ac-RRRXXXXXR(Ahx)FSQG-OH	Cytosol	20
13D2	Ac-RRZXXZXXR(Ahx)FSQG-OH	Cytosol	21

[069] In some embodiments, the compound may include the following sequences set forth in Table 2 (see also Figure 8):

Table 2:

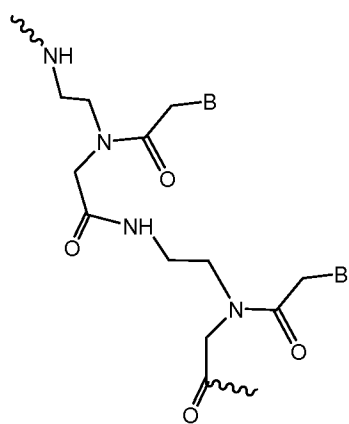
Hit	Sequence	SEQ ID NO:
9H8	Ac-RXXXXXXRXR(Ahx)	22
9H9	Ac-RZXXXXRXR(Ahx)	23
9H11	Ac-RXZXXRXRXR(Ahx)	24
1A2	Ac-RRRRRRRRRR(Ahx)	25
12D12	Ac-RZRXXXXXR(Ahx)	26
13D3	Ac-RXZXXZXXR(Ahx)	27
12D10	Ac-RRRXXXXXR(Ahx)	28
2C4	Ac-RXXXXXXRRR(Ahx)	29
4G9	Ac-RXXXXXXRR(Ahx)	30
11F4	Ac-RXXXXRXRXR(Ahx)	31
9F5	Ac-RXXXRXRXR(Ahx)	32
12D11	Ac-RXRXXXXXR(Ahx)	33
20B7	Ac-RXXXXXXXZR(Ahx)	34
20C4	Ac-RXXZXXXZR(Ahx)	35
5D4	Ac-RXXXRZXRR(Ahx)	36
9H7	Ac-RRXXXXRXR(Ahx)	37
5B1	Ac-RXXXZXRRR(Ahx)	38
4H6	Ac-RXXZXXXRR(Ahx)	39
12D4	Ac-RRRXXXXXR(Ahx)	40
15A8	Ac-RXXXXXXZXR(Ahx)	41
12D8	Ac-RXZRXXXXR(Ahx)	42

Hit	Sequence	SEQ ID NO:
12E3	Ac-RZXXXXXXR(Ahx)	43
12H2	Ac-RXXZZXXXR(Ahx)	44
4G10	Ac-RZXXXXXRR(Ahx)	45
15H5	Ac-RXXXXZZXR(Ahx)	46
11F1	Ac-RXRXXRXXR(Ahx)	47
21C7	Ac-RRXXZZXZR(Ahx)	48
12E2	Ac-RXXXXXXXR(Ahx)	49
12E4	Ac-RRZXXXXXR(Ahx)	50
12G4	Ac-RRXXZXXXR(Ahx)	51
21A5	Ac-RXXXXZXZR(Ahx)	52

[070] In some embodiments, the compound may include one or more cargo conjugated to the linker. In some embodiments, the cargo includes a chromophore, such as a fluorescein. The fluorescein may include Alexa Fluor® 488 5-Carboxamido-(6-Azidohexyl). In some embodiments, the cargo includes one or more therapeutic compounds, such as PMO. In some embodiments, the cargo may include a chromophore and one or more therapeutic compound. In some embodiments, the compound is capable of crossing the blood-brain barrier.

[071] In some embodiments, the cargo is any kind of compound or macromolecule, including small molecules, nucleic acids, peptides, and proteins.

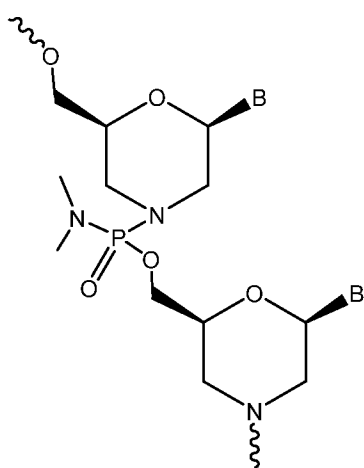
[072] In some embodiments, the therapeutic compound may include an antisense oligonucleotides (AON) AONs according to the present disclosure include PMO compounds as well as PNA compounds, phosphoramidate compounds, methylene methylimino ("MMI") compounds, 2-O-methyl compounds and 2-methoxy ethyl compounds, wherein the oligonucleobase of each subunit are set forth below and are known in the art. The oligonucleotide compounds are synthetic analogs of natural nucleic acids. In particular, instead of deoxyribose rings and phosphate-linkages, the oligonucleotide compounds comprise subunits comprised of the respective oligonucleotide subunits shown below:



[073]

PNA

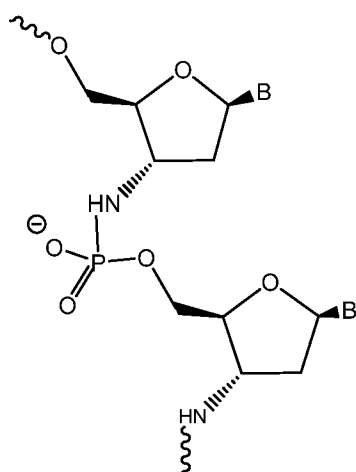
Formula I



[074]

Morpholino

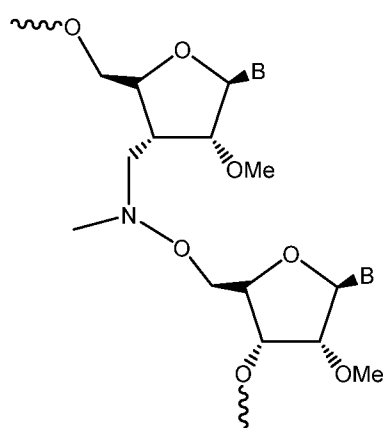
Formula II



[075]

Phosphoramidate

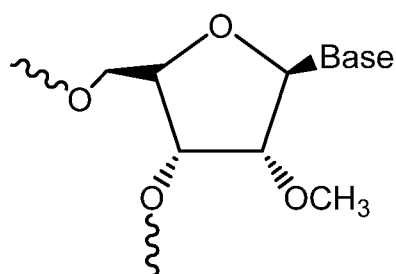
Formula III



[076]

MMI

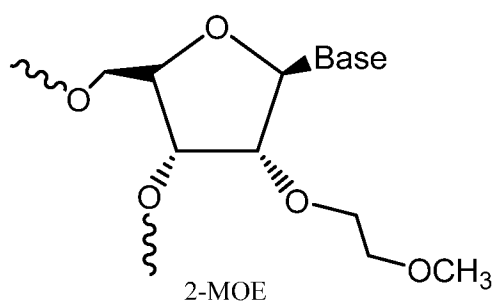
Formula IV



[077]

2-OMe

Formula V



[078]

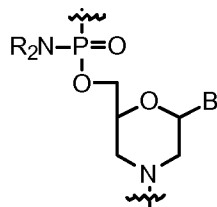
2-MOE

Formula VI

[079] In the case of each of Formula 1-VI, B is a nucleotide base. The primary nucleobases are cytosine (DNA and RNA), guanine (DNA and RNA), adenine (DNA and RNA), thymine (DNA) and uracil (RNA), abbreviated as C, G, A, T, and U, respectively. A, G, C, and T appear in DNA, these molecules are called DNA-bases; A, G, C, and U are called RNA-bases. Uracil replaces thymine in RNA. These two bases are identical except that uracil lacks the 5' methyl group. Adenine

and guanine belong to the double-ringed class of molecules called purines (abbreviated as R). Cytosine, thymine, and uracil are all pyrimidines (abbreviated as Y).

[080] PMO compounds possess subunits comprised of morpholine rings and phosphorodiamidate-linking groups, respectively. For example, the present disclosure includes a PMO compound comprising from 15 to 30 subunits of formula (I):



wherein R is an alkyl group and B is a naturally occurring purine or pyrimidine nucleotide base selected from cytosine (C), guanine (G), adenine (A), or thymine (T).

[081] In at least one embodiment, the PMO compound has from 15-25 subunits of formula (I). In another embodiment, the PMO compound has from 20-25 subunits of formula (I). In yet another embodiment, the PMO compound has about 25 subunits of formula (I), such as from 24-26 subunits.

[082] The therapeutic compound included in the compounds of the invention may be any protein where a deficiency, lack of, or aberrant expression of that protein gives rise to a disease or condition. In some embodiments, the therapeutic compound may be an enzyme. In other embodiments, the therapeutic compound is a compound that is not normally present or is not normally present in sufficient quantities in a subject to achieve the desired therapeutic effect.

[083] For example, a non-limiting selection of suitable therapeutic compounds includes erythropoietin, insulin, human growth hormone, cystic fibrosis

transmembrane conductance regulator (CFTR), insulin, alpha-galactosidase A, alpha-L-iduronidase, iduronate-2-sulfatase, N-acetylglucosamine-1-phosphate transferase, N-acetylglucosaminidase, alpha-glucosaminide acetyltransferase, N-acetylglucosamine 6-sulfatase, N-acetylgalactosamine-4-sulfatase, beta-glucosidase, galactose-6-sulfate sulfatase, beta-galactosidase, beta-glucuronidase, glucocerebrosidase, heparan sulfamidase, hyaluronidase, galactocerebrosidase, ornithine transcarbamylase (OTC), carbamoyl-phosphate synthetase 1 (CPS1), argininosuccinate synthetase (ASS1), argininosuccinate lyase (ASL), and arginase 1 (ARG1), glucose-6-phosphatase, glucose-6-phosphate translocase, glycogen debranching enzyme, lysosomal alpha-glucosidase, 1,4-alpha-glucan branching enzyme, glycogen phosphorylase, phosphofructokinase, liver phosphorylase, GLUT-2, UDP glycogen synthase, alpha-L-iduronidase, iduronate sulfate sulfatase, heparan sulfate sulfamidase, alpha-N-acetylglucose amidase, alpha-glucosaminid-N-acetyltransferase, and N-acetylglucosamine-6-sulfate sulfatase.

[084] [In some embodiments, the cargo may be conjugated to the peptide moiety or linker using a variety of chemical approaches. The cargo may be conjugated using hydrazine ligation, oxime ligation, and 1,2,3-triazole reactions. See, e.g., Figure 1 and Example 1.

[085] In some embodiments, the invention may include a composition of the compound, including the variable sequence – spacer – linker and cargo, and a carrier. A carrier may be any pharmaceutical carrier, vehicle, diluent, excipient and the like which are generally intended for use in connection with the administration of biologically active agents, including proteins and nucleic acids. In some embodiments, the carrier may comprise saline or a buffer, or may incorporate a molecular delivery system such as a plasmid or copolymer.

[086] In some embodiments, the compound may be lyophilized.

[087] Certain embodiments of the invention include methods for screening a compound library for compounds with desirable properties.

[088] In some embodiments, compounds with improved transduction into cells may be identified by incubating cells, a cell medium, a control compound, and one or more compounds according to embodiments of the invention; removing the cell medium; obtaining a measurement of the amount of transduced one or more compounds as measured by the level of fluorescence of the one or more compounds; obtaining a measurement of the amount of transduced control compound as measured by the level of fluorescence of the control compound; and identifying the one or more compounds wherein the level of fluorescence of the one or more compounds is greater than the level of fluorescence of the control compound. See, e.g., Example 3.

[089] In certain embodiments, fluorescence may be measured by including any type of chromophore as the cargo for each compound. In certain embodiments, other reporter signals may be used. For example, instead of fluorescence, a measurement of radioactivity may be used where the cargo is radioactive. In certain embodiments, more than one type of reporter signal may be used in the same assay.

[090] In certain embodiments, all of the cell medium is removed prior to measuring the reporter signal in the cells. In certain embodiments, most of the cell medium is removed.

[091] In certain embodiments, the cells, cell medium, control compound, and one or more compounds are incubated for less than two hours. In certain embodiments, the incubation is for less than 5, 10, 12, 18, or 24 hours. In certain embodiments, the incubation lasts for two days.

[092] In certain embodiments, compounds may be identified where the reporter signal (e.g., fluorescence or radioactivity) is 75% or greater when compared the reporter signal of a control compound. In certain embodiments, compounds may be identified where the reporter signal is greater than 100% of the reporter signal of the control compound.

[093] In certain embodiments, Ac-(RXR)₄XFSQKG-amide; X= 6-aminohexanoic acid (SEQ ID NO: 53) may be used in the control compound.

[094] In some embodiments, compounds with improved reduced toxicity may be identified by incubating a first sample of cells with one or more compounds according to according to embodiments of the invention; incubating a second sample of cells with a control compound; and identifying the one or more compounds in the first sample where the number of cells in the first sample of cells is greater than 75% of the number of cells in the second sample of cells. See, e.g., Example 4.

[095] In certain embodiments, the cell samples and compounds are incubated for less than two hours. In certain embodiments, the incubation is for less than 5, 10, 12, 18, or 24 hours. In certain embodiments, the incubation lasts for two days.

[096] In certain embodiments, compounds with nuclear localization may be identified by incubating a sample of cells with a compound according to embodiments of the invention; incubating the sample of cells with a nuclear stain; imaging the sample of cells; and identifying the compound wherein the fluorescent signal of the compound substantially overlaps with the location of the nuclear stain.

[097] A signal substantially overlaps where > 25% of the signal is overlapping across two or more signal images.

[098] Some embodiments, including compounds and compositions of the invention, may be administered using various techniques. Administration may be systemic or local. Local delivery may be effected by, for example, intravenous injection into the site of injury, disease manifestation, or pain. Administration may be by aerosol inhalation. Compositions may be provided in lozenges for oral, tracheal, or esophageal application; can be supplied in liquid, tablet or capsule form for administration to the stomach or intestines; can be supplied in suppository form for rectal or vaginal application; or can even be delivered to the eye by use of creams, drops.

[099] Additional embodiments of the invention include kits including compounds or compositions according to embodiments of the invention and instructions for use.

EXAMPLES

[0100] The following specific examples are to be construed as merely illustrative, and not limiting of the disclosure. Without further elaboration, it is believed that one skilled in the art can, based on the description herein, utilize the present invention to its fullest extent.

Example 1: Cargo conjugation techniques

[0101] Solulink Bis-arylhydrazone ligation. See Figure 1a.

[0102] HyNic - Peptide

[0103] Peptide is synthesized with Fmoc-Lysine- ϵ [hydrazinonicotinic acid]-OH (Solulink # S-3034) during solid phase peptide synthesis, to incorporate the HyNic(hydrazinonicotinic acid) moiety into the peptide structure. The peptide is cleaved from resin, HPLC purified and lyophilized.

[0104] 4FB - Oligonucleotide

[0105] The PMO is dissolved in 100mM PBS pH=8.0 and 20-fold molar excess of Sulfo-S-4FB (sulfo-succinimidyl-4-formylbenzamide) is solubilized in dimethylformamide and added as a 2:1 v/v ratio of PBS/DMF and allowed to react for 2hrs at room temperature. Excess S-4FB is removed by acetone precipitation followed by 3Kda dialysis exchange into 10mM NH_4HCO_3 and lyophilized.

[0106] Ligation

[0107] HyNic peptide is dissolved in 10mgs/ml of 100mM PBS, pH=6.0 and added to the 4FB-oligonucleotide. 1/10 volume of 100mM Aniline is added to catalyze the ligation and allowed to react for 2hrs at room temp. Conjugate is HPLC purified and bis-arylhydrazone bond is tracked using 354nm wavelength. Correct fractions are pooled and lyophilized.

[0108] Oxime Ligation. See Figure 1b.

[0109] Aminoxy – Peptide

[0110] Peptide is synthesized with N- α -Fmoc-N- β -(N-t.-Boc-amino-oxyacetyl)-L-diaminopropionic acid (Millipore# 852216) to introduce a side chain hydroxylamine during solid phase peptide synthesis. The peptide is cleaved from resin, HPLC purified and lyophilized.

[0111] 4FB - Oligonucleotide

[0112] The PMO is dissolved in 100mM PBS pH=8.0 and 20-fold molar excess of Sulfo-S-4FB (sulfo-succinimidyl-4-formylbenzamide) is solubilized in dimethylformamide and added as a 2:1 v/v ratio of PBS/DMF and allowed to react for 2hrs at room temperature. Excess S-4FB is removed by acetone precipitation followed by 3Kda dialysis exchange into 10mM NH_4HCO_3 and lyophilized.

[0113] Ligation

[0114] Aminoxy peptide is dissolved at 10mgs/ml of 10% Acetonitrile/water, pH=4.0 and the 4FB-oligonucleotide is added and allowed to react for 2hrs at room temp. Conjugate is HPLC purified and correct fractions verified by mass spectrometry are pooled and lyophilized.

[0115] Copper Catalyzed 1,2,3-Cycloaddition Reaction (Click Chemistry) See Figure 1c.

[0116] We have optimized the 1,2,3-triazole formation conditions using an alkyne bearing amino acid (Fmoc-d-Pra-OH) and azide bearing fluorescein (Alexa Fluor® 488 5-Carboxamido-(6-Azidohexyl)). We have applied this ligation approach to a blood brain barrier peptide pulled from the literature (Amino Acids, 2012 Jun;42(6):2373-81) where selective residues were replaced with (d-Pra) amino acid for ligation to a cargo. In tests so far the cargo has only been the Alexa Fluor®-azide; however ligation of a CXCR2 small molecule and measuring the conjugate's activity at the receptor in vitro upon and brain uptake in vivo is envisioned.

Example 2: Evaluating subcellular localization

[0117] 3e6 Hela cells (ATCC, CCL-2) were thawed into 50ml growth medium (RPMI 1640 (Gibco, 22400-089), 10% HI FBS (JRH, 12106-500M), 200 mM L-Glutamine (Gibco, 25030-081, 1:100), Pen-Strep (Gibco, 15140-122, 1:100)), in a vented T225 and incubated for 3 or 4 days at 37 degrees Celsius.

[0118] After 3 or 4 days, media was aspirated off, cells were washed with 10mls of Accutase® (eBioscience, 00-4555-56), 800ul of Accutase® was added, and flask was incubated for 3min at 37 degrees Celsius.

[0119] After 3min, flask was bumped on the side 20 times, 10ml of assay medium (RPMI 1640 (Gibco, 22400-089), 200 mM L-Glutamine (Gibco, 25030-081,

1:100), Pen-Strep (Gibco, 15140-122, 1:100)) was added to flask, and cells were counted.

[0120] Cells were spun down at 1500rpm for 5min, resuspended at 6×10^4 /100ul assay medium, and 100ul cell suspension was added to the appropriate wells on a 96-well Black/clear bottom TC plate (BD, 353219) that was coated for 24h with PDL (Sigma, P6407), thoroughly washed with sterile water, and allowed to dry, and stored at 4 degrees Celsius until use.

[0121] Cells were incubated overnight at 37 degrees Celsius before 20ul of a 6X stock of controls or peptides from the CPP library were thawed, mixed, and added carefully to cells or not to a final concentration of 50uM (17uM control) and incubated for 4h at room temp on a shaker at 400rpm under foil.

[0122] After 4h, 70ul of media was taken off.

[0123] 350ul of PBS (Gibco, 20012-027) was added on the slowest setting of a 12-channel Matrix pipettor and 350ul of medium was pipetted off 4 times.

[0124] Media was completely removed and cells fixed for 10 minutes in 100ul 10%Zn/Formalin.

[0125] Wash cells 3 x 2min in 200ul PBS.

[0126] Add 3-4 drops of antifade/DAPI mounting medium to each well.

[0127] Place the plate at -20C until imaging.

[0128] 96-well plates were scanned at 20x magnification using metamorph acquisition station with automated X,Y Z, autofocusing and a black and white 16-bit Orca-100 CCD camera. One image from each well was scanned for DAPI(Nuclear marker), then subsequently the FITC (peptide) channel using the positive control to set the exposure so all wells were collected at the same exposure (20 ms). Images

were imported into metamorph and the multiwavelength algorithm was run, utilizing the nuclear signal as a mask to quantify the level and area of FITC in each cell. Subsequent images were collected for presentations at 40x, imported into metamorph, assigned appropriate color channels (DAPI-Blue/FITC-Green) and merged into one image.

Example 3: Evaluating cell transduction

[0129] 3e6 Hela cells (ATCC, CCL-2) were thawed into 50ml growth medium (RPMI 1640 (Gibco, 22400-089), 10% HI FBS (JRH, 12106-500M), 200 mM L-Glutamine (Gibco, 25030-081, 1:100), Pen-Strep (Gibco, 15140-122, 1:100)), in a vented T225 and incubated for 3 or 4 days at 37 degrees Celsius.

[0130] After 3 or 4 days, media was aspirated off, cells were washed with 10mls of Accutase (eBioscience, 00-4555-56), 800ul of Accutase was added, and flask was incubated for 3min at 37 degrees Celsius.

[0131] After 3min, flask was bumped on the side 20 times, 10ml of assay medium (RPMI 1640 (Gibco, 22400-089), 200 mM L-Glutamine (Gibco, 25030-081, 1:100), Pen-Strep (Gibco, 15140-122, 1:100)) was added to flask, and cells were counted using a ViaCell.

[0132] Cells were spun down at 1500rpm for 5min, resuspended at 6e4/100ul assay medium, and 100ul cell suspension was added to the appropriate wells on a 96-well Black/clear bottom TC plate (BD, 353219) that was coated for 24h with PDL (Sigma, P6407), thoroughly washed with sterile water, and allowed to dry, and stored at 4 degrees Celsius until use.

[0133] Cells were incubated overnight at 37 degrees Celsius before 20ul of a 6X stock of controls or peptides from the CPP library were thawed, mixed, and

added carefully to cells or not to a final concentration of 50uM (17uM control) and incubated for 4h at room temp on a shaker at 400rpm under foil.

[0134] After 4h, 70ul of media was taken off.

[0135] 350ul of PBS (Gibco, 20012-027) was added on the slowest setting of a 12-channel Matrix pipettor and 350ul of medium was pipetted off 4 times.

[0136] Media was completely removed and 100ul 1% NP40 (Fluka, 74385) in 7.5% Sodium Bicarbonate (Gibco, 25080060) was added to every well and incubated for 30min at room temp on a shaker at 400rpm under foil.

[0137] Plate was read on Pherastar at 495nm wavelength.

Example 4: Toxicity assay

[0138] 3e6 Hela cells (ATCC, CCL-2) were thawed into 50ml growth medium (RPMI 1640 (Gibco, 22400-089), 10% HI FBS (JRH, 12106-500M), 200 mM L-Glutamine (Gibco, 25030-081, 1:100), Pen-Strep (Gibco, 15140-122, 1:100)), in a vented T225 and incubated for 3 or 4 days at 37 degrees Celsius.

[0139] After 3 or 4 days, media was aspirated off, cells were washed with 10mls of Accutase® (eBioscience, 00-4555-56), 850ul of Accutase® was added, and flask was incubated for 3min at 37 degrees Celsius.

[0140] After 3min, flask was bumped on the side 20 times, 10ml of assay medium (RPMI 1640 (Gibco, 22400-089), 200 mM L-Glutamine (Gibco, 25030-081, 1:100), Pen-Strep (Gibco, 15140-122, 1:100)) was added to flask, and cells were counted.

[0141] Cells were spun down at 1500rpm for 5min, resuspended at 6e4/100ul assay medium, and 100ul cell suspension was added to the appropriate wells on a 96-well Black/clear bottom TC plate (BD, 353219) that was coated for 24h with PDL

(Sigma, P6407), thoroughly washed with sterile water, and allowed to dry, and stored at 4 degrees Celsius until use.

[0142] Cells were incubated overnight at 37 degrees Celsius before 20ul of a 6X stock of controls and top 48hits from the CPP library were thawed, mixed, and added carefully to cells or not to a final concentration of 50uM (17uM control) and incubated for 24h (1/9/13 and 1/10/13 assays) or 6h (1/15/13 and 1/22/13 assays) at 37 degrees Celsius compared to .2%DMSO in 1M sodium bicarbonate (Gibco, 25080).

[0143] After 24h, 120ul of Cell Titer Glo® (Promega, G7571) was added and plate was incubated on a shaker (400rpm) for 10min.

[0144] After 10min, luminescence was read on Pherastar.

[0145] The specification is most thoroughly understood in light of the teachings of the references cited within the specification. The embodiments within the specification provide an illustration of embodiments of the invention and should not be construed to limit the scope of the invention. The skilled artisan readily recognizes that many other embodiments are encompassed by the invention. The citation of any references herein is not an admission that such references are prior art to the present invention.

[0146] Unless otherwise indicated, all numbers expressing quantities of ingredients, reaction conditions, and so forth used in the specification, including claims, are to be understood as approximations and may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should be construed in light of the number of significant digits and ordinary rounding approaches. The recitation of

series of numbers with differing amounts of significant digits in the specification is not to be construed as implying that numbers with fewer significant digits given have the same precision as numbers with more significant digits given.

[0147] The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims and/or the specification may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.” The use of the term “or” in the claims is used to mean “and/or” unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and “and/or.”

[0148] Unless otherwise indicated, the term “at least” preceding a series of elements is to be understood to refer to every element in the series. 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 claims.

[0149] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described.

[0150] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided

may be different from the actual publication dates which may need to be independently confirmed.

[0151] Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

WHAT IS CLAIMED IS:

1. A compound comprising the formula variable sequence – spacer - linker,
wherein the variable sequence is $\text{Ac-R(O)}_n\text{R}$, wherein $n \geq 7$;
wherein O is a sequence of residues selected from R, X, and Z;
wherein R is L-arginine, X is 3-cis-aminocyclohexane, and Z is
cis-2-aminocyclopentane-1-carbonyl.
2. The compound of any of the preceding claims, wherein the variable sequence
comprises an α -, β -, γ -, or δ -amino acid, or a cycloalkane structure.
3. The compound of claim 1, wherein the variable sequence causes the
compound to be targeted to the nucleus.
4. The compound of any of the preceding claims, wherein the variable sequence
causes the compound to be targeted to the cytosol.
5. The compound of any of the preceding claims, wherein the variable sequence
causes the compound to be targeted to the mitochondria.
6. The compound of any of the preceding claims, wherein the spacer comprises
an aminohexanoic acid (Ahx).
7. The compound of any of the preceding claims, wherein the spacer comprises
(Ahx)B, wherein B is selected from β -alanine or β -glycine.
8. The compound of any of the preceding claims, wherein the linker comprises a
cleavage motif.
9. The compound of claim 8, wherein the cleavage motif is cleavable by a
hydrolytic enzyme.
10. The compound of any of the preceding claims, wherein the linker comprises
FS.
11. The compound of any of the preceding claims, wherein the linker comprises
FSQ.
12. The compound of any of the preceding claims, wherein the linker comprises
FSQK.
13. The compound of any of the preceding claims, wherein the linker comprises
FxyB, x is any amino acid, y is selected from glutamic acid (E), aspartic acid
(D), lysine (K), serine (S), and threonine (T), and B is β -alanine.
14. The compound of claim 13, wherein y is a non-natural analog of glutamic acid.

15. The compound of claim 13, wherein y is a non-natural analog of aspartic acid.
16. The compound of claim 13, wherein y is a non-natural analog of lysine.
17. The compound of claim 13, wherein y is a non-natural analog of serine.
18. The compound of claim 13, wherein y is a non-natural analog of threonine.
19. The compound of any of the preceding claims, wherein the linker comprises FSQG-OH.
20. The compound of any of the preceding claims, wherein n is 7, and the spacer is (Ahx).
21. The compound of any of the preceding claims, wherein the variable sequence, the spacer, and the linker comprise a sequence selected from: SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11.
22. The compound of any of the preceding claims, further comprising a cargo conjugated to the linker.
23. The compound of claim 22, wherein the cargo comprises a chromophore.
24. The compound of claims 22 or 23, wherein the cargo comprises fluorescein.
25. The compound of claim 22, wherein the cargo comprises one or more therapeutic compounds.
26. The compound of claim 22, wherein the cargo comprises a phosphorodiamidate morpholino oligomer (PMO).
27. The compound of claim 26, wherein the cargo further comprises one or more additional PMOs.
28. The compound of any of the preceding claims, wherein the compound is capable of crossing the blood-brain barrier.
29. A composition comprising the compound of any of the preceding claims and a carrier.
30. The composition of claim 29, wherein the carrier comprises saline.
31. The composition of claim 29, wherein the compound is lyophilized.
32. A method of screening for compounds with improved transduction, comprising:
 - incubating cells, a cell medium, a control compound, and one or more compounds according to claim 23;
 - removing the cell medium;

- obtaining a measurement of the amount of transduced one or more compounds as measured by the level of fluorescence of the one or more compounds;
 - obtaining a measurement of the amount of transduced control compound as measured by the level of fluorescence of the control compound; and
 - identifying the one or more compounds wherein the level of fluorescence of the one or more compounds is greater than the level of fluorescence of the control compound.
33. A method of screening for cell penetrating peptide sequences with reduced toxicity, comprising:
- incubating a first sample of cells with one or more compounds according to claim 23;
 - incubating a second sample of cells with a control compound; and
 - identifying the one or more compounds in the first sample where the number of cells in the first sample of cells is greater than 75% of the number of cells in the second sample of cells.
34. A method of screening for compounds with nuclear localization, comprising:
- incubating a sample of cells with a compound according to claim 23;
 - incubating the sample of cells with a nuclear stain;
 - imaging the sample of cells; and
 - identifying the compound wherein the fluorescent signal of the compound substantially overlaps with the location of the nuclear stain.
35. A kit comprising the compound of any of claims 1-28.
36. The kit of claim 36, further comprising instructions for use.

Solulink Bis-arylhydrazone ligation

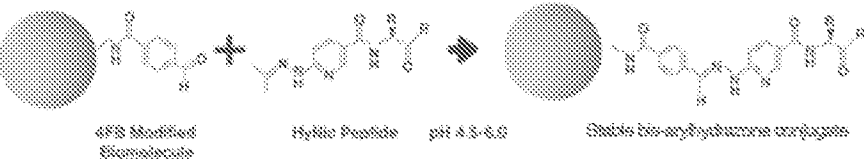


FIG. 1A

Oxime ligation

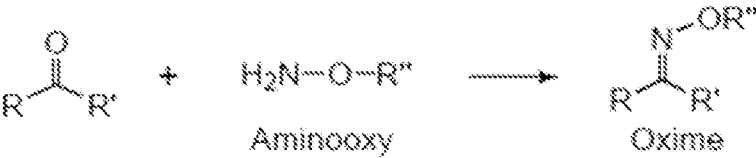
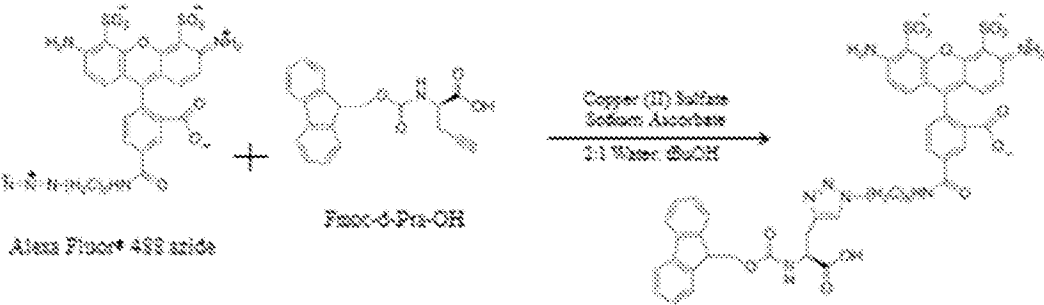


FIG. 1B



Alkyne to Azide Ratio	1:1	1:1.5	1:2	1.5:1	2:1
Percent Yield	80%	85%	85%	95%	97%

FIG. 1C

Peptide #	Sequence	SEQ. ID No.
1370	Ac-(RXR) ₂ XFSEK(Fluor)G-OH	54
1371	Ac-(RXR) ₂ XFSQK(Fluor)G-OH	55
1372	Ac-(RXR) ₃ XFSEK(Fluor)G-OH	56
1373	Ac-(RXR) ₃ XFSQK(Fluor)G-OH	57
1374	GQNDTSQTSSPSADK(Fluor)-OH	58
1226-F	Ac-(RXR) ₄ XFK(Fluor)G-OH	59
Fluor	5'-Carboxyfluorescein	
1069	Fluor-EDTAVYYCTPVDFWG-OH	60

FIG. 2

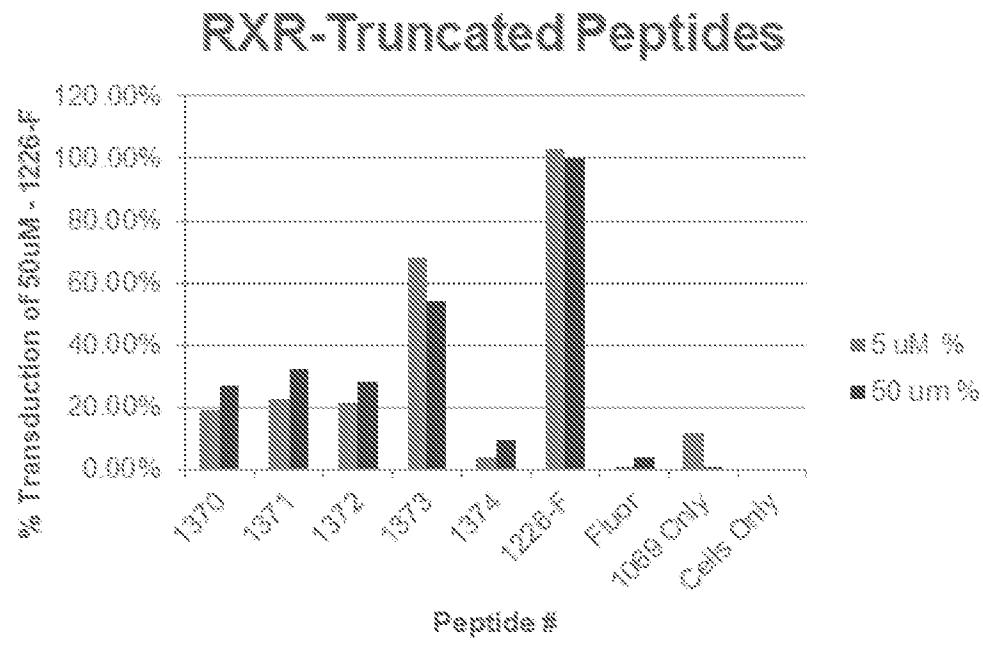


FIG. 3A

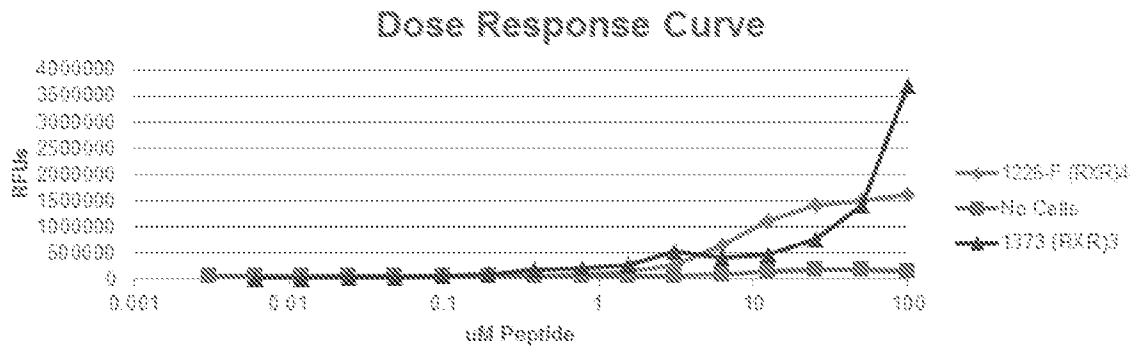


FIG. 3B

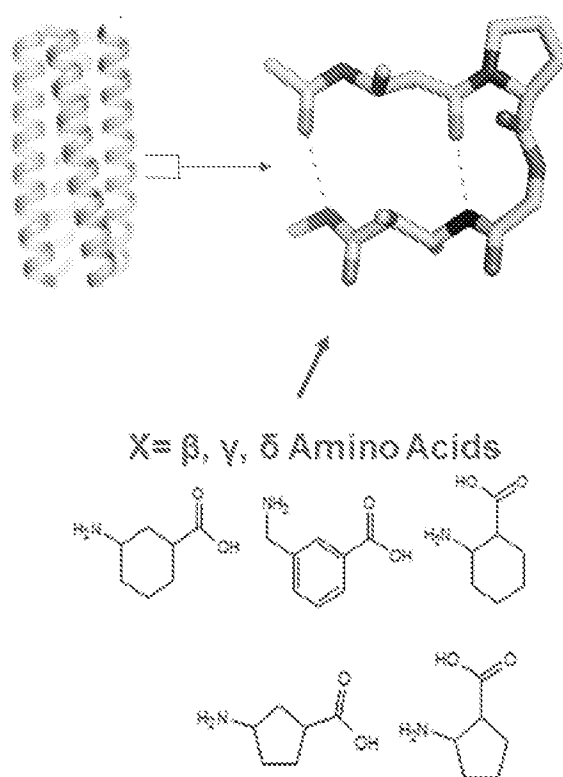


FIG. 4A

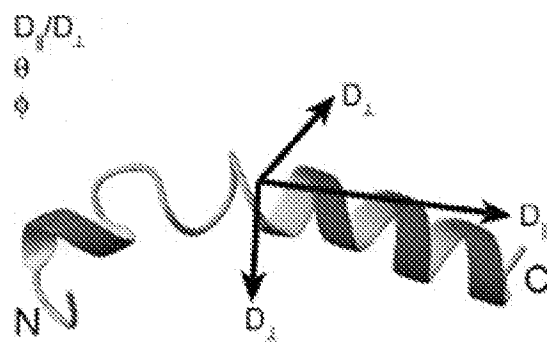


FIG. 4B

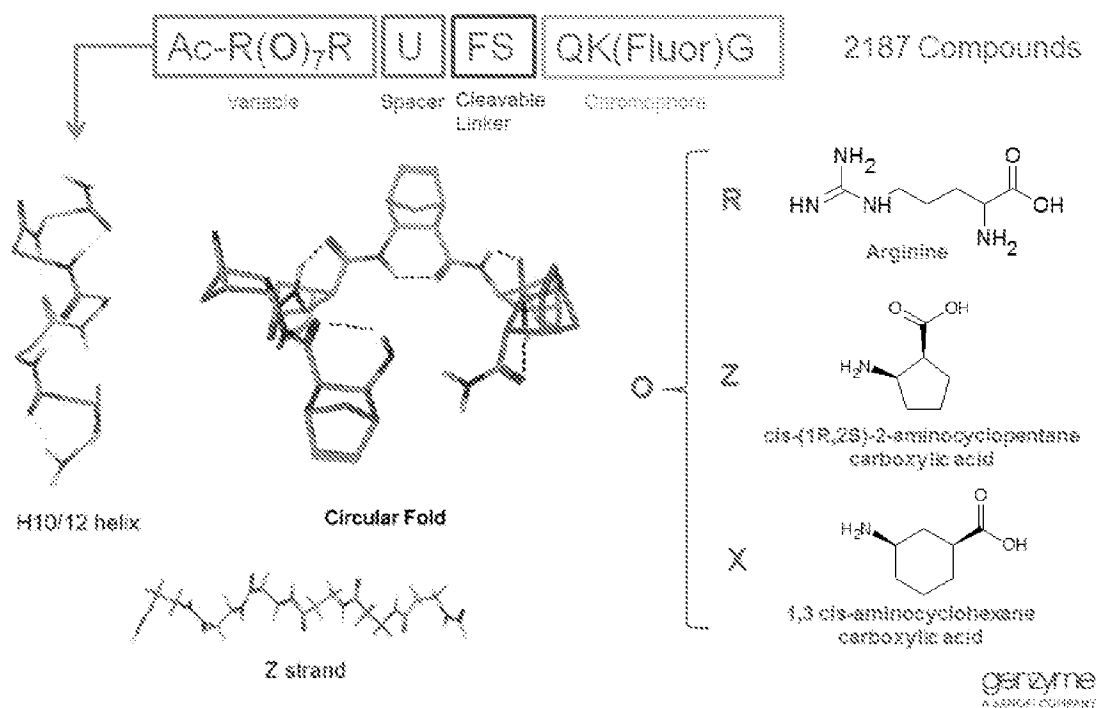
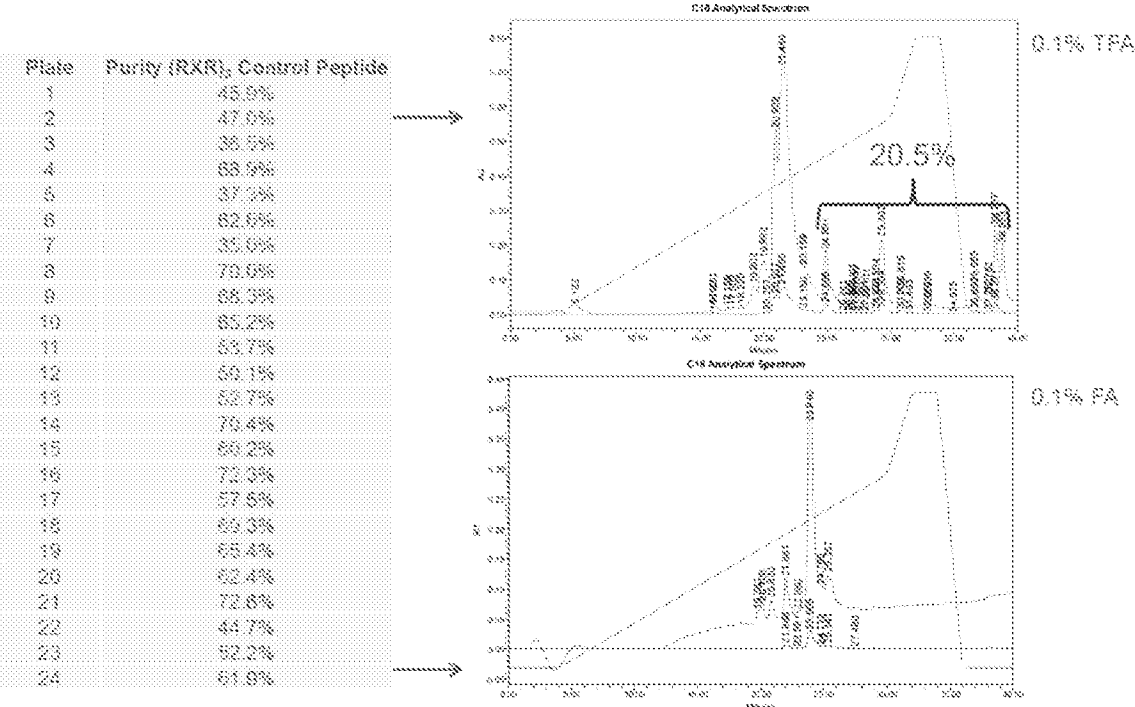


FIG. 5



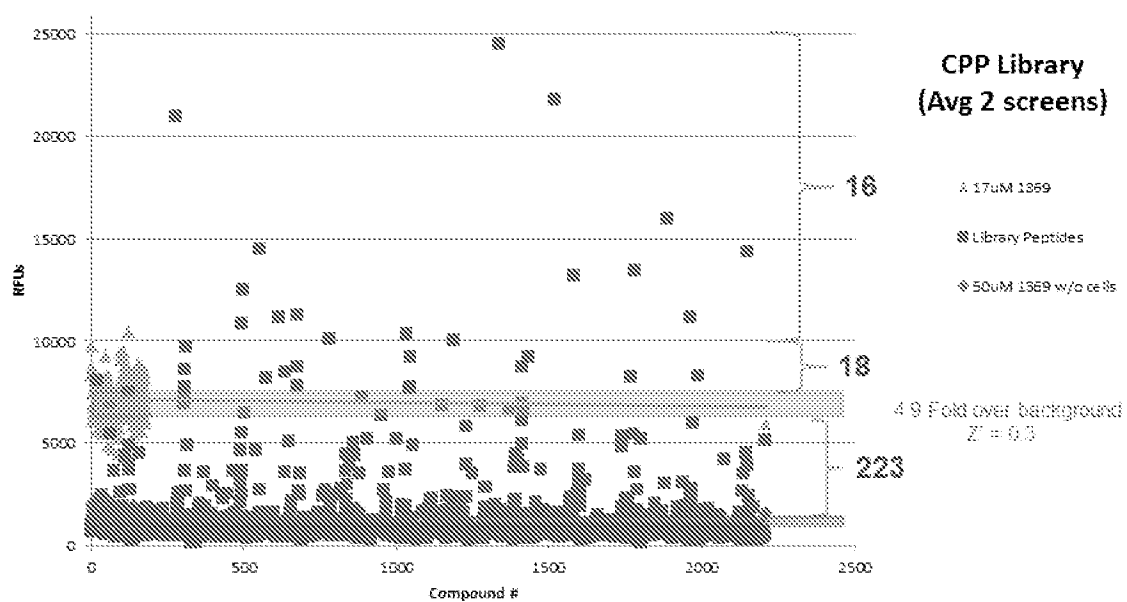


FIG. 7

Address	Sequence	MV	Run 1 RFUs	Run 2 RFUs	Average RFUs	Seq. ID No.
9H8	Ac-RXXXXXRXR(Ahx)FSQK(Fluor)G-OH	2294 34	22904	26198	24551	61
9H9	Ac-RZXXXXXRXR(Ahx)FSQK(Fluor)G-OH	2280.34	21258	22331	21795	62
9H11	Ac-RXZXXXXXR(Ahx)FSQK(Fluor)G-OH	2280 34	9872	22051	15962	63
1A2	Ac-RRRRRRRRR(Ahx)FSQK(Fluor)G-OH	2484.66	20554	21448	21001	64
12D12	Ac-RZRXXXXXR(Ahx)FSQK(Fluor)G-OH	2280.34	16534	12144	14339	65
13D3	Ac-RXZXXZXR(Ahx)FSQK(Fluor)G-OH	2234.62	10932	14046	12489	66
12D10	Ac-RRRXXXXXR(Ahx)FSQK(Fluor)G-OH	2326.06	13473	13450	13462	67
2C4	Ac-RXXXXRRR(Ahx)FSQK(Fluor)G-OH	2326.06	15603	13349	14476	68
4G9	Ac-RXXXXRRR(Ahx)FSQK(Fluor)G-OH	2294.34	8930	17423	13177	69
11F4	Ac-RXXXXRRR(Ahx)FSQK(Fluor)G-OH	2294.34	11346	10909	11128	70
9F5	Ac-RXXRXR(Ahx)FSQK(Fluor)G-OH	2326.06	14172	6000	10086	71
12D11	Ac-RXRXXXXXR(Ahx)FSQK(Fluor)G-OH	2294.34	6526	15753	11140	72
20B7	Ac-RXXXXXZR(Ahx)FSQK(Fluor)G-OH	2248 62	10206	9831	10019	73
20C4	Ac-RXXZXXZXR(Ahx)FSQK(Fluor)G-OH	2234.62	10209	6721	8465	74
5D4	Ac-RXXRXZRR(Ahx)FSQK(Fluor)G-OH	2312 06	4950	11367	8159	75
9H7	Ac-RRXXXXXR(Ahx)FSQK(Fluor)G-OH	2326.06	7103	6666	6885	76
5B1	Ac-RXXZXXRR(Ahx)FSQK(Fluor)G-OH	2280.34	5140	10929	8035	77
4H6	Ac-RXXZXXRR(Ahx)FSQK(Fluor)G-OH	2280.34	6784	13892	10338	78
12D4	Ac-RRRXXXXXR(Ahx)FSQK(Fluor)G-OH	2326.06	7960	14637	11299	79
15A8	Ac-RXXXXXZR(Ahx)FSQK(Fluor)G-OH	2248.62	13239	5201	9220	80
12D8	Ac-RXZRRXXXXXR(Ahx)FSQK(Fluor)G-OH	2280.34	5560	11949	8755	81
12E3	Ac-RZXXXXXR(Ahx)FSQK(Fluor)G-OH	2248 62	5922	15789	10856	82
12H2	Ac-RXXZXXXR(Ahx)FSQK(Fluor)G-OH	2234 62	10543	8920	9732	83
4G10	Ac-RZXXXXRR(Ahx)FSQK(Fluor)G-OH	2280 34	8161	8332	8247	84
15H5	Ac-RXXXXZXR(Ahx)FSQK(Fluor)G-OH	2234.62	9278	5229	7254	85
11F1	Ac-RXRXXR(Ahx)FSQK(Fluor)G-OH	2326.06	4590	9302	6946	86
21C7	Ac-RRXXZZXR(Ahx)FSQK(Fluor)G-OH	2252.34	7936	5730	6833	87
12E2	Ac-RXXXXXXXR(Ahx)FSQK(Fluor)G-OH	2262.62	10724	6494	8609	88
12E4	Ac-RRZXXXXXR(Ahx)FSQK(Fluor)G-OH	2280.34	8434	8960	8697	89
12G4	Ac-RRXXZXXXR(Ahx)FSQK(Fluor)G-OH	2280.34	6102	9543	7823	90
21A5	Ac-RXXXXZXZR(Ahx)FSQK(Fluor)G-OH	2234 62	6427	4006	5217	91

FIG. 8

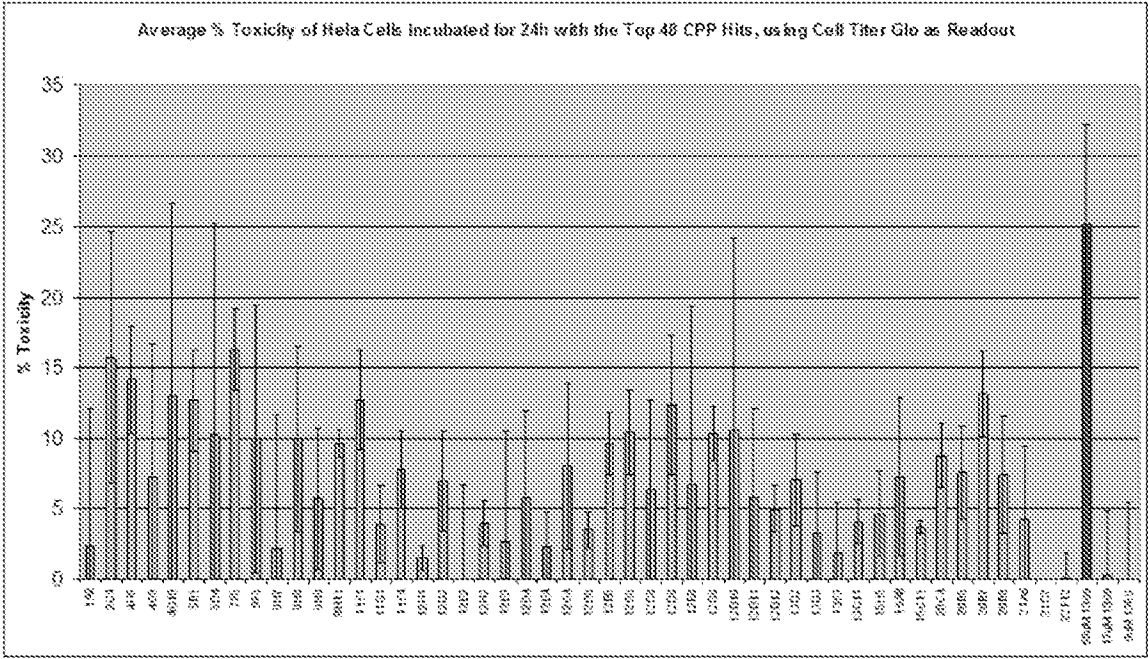


FIG. 9

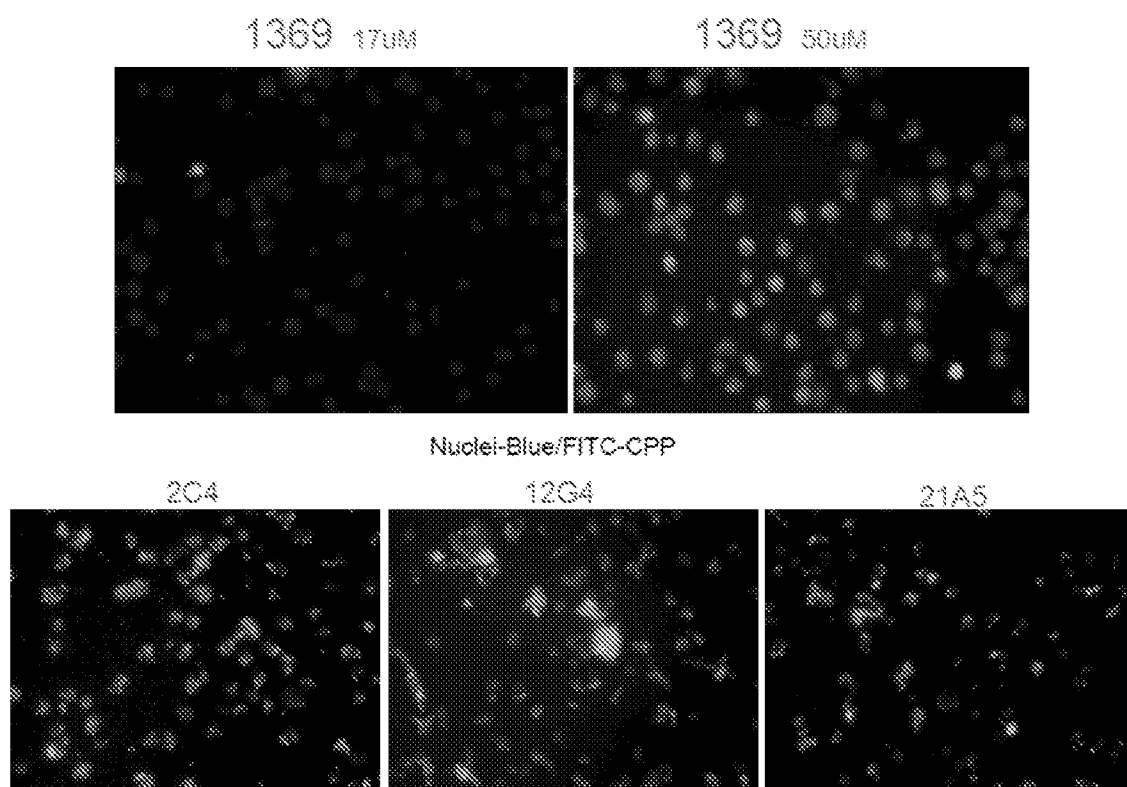


FIG. 10

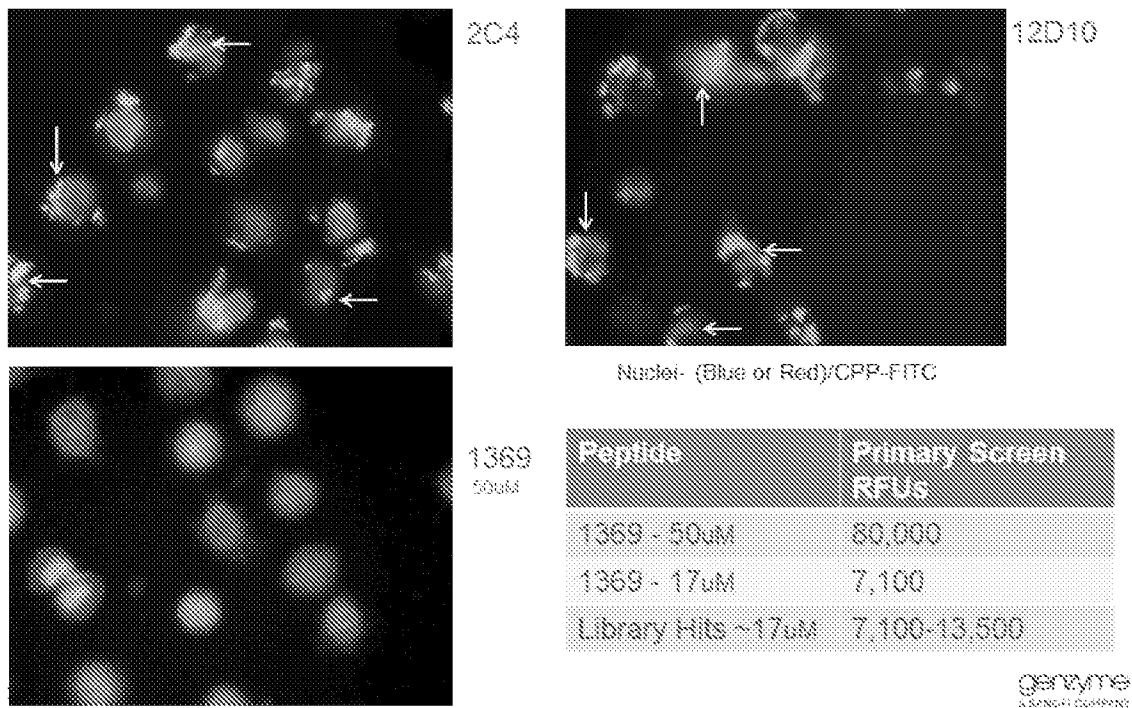


FIG. 11

Hit	Sequence	SEQ. ID No.
2C4	Ac-RXXXXXXRRR(Ahx)FSQG-OH	11
49G	Ac-RXXXXXXRR(Ahx)FSQG-OH	12
9F5	Ac-RXXXRXRXR(Ahx)FSQG-OH	13
12G4	Ac-RRXXZXXXR(Ahx)FSQG-OH	14
12D10	Ac-RRRXXXXXR(Ahx)FSQG-OH	15
12D11	Ac-RXRXXXXXR(Ahx)FSQG-OH	16
12E4	Ac-RRZXXXXXR(Ahx)FSQG-OH	17
21A5	Ac-RXXXXZXZR(Ahx)FSQG-OH	18
11G1	Ac-RXXZXRXRXR(Ahx)FSQG-OH	19
12D4	Ac-RRXRXXXXXR(Ahx)FSQG-OH	20
13D2	Ac-RRZXXZXXR(Ahx)FSQG-OH	21

FIG. 12A

Hit	Sequence	SEQ. ID No.
11G1	Ac-RXXZXRXRXR(Ahx)FSQG-OH	19
12D4	Ac-RRXRXXXXXR(Ahx)FSQG-OH	20
13D2	Ac-RRZXXZXXR(Ahx)FSQG-OH	21

FIG. 12B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/032143

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07C229/48 C12N15/113 C07K7/04 C07K19/00 A61K47/48
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07C C12N C07K A61K

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

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

EPO-Internal, BEILSTEIN Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GAMMON S T ET AL: "Quantitative analysis of permeation peptide complexes labeled with technetium-99m: Chiral and sequence-specific effects on net cell uptake", BIOCONJUGATE CHEMISTRY, ACS, WASHINGTON, DC, US, vol. 14, no. 2, 1 January 2003 (2003-01-01), pages 368-376, XP002347554, ISSN: 1043-1802, DOI: 10.1021/BC0256291 table 1; compound 32	1,2,4,6, 20,22, 25,29, 30,35,36
X	WO 2014/052276 A1 (GENZYME CORP [US]) 3 April 2014 (2014-04-03) claims ----- -/-	1,2, 25-27, 29,30, 35,36



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

8 September 2015

Date of mailing of the international search report

16/09/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Vogt, Titus

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2015/032143

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PETER JÄRVER ET AL: "Peptide-mediated Cell and In Vivo Delivery of Antisense Oligonucleotides and siRNA", MOLECULAR THERAPY - NUCLEIC ACIDS, vol. 1, no. 6, 1 June 2012 (2012-06-01), page e27, XP55210945, ISSN: 2162-2531, DOI: 10.1038/mtna.2012.18 the whole document -----	1-36

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/032143

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014052276 A1	03-04-2014	AU 2013323820 A1	16-04-2015
		CA 2886017 A1	03-04-2014
		CN 104837997 A	12-08-2015
		EP 2900821 A1	05-08-2015
		KR 20150060738 A	03-06-2015
		TW 201424752 A	01-07-2014
		US 2015238627 A1	27-08-2015
		WO 2014052276 A1	03-04-2014
