

COMPOSITIONS AND METHODS FOR MODULATING TISSUE DISTRIBUTION OF INTRACELLULAR THERAPEUTICS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. provisional patent application No. 63/186,664, which was filed on May 10, 2021, U.S. provisional patent application No. 63/214,085, which was filed on June 23, 2021, U.S. provisional patent application No. 63/239,671, which was filed on September 1, 2021, U.S. provisional patent application No. 63/290,960, which was filed on December 17, 2021, U.S. provisional patent application No. 63/298,565, which was filed on January 11, 2022, U.S. provisional patent application No. 63/268,577, which was filed on February 25, 2022, U.S. provisional patent application No. 63/362,295, which was filed on March 31, 2022, the disclosures of each of which are hereby incorporated by reference in their entireties.

BACKGROUND

[0002] Biologics such as proteins, peptides and nucleic acids are promising approaches for treatment of a wide variety of diseases and disorders. In particular, the therapeutic applications of oligonucleotides such as antisense compounds are extremely broad, since these compounds can be synthesized with any nucleotide sequence directed against virtually any target gene or genomic segments. However, the plasma membrane presents a major challenge in both drug discovery and therapy, especially for therapeutic agents such as biologics. For example, a major problem for the use of oligonucleotide-based biologics in therapy is their limited ability to gain access to the intracellular compartment when administered systemically. Intracellular delivery of oligonucleotide compounds can be facilitated by use of carrier systems such as polymers, cationic liposomes or by chemical modification of the construct, for example by the covalent attachment of cholesterol molecules. However, intracellular delivery efficiency is still low.

[0003] One potential strategy to subvert the membrane barrier and deliver therapeutic agents such as biologics into cells is to attach them to "cell-penetrating peptides (CPPs)". CPPs that enter cells via endocytosis must exit from endocytic vesicles in order to reach the cytosol. Unfortunately, the endosomal membrane has proven to be a significant barrier towards cytoplasmic delivery by these CPPs; often a negligible fraction of the peptides escapes into the cell interior (see e.g., El-Sayed, A et al. AAPS J., 2009, 11, 13-22; Varkouhi, A K et al. J. Controlled Release, 2011, 151, 220-228;

Appelbaum, J S et al. Chem. Biol., 2012, 19, 819-830). Cyclic CPPs (cCPPs) with improved properties have been described for use in intracellular delivery of a cargo moiety (U.S. Patent Publication Nos. 2017/0190743 and 2017/0355730).

5 [0004] There is still an unmet need for effective compositions and methods for intracellular delivery of a therapeutic agent, particularly in a manner that allows for modulation of the tissue distribution and/or retention of the agent. The compositions and methods disclosed herein address these and other needs.

SUMMARY

10 [0005] Compositions and methods for modulating tissue distribution and/or retention of intracellular therapeutic agents are described herein. Compounds comprising a cell penetrating peptide (CPP) linked to a therapeutic moiety (TM) have been found to have altered tissue distribution and/or retention when the compound further comprises an exocyclic peptide (EP) as described herein. The EP typically is a lysine-containing peptide. EPs have been identified that
15 previously were known in the art as “nuclear localization signals” (NLS), such as the nuclear localization sequence of the SV40 virus large T-antigen, the minimal functional unit of which is the seven amino acid sequence PKKKRKV. Inclusion of an EP in the CPP-TM compound can, for example, alter the level of expression, activity or function of the TM in different tissues, such as different types of muscle tissue or different types of central nervous system tissue (see e.g.,
20 Examples 5 and 7). In one embodiment, the compound is administered intrathecally and the tissue distribution and/or retention of the compound in tissues of the CNS is modulated.

[0006] In embodiments, a compound is provided that comprises:

- (a) a cell penetrating peptide (CPP);
- (b) a therapeutic moiety (TM); and
- 25 (c) an exocyclic peptide (EP), wherein tissue distribution or retention of the compound is modulated as compared to a compound comprising the CPP and TM but lacking the EP.

[0007] In one embodiment, the EP is conjugated to the CPP. In one embodiment, the EP is conjugated to the TM. In one embodiment, the CPP is a cyclic cell penetrating peptide (cCPP). In various embodiments, the therapeutic moiety is a protein, a polypeptide, an oligonucleotide or a

small molecule. In one embodiment, the oligonucleotide is an antisense compound (AC). In one embodiment, the oligonucleotide is other than an antisense compound (AC).

[0008] In embodiments, the therapeutic compound comprises an AC that modulates splicing of exon 2, 8, 11, 17, 19, 23, 29, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 51, 52, 53, 55, and 59 of DMD.

5 In embodiments, the compound comprises an AC that modulates splicing of exon 2, 8, 11, 23 43, 44, 45, 50, 51, 53, and 55 of DMD. In embodiments, the compound comprises an AC that modulates splicing of exon 2, 23, 44, or 51 of DMD.

[0009] In embodiments, the compound comprises an AC that modulates splicing of exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7a, and exon 7b of CD33.

10 [0010] In embodiments, the compound is selected from the group consisting of EEV-PMO-MDX23-1,2,3, EEV-PMO-CD33-1 and the compounds shown in Table C of Example 5, having the structures shown herein.

[0011] In embodiments, a pharmaceutical composition is provided that includes the compound described herein and a pharmaceutically acceptable carrier.

15 [0012] In embodiments, a cell is provided that includes a compound described herein.

[0013] The disclosure also relates to a method of modulating tissue distribution or retention of a therapeutic agent in a subject in need thereof, comprising administering a compound of the disclosure. In embodiments, the compound is administered to the subject intrathecally and the compound modulates tissue distribution or retention of the therapeutic agent in tissues of the
20 central nervous system (CNS). In embodiments, the compound modulates tissue distribution or retention of the therapeutic agent in muscle tissue.

[0014] The disclosure also pertains to a method of treating a disease or disorder in a subject in need thereof, comprising administering a compound of the disclosure. In embodiments, the therapeutic agent is an antisense compound (AC) and administration of the compound modulates
25 splicing or expression of a target gene, degrades mRNA, stabilizes mRNA, or sterically blocks mRNA. In embodiments, administration of the compound modulates splicing of the target pre-mRNA.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] **FIG. 1** shows modified nucleotides used in antisense oligonucleotides described herein.

[0016] **FIGS. 2A-2D** provide structures for morpholino subunit monomers used in synthesizing phosphorodiamidate-linked morpholino oligomers. **FIG. 2A** provides the structure for adenine morpholino monomer. **FIG. 2B** provides the structure for cytosine morpholino monomer. **FIG. 2C** provides the structure for guanine morpholino monomer. **FIG. 2D** provides the structure for thymine morpholino monomer.

[0017] **FIGS. 3A-3D** illustrate conjugation chemistries for connecting AC to a cyclic cell penetrating peptide. **FIG. 3A** shows the amide bond formation between peptides with carboxylic acid group or with TFP activated ester and primary amine residues at the 5' end of AC. **FIG. 3B** shows the conjugation of secondary amine or primary amine modified AC at 3' and peptide-TFP ester through amide bond formation. **FIG. 3C** shows the conjugation of peptide-azide to the 5' cyclooctyne modified AC via copper-free azide-alkyne cycloaddition. **FIG. 3D** demonstrates another exemplary conjugation between 3' modified cyclooctyne ACs or 3' modified azide ACs and CPP containing linker-azide or linker-alkyne/cyclooctyne moiety, via a copper-free azide-alkyne cycloaddition or copper catalyzed azide-alkyne cycloaddition, respectively (click reaction).

[0018] **FIG. 4** shows the conjugation chemistry for connecting AC and CPP with an additional linker modality containing a polyethylene glycol (PEG) moiety.

[0019] **FIG. 5A-5C** show the structures of three conjugates used in the examples. **FIG. 5A** shows the structure of an oligonucleotide conjugate that includes a PMO oligonucleotide and a bonding group. **FIG. 5B** shows a structure of a cell penetrating peptide-oligonucleotide conjugate that includes a PMO oligonucleotide, a bonding group, a cyclic cell penetrating peptide and a nuclear localization signal (PMO-NLS-EEV). **FIG. 5C** shows a structure of a conjugate that includes a bonding group, a polyethylene glycol (PEG) linker and a cyclic cell penetrating peptide - CPP12-PEG4-dk(LSR).

[0020] **FIG. 6** shows representative structures of cyclic peptide-linker conjugates of the present disclosure.

[0021] **FIG. 7** shows a further representative structure of a cyclic peptide-linker conjugate of the present disclosure.

[0022] FIG. 8 provides general schematics for an antigen-degradation construct. The hashed boxes indicate optional linker sequences. The orientation of the constructs shown in these figures is not limiting. Various other general conformations of these constructs are contemplated herein. For example, the CPP can be located at any suitable location in the construct (N-terminal as shown, C-terminal, or at an internal location of the construct).

[0023] FIG. 9 illustrates direct action of degradation moieties leading to target antigen degradation.

[0024] FIG. 10 illustrates indirect action of degradation moieties leading to target antigen degradation.

[0025] FIG. 11 shows the conjugation of an exemplary CPP and an exemplary AC, as described in Example 3.

[0026] FIGS. 12A-12D show the level of exon 23 correction in muscle tissue of MDX mice five days after treatment with the indicated compounds. Results are shown for diaphragm (FIG. 12A), heart (FIG. 12B), tibialis anterior (FIG. 12C) and triceps (FIG. 12D).

[0027] FIGS. 13A-13C show the level of dystrophin expression in muscle tissue of MDX mice five days after treatment with the indicated compounds. Results are shown for diaphragm (FIG. 13A), heart (FIG. 13B) and tibialis anterior (FIG. 13C).

[0028] FIGS. 14A-14B are schematic illustrations of the synthesis of the compounds EEV-PMO-CD33-1 (FIG. 14A) and EEV-PMO-CD33-2 (FIG. 14B).

[0029] FIG. 15 shows results from experiments conducted by intrathecal (IT) injection of PMO-CD33, EEV-PMO-CD33-2, and EEV-PMO-CD33-1 to rats, followed by analysis of expression in the cerebellum, cortex, hippocampus and olfactory bulb of the rat brains.

[0030] FIGS. 16A-16C show results from experiments conducted by intrathecal (IT) injection of PMO-CD33, EEV-PMO-CD33-2, and EEV-PMO-CD33-1 to rats, followed by analysis of expression in the spinal cord, DRG and CSF of the rat brains.

DETAILED DESCRIPTION

Compounds

Endosomal Escape Vehicles (EEVs)

[0031] An endosomal escape vehicle (EEV) is provided herein that can be used to transport a cargo across a cellular membrane, for example, to deliver the cargo to the cytosol or nucleus of a cell. Cargo can include a macromolecule, for example, a peptide or oligonucleotide, or a small molecule. The EEV can comprise a cell penetrating peptide (CPP), for example, a cyclic cell penetrating peptide (cCPP), which is conjugated to an exocyclic peptide (EP). The EP can be referred to interchangeably as a modulatory peptide (MP). The EP can comprise a sequence of a nuclear localization signal (NLS). The EP can be coupled to the cargo. The EP can be coupled to the cCPP. The EP can be coupled to the cargo and the cCPP. Coupling between the EP, cargo, cCPP, or combinations thereof, may be non-covalent or covalent. The EP can be attached through a peptide bond to the N-terminus of the cCPP. The EP can be attached through a peptide bond to the C-terminus of the cCPP. The EP can be attached to the cCPP through a side chain of an amino acid in the cCPP. The EP can be attached to the cCPP through a side chain of a lysine which can be conjugated to the side chain of a glutamine in the cCPP. The EP can be conjugated to the 5' or 3' end of an oligonucleotide cargo. The EP can be coupled to a linker. The exocyclic peptide can be conjugated to an amino group of the linker. The EP can be coupled to a linker via the C-terminus of an EP and a cCPP through a side chain on the cCPP and/or EP. For example, an EP may comprise a terminal lysine which can then be coupled to a cCPP containing a glutamine through an amide bond. When the EP contains a terminal lysine, and the side chain of the lysine can be used to attach the cCPP, the C- or N-terminus may be attached to a linker on the cargo.

Exocyclic Peptides

[0032] The exocyclic peptide (EP) can comprise from 2 to 10 amino acid residues e.g., 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues, inclusive of all ranges and values therebetween. The EP can comprise 6 to 9 amino acid residues. The EP can comprise from 4 to 8 amino acid residues.

[0033] Each amino acid in the exocyclic peptide may be a natural or non-natural amino acid. The term "non-natural amino acid" refers to an organic compound that is a congener of a natural amino acid in that it has a structure similar to a natural amino acid so that it mimics the structure and

reactivity of a natural amino acid. The non-natural amino acid can be a modified amino acid, and/or amino acid analog, that is not one of the 20 common naturally occurring amino acids or the rare natural amino acids selenocysteine or pyrrolysine. Non-natural amino acids can also be the D-isomer of the natural amino acids. Examples of suitable amino acids include, but are not limited to, alanine, allosoleucine, arginine, citrulline, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, naphthylalanine, phenylalanine, proline, pyroglutamic acid, serine, threonine, tryptophan, tyrosine, valine, a derivative thereof, or combinations thereof. These, and others amino acids, are listed in the **Table 1** along with their abbreviations used herein. For example, the amino acids can be A, G, P, K, R, V, F, H, Nal, or citrulline.

[0034] The EP can comprise at least one positively charged amino acid residue, e.g., at least one lysine residue and/or at least one amine acid residue comprising a side chain comprising a guanidine group, or a protonated form thereof. The EP can comprise 1 or 2 amino acid residues comprising a side chain comprising a guanidine group, or a protonated form thereof. The amino acid residue comprising a side chain comprising a guanidine group can be an arginine residue. Protonated forms can mean salt thereof throughout the disclosure.

[0035] The EP can comprise at least two, at least three or at least four or more lysine residues. The EP can comprise 2, 3, or 4 lysine residues. The amino group on the side chain of each lysine residue can be substituted with a protecting group, including, for example, trifluoroacetyl ($-\text{COCF}_3$), allyloxycarbonyl (Alloc), 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde), or (4,4-dimethyl-2,6-dioxocyclohex-1-ylidene-3)-methylbutyl (ivDde) group. The amino group on the side chain of each lysine residue can be substituted with a trifluoroacetyl ($-\text{COCF}_3$) group. The protecting group can be included to enable amide conjugation. The protecting group can be removed after the EP is conjugated to a cCPP.

[0036] The EP can comprise at least 2 amino acid residues with a hydrophobic side chain. The amino acid residue with a hydrophobic side chain can be selected from valine, proline, alanine, leucine, isoleucine, and methionine. The amino acid residue with a hydrophobic side chain can be valine or proline.

[0037] The EP can comprise at least one positively charged amino acid residue, e.g., at least one lysine residue and/or at least one arginine residue. The EP can comprise at least two, at least three or at least four or more lysine residues and/or arginine residues.

[0038] The EP can comprise KK, KR, RR, HH, HK, HR, RH, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKH, KHK, HKK, HRR, HRH, HHR, HBH, HHH, HHHH, KHKK, KKKH, KKKH, KKKH, HKHK, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, HBHBH, HBKBH, RRRRR, KKKKK, KKKRK, RKKKK, KRKKK, KKKRK, KKKRK, KBKKB, RKKKKG, KRKKKG, KKKKKG, KKKKRG, RKKKKB, KRKKKB, KKKKKB, KKKKRB, KKKRKV, RRRRRR, HHHHHH, RHRHRH, HRHRHR, KRKRKR, RKRKRK, RBRBRB, KBKKB, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV or PKKKRKG, wherein B is beta-alanine. The amino acids in the EP can have D or L stereochemistry.

[0039] The EP can comprise KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKKB, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV or PKKKRKG. The EP can comprise PKKKRKV, RR, RRR, RHR, RBR, RBRBR, RBHBR, or HBRBH, wherein B is beta-alanine. The amino acids in the EP can have D or L stereochemistry.

[0040] The EP can consist of KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKKB, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV or PKKKRKG. The EP can consist of PKKKRKV, RR, RRR, RHR, RBR, RBRBR, RBHBR, or HBRBH, wherein B is beta-alanine. The amino acids in the EP can have D or L stereochemistry.

[0041] The EP can comprise an amino acid sequence identified in the art as a nuclear localization sequence (NLS). The EP can consist of an amino acid sequence identified in the art as a nuclear localization sequence (NLS). The EP can comprise an NLS comprising the amino acid sequence PKKKRKV. The EP can consist of an NLS comprising the amino acid sequence PKKKRKV. The EP can comprise an NLS comprising an amino acid sequence selected from NLSKRPAAIKKAGQAKKKK, PAAKRVKLD, RQRRNELKRSF, RMRKFKNKGKDTAELRRRRVEVSVELR, KAKKDEQILKRRNV, VSRKRPRP, PPKKARED, PPKKKPL, SALIKKKKKMAP, DRLRR, PKQKKRK, RKLKKIKKL,

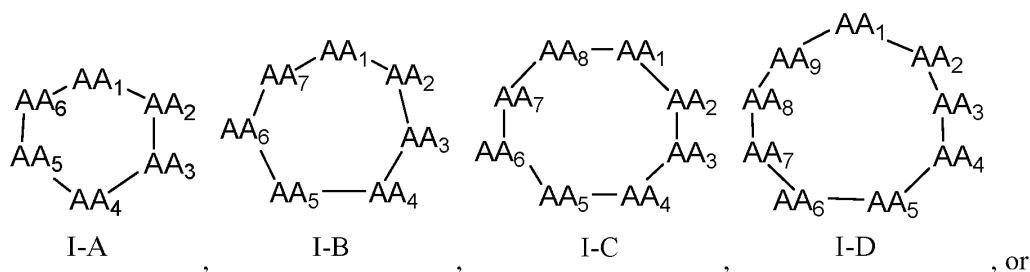
REKKKFLKRR, KRKGDEV DGVDEVAKKKSKK and RKCLQAGMNLEARKTKK. The EP can consist of an NLS comprising an amino acid sequence selected from NLSKRPAAIKKAGQAKKKK, PAAKRVKLD, RQRRNELKRSF, RMRKFKNKGKDTAELRRRRVEVSVELR, KAKKDEQILKRRNV, VSRKRPRP, 5 PPKKARED, PQPKKKPL, SALIKKKKKMAP, DRLRR, PKQKKRK, RKLKKKIKKL, REKKKFLKRR, KRKGDEV DGVDEVAKKKSKK and RKCLQAGMNLEARKTKK

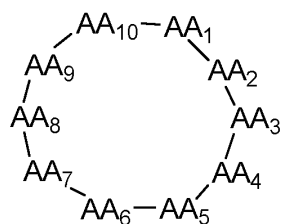
[0042] All exocyclic sequences can also contain an N-terminal acetyl group. Hence, for example, the EP can have the structure: Ac-PKKKKRKV.

Cell Penetrating Peptides (CPP)

10 [0043] The cell penetrating peptide (CPP) can comprise 6 to 20 amino acid residues. The cell penetrating peptide can be a cyclic cell penetrating peptide (cCPP). The cCPP is capable of penetrating a cell membrane. An exocyclic peptide (EP) can be conjugated to the cCPP, and the resulting construct can be referred to as an endosomal escape vehicle (EEV). The cCPP can direct a cargo (e.g., a therapeutic moiety (TM) such as an oligonucleotide, peptide or small molecule) to 15 penetrate the membrane of a cell. The cCPP can deliver the cargo to the cytosol of the cell. The cCPP can deliver the cargo to a cellular location where a target (e.g., pre-mRNA) is located. To conjugate the cCPP to a cargo (e.g., peptide, oligonucleotide, or small molecule), at least one bond or lone pair of electrons on the cCPP can be replaced.

[0044] The total number of amino acid residues in the cCPP is in the range of from 6 to 20 amino acid residues, e.g., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid residues, 20 inclusive of all ranges and subranges therebetween. The cCPP can comprise 6 to 13 amino acid residues. The cCPP disclosed herein can comprise 6 to 10 amino acids. By way of example, cCPP comprising 6-10 amino acid residues can have a structure according to any of Formula I-A to I-E:





I-E

, wherein AA₁, AA₂, AA₃, AA₄, AA₅, AA₆, AA₇, AA₈, AA₉, and AA₁₀

are amino acid residues.

[0045] The cCPP can comprise 6 to 8 amino acids. The cCPP can comprise 8 amino acids.

[0046] Each amino acid in the cCPP may be a natural or non-natural amino acid. The term “non-natural amino acid” refers to an organic compound that is a congener of a natural amino acid in that it has a structure similar to a natural amino acid so that it mimics the structure and reactivity of a natural amino acid. The non-natural amino acid can be a modified amino acid, and/or amino acid analog, that is not one of the 20 common naturally occurring amino acids or the rare natural amino acids selenocysteine or pyrrolysine. Non-natural amino acids can also be a D-isomer of a natural amino acid. Examples of suitable amino acids include, but are not limited to, alanine, allosoleucine, arginine, citrulline, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, naphthylalanine, phenylalanine, proline, pyroglutamic acid, serine, threonine, tryptophan, tyrosine, valine, a derivative thereof, or combinations thereof. These, and others amino acids, are listed in the **Table 1** along with their abbreviations used herein.

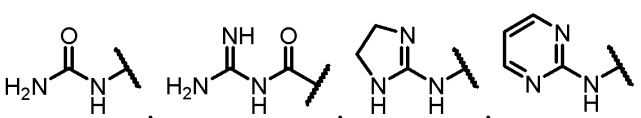
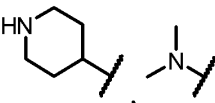
Table 1. Amino Acid Abbreviations

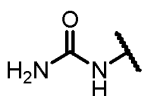
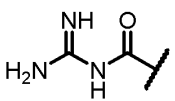
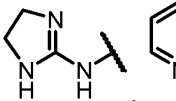
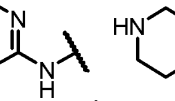
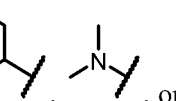
Amino Acid	Abbreviations* L-amino acid	Abbreviations* D-amino acid
2-[2-[2-aminoethoxy]ethoxy]acetic acid	AEEA, miniPEG, PEG2	NA
Alanine	Ala (A)	ala (a)
Allo-isoleucine	Aile	Aile
Arginine	Arg (R)	arg (r)
Asparagine	Asn (N)	asn (n)
aspartic acid	Asp (D)	asp (d)
Cysteine	Cys (C)	cys (c)
Citrulline	Cit	Cit
Cyclohexylalanine	Cha	cha
2,3-diaminopropionic acid	Dap	dap
4-fluorophenylalanine	Fpa (Σ)	pfa
glutamic acid	Glu (E)	glu (e)
glutamine	Gln (Q)	gln (q)

Amino Acid	Abbreviations* L-amino acid	Abbreviations* D-amino acid
glycine	Gly (G)	gly (g)
histidine	His (H)	his (h)
Homoproline (aka pipecolic acid)	Pip (Θ)	pip (θ)
isoleucine	Ile (I)	ile (i)
leucine	Leu (L)	leu (l)
lysine	Lys (K)	lys (k)
methionine	Met (M)	met (m)
3-(2-naphthyl)-alanine	Nal (Φ)	nal (φ)
3-(1-naphthyl)-alanine	1-Nal	1-nal
norleucine	Nle (Ω)	nle
phenylalanine	Phe (F)	phe (f)
phenylglycine	Phg (Ψ)	phg
4-(phosphonodifluoromethyl)phenylalanine	F ₂ Pmp (Λ)	f ₂ pmp
proline	Pro (P)	pro (p)
sarcosine	Sar (Ξ)	sar
selenocysteine	Sec (U)	sec (u)
serine	Ser (S)	ser (s)
threonine	Thr (T)	thr (y)
tyrosine	Tyr (Y)	tyr (y)
tryptophan	Trp (W)	trp (w)
valine	Val (V)	val (v)
Tert-butyl-alanine	Tle	tle
Penicillamine	Pen	Pen
Homoarginine	HomoArg	homoarg
Nicotinyl-lysine	Lys(NIC)	lys(NIC)
Trifluoroacetyl-lysine	Lys(TFA)	lys(TFA)
Methyl-leucine	MeLeu	meLeu
3-(3-benzothienyl)-alanine	Bta	bta

* single letter abbreviations: when shown in capital letters herein it indicates the L-amino acid form, when shown in lower case herein it indicates the D-amino acid form.

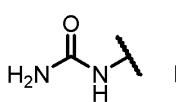
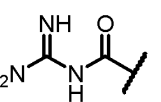
[0047] The cCPP can comprise 4 to 20 amino acids, wherein: (i) at least one amino acid has a side chain comprising a guanidine group, or a protonated form thereof; (ii) at least one amino acid has

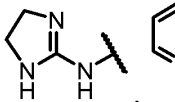
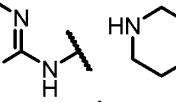
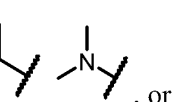
- 5 no side chain or a side chain comprising , , , or a protonated form thereof; and (iii) at least two amino acids independently have a side chain comprising an aromatic or heteroaromatic group.

[0048] At least two amino acids can have no side chain or a side chain comprising , , , , , or a protonated form thereof. As used herein, when no side chain is present, the amino acid has two hydrogen atoms on the carbon atom(s) (e.g., -CH₂-) linking the amine and carboxylic acid.

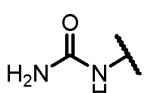
5 [0049] The amino acid having no side chain can be glycine or β-alanine.

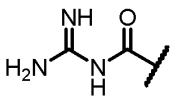
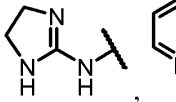
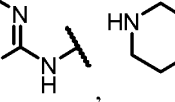
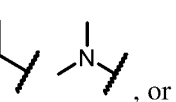
[0050] The cCPP can comprise from 6 to 20 amino acid residues which form the cCPP, wherein: (i) at least one amino acid can be glycine, β-alanine, or 4-aminobutyric acid residues; (ii) at least one amino acid can have a side chain comprising an aryl or heteroaryl group; and (iii) at least one

amino acid has a side chain comprising a guanidine group, , ,

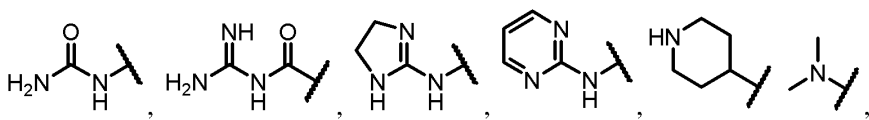
10 , , , or a protonated form thereof.

[0051] The cCPP can comprise from 6 to 20 amino acid residues which form the cCPP, wherein: (i) at least two amino acid can independently be glycine, β-alanine, or 4-aminobutyric acid residues; (ii) at least one amino acid can have a side chain comprising an aryl or heteroaryl group;

and (iii) at least one amino acid has a side chain comprising a guanidine group, ,

15 , , , , or a protonated form thereof.

[0052] The cCPP can comprise from 6 to 20 amino acid residues which form the cCPP, wherein: (i) at least three amino acids can independently be glycine, β-alanine, or 4-aminobutyric acid residues; (ii) at least one amino acid can have a side chain comprising an aromatic or heteroaromatic group; and (iii) at least one amino acid can have a side chain comprising a

guanidine group, , or a protonated form thereof.

Glycine and Related Amino Acid Residues

[0053] The cCPP can comprise (i) 1, 2, 3, 4, 5, or 6 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 2 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 4 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 5 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 6 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3, 4, or 5 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3 or 4 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof.

[0054] The cCPP can comprise (i) 1, 2, 3, 4, 5, or 6 glycine residues. The cCPP can comprise (i) 2 glycine residues. The cCPP can comprise (i) 3 glycine residues. The cCPP can comprise (i) 4 glycine residues. The cCPP can comprise (i) 5 glycine residues. The cCPP can comprise (i) 6 glycine residues. The cCPP can comprise (i) 3, 4, or 5 glycine residues. The cCPP can comprise (i) 3 or 4 glycine residues. The cCPP can comprise (i) 2 or 3 glycine residues. The cCPP can comprise (i) 1 or 2 glycine residues.

[0055] The cCPP can comprise (i) 3, 4, 5, or 6 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 4 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 5 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 6 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3, 4, or 5 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof. The cCPP can comprise (i) 3 or 4 glycine, β -alanine, 4-aminobutyric acid residues, or combinations thereof.

[0056] The cCPP can comprise at least three glycine residues. The cCPP can comprise (i) 3, 4, 5, or 6 glycine residues. The cCPP can comprise (i) 3 glycine residues. The cCPP can comprise (i) 4 glycine residues. The cCPP can comprise (i) 5 glycine residues. The cCPP can comprise (i) 6 glycine residues. The cCPP can comprise (i) 3, 4, or 5 glycine residues. The cCPP can comprise (i) 3 or 4 glycine residues

[0057] In embodiments, none of the glycine, β -alanine, or 4-aminobutyric acid residues in the cCPP are contiguous. Two or three glycine, β -alanine, 4-or aminobutyric acid residues can be contiguous. Two glycine, β -alanine, or 4-aminobutyric acid residues can be contiguous.

[0058] In embodiments, none of the glycine residues in the cCPP are contiguous. Each glycine residues in the cCPP can be separated by an amino acid residue that cannot be glycine. Two or three glycine residues can be contiguous. Two glycine residues can be contiguous

Amino Acid Side Chains with an Aromatic or Heteroaromatic Group

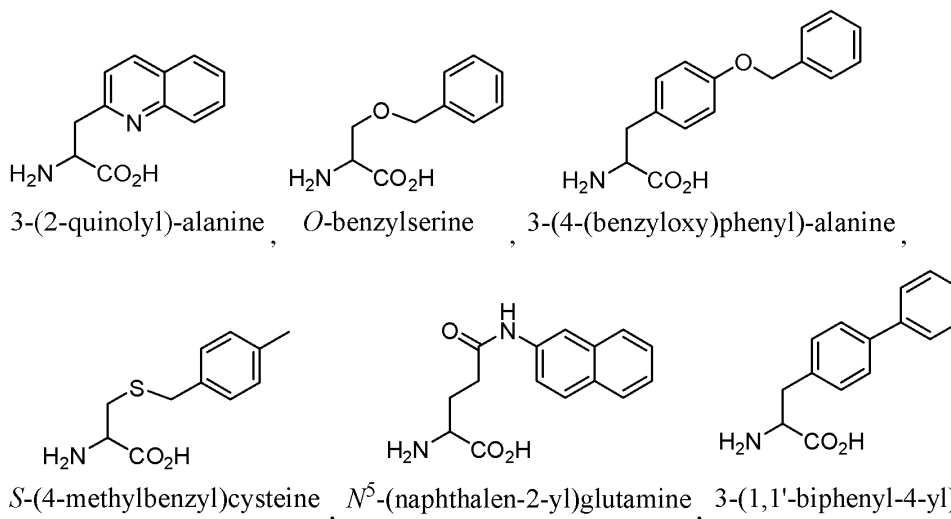
[0059] The cCPP can comprise (ii) 2, 3, 4, 5 or 6 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 2 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 3 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 4 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 5 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 6 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 2, 3, or 4 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group. The cCPP can comprise (ii) 2 or 3 amino acid residues independently having a side chain comprising an aromatic or heteroaromatic group.

[0060] The cCPP can comprise (ii) 2, 3, 4, 5 or 6 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 2 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 3 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 4 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 5 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 6 amino acid residues independently

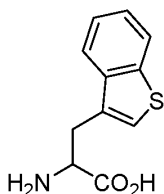
having a side chain comprising an aromatic group. The cCPP can comprise (ii) 2, 3, or 4 amino acid residues independently having a side chain comprising an aromatic group. The cCPP can comprise (ii) 2 or 3 amino acid residues independently having a side chain comprising an aromatic group.

5 **[0061]** The aromatic group can be a 6- to 14-membered aryl. Aryl can be phenyl, naphthyl or anthracenyl, each of which is optionally substituted. Aryl can be phenyl or naphthyl, each of which is optionally substituted. The heteroaromatic group can be a 6- to 14-membered heteroaryl having 1, 2, or 3 heteroatoms selected from N, O, and S. Heteroaryl can be pyridyl, quinolyl, or isoquinolyl.

10 **[0062]** The amino acid residue having a side chain comprising an aromatic or heteroaromatic group can each independently be bis(homonaphthylalanine), homonaphthylalanine, naphthylalanine, phenylglycine, bis(homophenylalanine), homophenylalanine, phenylalanine, tryptophan, 3-(3-benzothienyl)-alanine, 3-(2-quinolyl)-alanine, O-benzylserine, 3-(4-(benzyloxy)phenyl)-alanine, S-(4-methylbenzyl)cysteine, N-(naphthalen-2-yl)glutamine, 3-(1,1'-
15 biphenyl-4-yl)-alanine, 3-(3-benzothienyl)-alanine or tyrosine, each of which is optionally substituted with one or more substituents. The amino acid having a side chain comprising an aromatic or heteroaromatic group can each independently be selected from:



20 and



3-(3-benzothieryl)-alanine, wherein the H on the N-terminus and/or the H on the C-terminus are replaced by a peptide bond.

[0063] The amino acid residue having a side chain comprising an aromatic or heteroaromatic group can each be independently a residue of phenylalanine, naphthylalanine, phenylglycine, homophenylalanine, homonaphthylalanine, bis(homophenylalanine), bis-(homonaphthylalanine), tryptophan, or tyrosine, each of which is optionally substituted with one or more substituents. The amino acid residue having a side chain comprising an aromatic group can each independently be a residue of tyrosine, phenylalanine, 1-naphthylalanine, 2-naphthylalanine, tryptophan, 3-benzothierylalanine, 4-phenylphenylalanine, 3,4-difluorophenylalanine, 4-trifluoromethylphenylalanine, 2,3,4,5,6-pentafluorophenylalanine, homophenylalanine, β -homophenylalanine, 4-tert-butyl-phenylalanine, 4-pyridinylalanine, 3-pyridinylalanine, 4-methylphenylalanine, 4-fluorophenylalanine, 4-chlorophenylalanine, 3-(9-anthryl)-alanine. The amino acid residue having a side chain comprising an aromatic group can each independently be a residue of phenylalanine, naphthylalanine, phenylglycine, homophenylalanine, or homonaphthylalanine, each of which is optionally substituted with one or more substituents. The amino acid residue having a side chain comprising an aromatic group can each be independently a residue of phenylalanine, naphthylalanine, homophenylalanine, homonaphthylalanine, bis(homonaphthylalanine), or bis(homonaphthylalanine), each of which is optionally substituted with one or more substituents. The amino acid residue having a side chain comprising an aromatic group can each be independently a residue of phenylalanine or naphthylalanine, each of which is optionally substituted with one or more substituents. At least one amino acid residue having a side chain comprising an aromatic group can be a residue of phenylalanine. At least two amino acid residues having a side chain comprising an aromatic group can be residues of phenylalanine. Each amino acid residue having a side chain comprising an aromatic group can be a residue of phenylalanine.

[0064] In embodiments, none of the amino acids having the side chain comprising the aromatic or heteroaromatic group are contiguous. Two amino acids having the side chain comprising the

aromatic or heteroaromatic group can be contiguous. Two contiguous amino acids can have opposite stereochemistry. The two contiguous amino acids can have the same stereochemistry. Three amino acids having the side chain comprising the aromatic or heteroaromatic group can be contiguous. Three contiguous amino acids can have the same stereochemistry. Three contiguous amino acids can have alternating stereochemistry.

[0065] The amino acid residues comprising aromatic or heteroaromatic groups can be L-amino acids. The amino acid residues comprising aromatic or heteroaromatic groups can be D-amino acids. The amino acid residues comprising aromatic or heteroaromatic groups can be a mixture of D- and L-amino acids.

[0066] The optional substituent can be any atom or group which does not significantly reduce (e.g., by more than 50%) the cytosolic delivery efficiency of the cCPP, e.g., compared to an otherwise identical sequence which does not have the substituent. The optional substituent can be a hydrophobic substituent or a hydrophilic substituent. The optional substituent can be a hydrophobic substituent. The substituent can increase the solvent-accessible surface area (as defined herein) of the hydrophobic amino acid. The substituent can be halogen, alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkenyl, cycloalkynyl, heterocyclyl, aryl, heteroaryl, alkoxy, aryloxy, acyl, alkylcarbamoyl, alkylcarboxamidyl, alkoxycarbonyl, alkylthio, or arylthio. The substituent can be halogen.

[0067] While not wishing to be bound by theory, it is believed that amino acids having an aromatic or heteroaromatic group having higher hydrophobicity values (i.e., amino acids having side chains comprising aromatic or heteroaromatic groups) can improve cytosolic delivery efficiency of a cCPP relative to amino acids having a lower hydrophobicity value. Each hydrophobic amino acid can independently have a hydrophobicity value greater than that of glycine. Each hydrophobic amino acid can independently be a hydrophobic amino acid having a hydrophobicity value greater than that of alanine. Each hydrophobic amino acid can independently have a hydrophobicity value greater or equal to phenylalanine. Hydrophobicity may be measured using hydrophobicity scales known in the art. **Table 2** lists hydrophobicity values for various amino acids as reported by Eisenberg and Weiss (Proc. Natl. Acad. Sci. U. S. A. 1984;81(1):140–144), Engleman, et al. (Ann. Rev. of Biophys. Biophys. Chem.. 1986;1986(15):321–53), Kyte and Doolittle (J. Mol. Biol. 1982;157(1):105–132), Hoop and Woods (Proc. Natl. Acad. Sci. U. S. A. 1981;78(6):3824–3828), and Janin (Nature. 1979;277(5696):491–492), the entirety of each of which is herein

incorporated by reference. Hydrophobicity can be measured using the hydrophobicity scale reported in Engleman, et al.

Table 2. Amino Acid Hydrophobicity

Amino Acid	Group	Eisenberg and Weiss	Engleman et al.	Kyrie and Doolittle	Hoop and Woods	Janin
Ile	Nonpolar	0.73	3.1	4.5	-1.8	0.7
Phe	Nonpolar	0.61	3.7	2.8	-2.5	0.5
Val	Nonpolar	0.54	2.6	4.2	-1.5	0.6
Leu	Nonpolar	0.53	2.8	3.8	-1.8	0.5
Trp	Nonpolar	0.37	1.9	-0.9	-3.4	0.3
Met	Nonpolar	0.26	3.4	1.9	-1.3	0.4
Ala	Nonpolar	0.25	1.6	1.8	-0.5	0.3
Gly	Nonpolar	0.16	1.0	-0.4	0.0	0.3
Cys	Unch/Polar	0.04	2.0	2.5	-1.0	0.9
Tyr	Unch/Polar	0.02	-0.7	-1.3	-2.3	-0.4
Pro	Nonpolar	-0.07	-0.2	-1.6	0.0	-0.3
Thr	Unch/Polar	-0.18	1.2	-0.7	-0.4	-0.2
Ser	Unch/Polar	-0.26	0.6	-0.8	0.3	-0.1
His	Charged	-0.40	-3.0	-3.2	-0.5	-0.1
Glu	Charged	-0.62	-8.2	-3.5	3.0	-0.7
Asn	Unch/Polar	-0.64	-4.8	-3.5	0.2	-0.5
Gln	Unch/Polar	-0.69	-4.1	-3.5	0.2	-0.7
Asp	Charged	-0.72	-9.2	-3.5	3.0	-0.6
Lys	Charged	-1.10	-8.8	-3.9	3.0	-1.8
Arg	Charged	-1.80	-12.3	-4.5	3.0	-1.4

- 5 [0068] The size of the aromatic or heteroaromatic groups may be selected to improve cytosolic delivery efficiency of the cCPP. While not wishing to be bound by theory, it is believed that a larger aromatic or heteroaromatic group on the side chain of amino acid may improve cytosolic delivery efficiency compared to an otherwise identical sequence having a smaller hydrophobic amino acid. The size of the hydrophobic amino acid can be measured in terms of molecular weight
- 10 of the hydrophobic amino acid, the steric effects of the hydrophobic amino acid, the solvent-accessible surface area (SASA) of the side chain, or combinations thereof. The size of the hydrophobic amino acid can be measured in terms of the molecular weight of the hydrophobic amino acid, and the larger hydrophobic amino acid has a side chain with a molecular weight of at least about 90 g/mol, or at least about 130 g/mol, or at least about 141 g/mol. The size of the amino
- 15 acid can be measured in terms of the SASA of the hydrophobic side chain. The hydrophobic amino acid can have a side chain with a SASA of greater than or equal to alanine, or greater than or equal to glycine. Larger hydrophobic amino acids can have a side chain with a SASA greater than alanine, or greater than glycine. The hydrophobic amino acid can have an aromatic or

heteroaromatic group with a SASA greater than or equal to about piperidine-2-carboxylic acid, greater than or equal to about tryptophan, greater than or equal to about phenylalanine, or greater than or equal to about naphthylalanine. A first hydrophobic amino acid (AA_{H1}) can have a side chain with a SASA of at least about 200 Å², at least about 210 Å², at least about 220 Å², at least about 240 Å², at least about 250 Å², at least about 260 Å², at least about 270 Å², at least about 280 Å², at least about 290 Å², at least about 300 Å², at least about 310 Å², at least about 320 Å², or at least about 330 Å². A second hydrophobic amino acid (AA_{H2}) can have a side chain with a SASA of at least about 200 Å², at least about 210 Å², at least about 220 Å², at least about 240 Å², at least about 250 Å², at least about 260 Å², at least about 270 Å², at least about 280 Å², at least about 290 Å², at least about 300 Å², at least about 310 Å², at least about 320 Å², or at least about 330 Å². The side chains of AA_{H1} and AA_{H2} can have a combined SASA of at least about 350 Å², at least about 360 Å², at least about 370 Å², at least about 380 Å², at least about 390 Å², at least about 400 Å², at least about 410 Å², at least about 420 Å², at least about 430 Å², at least about 440 Å², at least about 450 Å², at least about 460 Å², at least about 470 Å², at least about 480 Å², at least about 490 Å², greater than about 500 Å², at least about 510 Å², at least about 520 Å², at least about 530 Å², at least about 540 Å², at least about 550 Å², at least about 560 Å², at least about 570 Å², at least about 580 Å², at least about 590 Å², at least about 600 Å², at least about 610 Å², at least about 620 Å², at least about 630 Å², at least about 640 Å², greater than about 650 Å², at least about 660 Å², at least about 670 Å², at least about 680 Å², at least about 690 Å², or at least about 700 Å². AA_{H2} can be a hydrophobic amino acid residue with a side chain having a SASA that is less than or equal to the SASA of the hydrophobic side chain of AA_{H1}. By way of example, and not by limitation, a cCPP having a Nal-Arg motif may exhibit improved cytosolic delivery efficiency compared to an otherwise identical cCPP having a Phe-Arg motif; a cCPP having a Phe-Nal-Arg motif may exhibit improved cytosolic delivery efficiency compared to an otherwise identical cCPP having a Nal-Phe-Arg motif; and a phe-Nal-Arg motif may exhibit improved cytosolic delivery efficiency compared to an otherwise identical cCPP having a nal-Phe-Arg motif.

[0069] As used herein, “hydrophobic surface area” or “SASA” refers to the surface area (reported as square Ångstroms; Å²) of an amino acid side chain that is accessible to a solvent., SASA can be calculated using the 'rolling ball' algorithm developed by Shrake & Rupley (*J Mol Biol.* **79** (2): 351–71), which is herein incorporated by reference in its entirety for all purposes. This algorithm

uses a “sphere” of solvent of a particular radius to probe the surface of the molecule. A typical value of the sphere is 1.4 Å, which approximates to the radius of a water molecule.

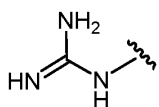
[0070] SASA values for certain side chains are shown below in **Table 3**. The SASA values described herein are based on the theoretical values listed in **Table 3** below, as reported by Tien, et al. (PLOS ONE 8(11): e80635. <https://doi.org/10.1371/journal.pone.0080635>), which is herein incorporated by reference in its entirety for all purposes.

Table 3. Amino Acid SASA Values

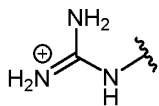
Residue	Theoretical	Empirical	Miller <i>et al.</i> (1987)	Rose <i>et al.</i> (1985)
Alanine	129.0	121.0	113.0	118.1
Arginine	274.0	265.0	241.0	256.0
Asparagine	195.0	187.0	158.0	165.5
Aspartate	193.0	187.0	151.0	158.7
Cysteine	167.0	148.0	140.0	146.1
Glutamate	223.0	214.0	183.0	186.2
Glutamine	225.0	214.0	189.0	193.2
Glycine	104.0	97.0	85.0	88.1
Histidine	224.0	216.0	194.0	202.5
Isoleucine	197.0	195.0	182.0	181.0
Leucine	201.0	191.0	180.0	193.1
Lysine	236.0	230.0	211.0	225.8
Methionine	224.0	203.0	204.0	203.4
Phenylalanine	240.0	228.0	218.0	222.8
Proline	159.0	154.0	143.0	146.8
Serine	155.0	143.0	122.0	129.8
Threonine	172.0	163.0	146.0	152.5
Tryptophan	285.0	264.0	259.0	266.3
Tyrosine	263.0	255.0	229.0	236.8
Valine	174.0	165.0	160.0	164.5

Amino Acid Residues Having a Side Chain Comprising a Guanidine Group, Guanidine Replacement Group, or Protonated Form Thereof

10 [0071] As used herein, guanidine refers to the structure:

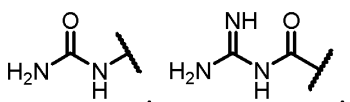
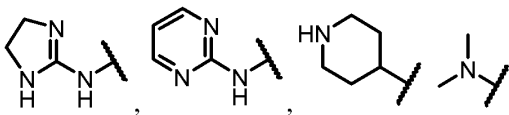


[0072] As used herein, a protonated form of guanidine refers to the structure:

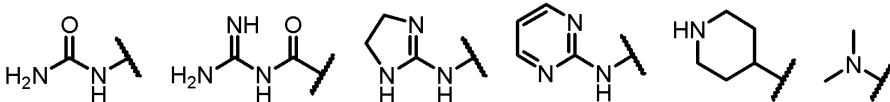


[0073] Guanidine replacement groups refer to functional groups on the side chain of amino acids that will be positively charged at or above physiological pH or those that can recapitulate the hydrogen bond donating and accepting activity of guanidinium groups.

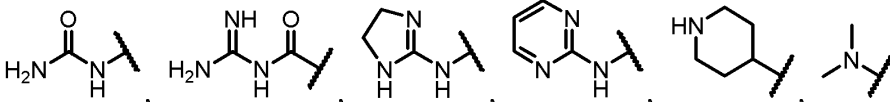
- [0074] The guanidine replacement groups facilitate cell penetration and delivery of therapeutic agents while reducing toxicity associated with guanidine groups or protonated forms thereof. The cCPP can comprise at least one amino acid having a side chain comprising a guanidine or guanidinium replacement group. The cCPP can comprise at least two amino acids having a side chain comprising a guanidine or guanidinium replacement group. The cCPP can comprise at least three amino acids having a side chain comprising a guanidine or guanidinium replacement group
- [0075] The guanidine or guanidinium group can be an isostere of guanidine or guanidinium. The guanidine or guanidinium replacement group can be less basic than guanidine.

[0076] As used herein, a guanidine replacement group refers to , , or a protonated form thereof.

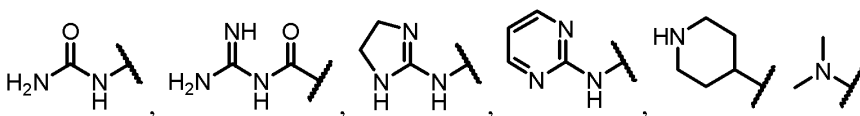
- [0077] The disclosure relates to a cCPP comprising from 4 to 20 amino acids residues, wherein:
- (i) at least one amino acid has a side chain comprising a guanidine group, or a protonated form thereof; (ii) at least one amino acid residue has no side chain or a side chain comprising

, or a protonated form thereof; and (iii) at least two amino acids residues independently have a side chain comprising an aromatic or heteroaromatic group.

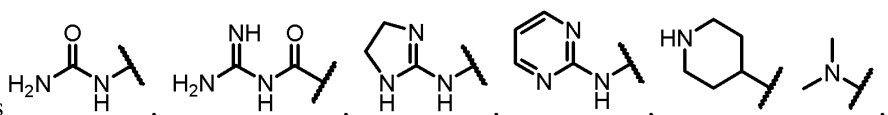
- [0078] At least two amino acids residues can have no side chain or a side chain comprising

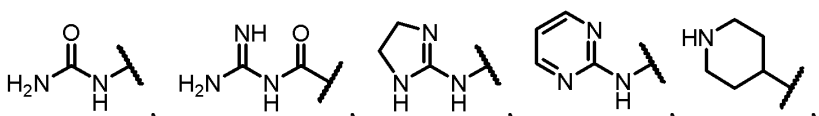
, or a protonated form thereof. As used herein, when no side chain is present, the amino acid residue have two hydrogen atoms on the carbon atom(s) (e.g., -CH₂-) linking the amine and carboxylic acid.

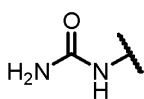
[0079] The cCPP can comprise at least one amino acid having a side chain comprising one of the

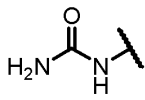
following moieties: , or a protonated form thereof.

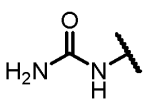
[0080] The cCPP can comprise at least two amino acids each independently having one of the

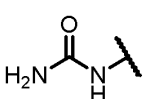
5 following moieties , or a protonated form thereof. At least two amino acids can have a side chain comprising the same

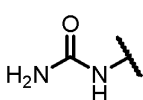
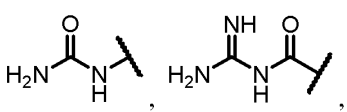
moiety selected from: , , or a protonated form thereof. At least one amino acid can have a side chain comprising

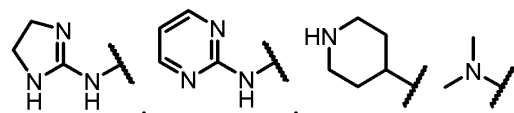
, or a protonated form thereof. At least two amino acids can have a side chain

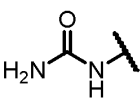
10 comprising , or a protonated form thereof. One, two, three, or four amino acids can

have a side chain comprising , or a protonated form thereof. One amino acid can have

a side chain comprising , or a protonated form thereof. Two amino acids can have a

side chain comprising , or a protonated form thereof. ,

, or a protonated form thereof, can be attached to the

terminus of the amino acid side chain.  can be attached to the terminus of the amino acid side chain.

[0081] The cCPP can comprise (iii) 2, 3, 4, 5 or 6 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof.

5 The cCPP can comprise (iii) 2 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 3 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 4 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 5 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 6 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 2, 3, 4, or 5 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 2, 3, or 4 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) 2 or 3 amino acid residues independently having a side chain comprising a guanidine group, guanidine replacement group, or a protonated form thereof. The cCPP can comprise (iii) at least one amino acid residue having a side chain comprising a guanidine group or protonated form thereof. The cCPP can comprise (iii) two amino acid residues having a side chain comprising a guanidine group or protonated form thereof. The cCPP can comprise (iii) three amino acid residues having a side chain comprising a guanidine group or protonated form thereof.

25 **[0082]** The amino acid residues can independently have the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof that are not contiguous. Two amino acid residues can independently have the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof can be contiguous. Three amino acid residues can independently have the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof can be contiguous. Four amino acid residues

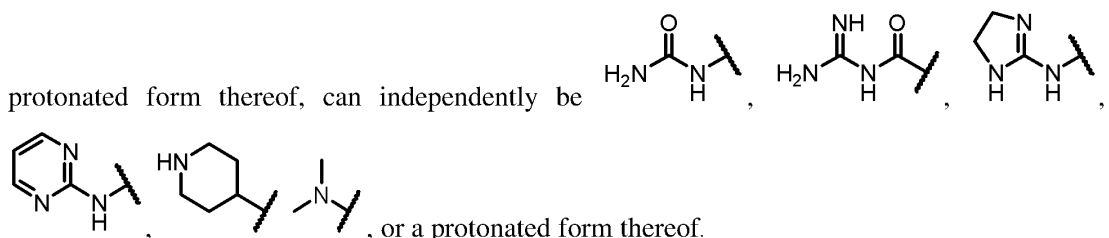
can independently have the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof can be contiguous. The contiguous amino acid residues can have the same stereochemistry. The contiguous amino acids can have alternating stereochemistry.

[0083] The amino acid residues independently having the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof, can be L-amino acids. The amino acid residues independently having the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof, can be D-amino acids. The amino acid residues independently having the side chain comprising the guanidine group, guanidine replacement group, or the protonated form thereof, can be a mixture of L- or D-amino acids.

[0084] Each amino acid residue having the side chain comprising the guanidine group, or the protonated form thereof, can independently be a residue of arginine, homoarginine, 2-amino-3-propionic acid, 2-amino-4-guanidinobutyric acid or a protonated form thereof. Each amino acid residue having the side chain comprising the guanidine group, or the protonated form thereof, can independently be a residue of arginine or a protonated form thereof.

[0085] Each amino acid having the side chain comprising a guanidine replacement group, or

protonated form thereof, can independently be



, or a protonated form thereof.

[0086] Without being bound by theory, it is hypothesized that guanidine replacement groups have reduced basicity, relative to arginine and in some cases are uncharged at physiological pH (e.g., a -N(H)C(O)), and are capable of maintaining the bidentate hydrogen bonding interactions with phospholipids on the plasma membrane that is believed to facilitate effective membrane association and subsequent internalization. The removal of positive charge is also believed to reduce toxicity of the cCPP.

[0087] Those skilled in the art will appreciate that the N- and/or C-termini of the above non-natural aromatic hydrophobic amino acids, upon incorporation into the peptides disclosed herein, form amide bonds.

[0088] The cCPP can comprise a first amino acid having a side chain comprising an aromatic or heteroaromatic group and a second amino acid having a side chain comprising an aromatic or

heteroaromatic group, wherein an N-terminus of a first glycine forms a peptide bond with the first amino acid having the side chain comprising the aromatic or heteroaromatic group, and a C-terminus of the first glycine forms a peptide bond with the second amino acid having the side chain comprising the aromatic or heteroaromatic group. Although by convention, the term “first amino acid” often refers to the N-terminal amino acid of a peptide sequence, as used herein “first amino acid” is used to distinguish the referent amino acid from another amino acid (e.g., a “second amino acid”) in the cCPP such that the term “first amino acid” may or may refer to an amino acid located at the N-terminus of the peptide sequence.

[0089] The cCPP can comprise an N-terminus of a second glycine forms a peptide bond with an amino acid having a side chain comprising an aromatic or heteroaromatic group, and a C-terminus of the second glycine forms a peptide bond with an amino acid having a side chain comprising a guanidine group, or a protonated form thereof.

[0090] The cCPP can comprise a first amino acid having a side chain comprising a guanidine group, or a protonated form thereof, and a second amino acid having a side chain comprising a guanidine group, or a protonated form thereof, wherein an N-terminus of a third glycine forms a peptide bond with a first amino acid having a side chain comprising a guanidine group, or a protonated form thereof, and a C-terminus of the third glycine forms a peptide bond with a second amino acid having a side chain comprising a guanidine group, or a protonated form thereof.

[0091] The cCPP can comprise a residue of asparagine, aspartic acid, glutamine, glutamic acid, or homoglutamine. The cCPP can comprise a residue of asparagine. The cCPP can comprise a residue of glutamine.

[0092] The cCPP can comprise a residue of tyrosine, phenylalanine, 1-naphthylalanine, 2-naphthylalanine, tryptophan, 3-benzothienylalanine, 4-phenylphenylalanine, 3,4-difluorophenylalanine, 4-trifluoromethylphenylalanine, 2,3,4,5,6-pentafluorophenylalanine, homophenylalanine, β -homophenylalanine, 4-tert-butyl-phenylalanine, 4-pyridinylalanine, 3-pyridinylalanine, 4-methylphenylalanine, 4-fluorophenylalanine, 4-chlorophenylalanine, 3-(9-anthryl)-alanine.

[0093] While not wishing to be bound by theory, it is believed that the chirality of the amino acids in the cCPPs may impact cytosolic uptake efficiency. The cCPP can comprise at least one D amino acid. The cCPP can comprise one to fifteen D amino acids. The cCPP can comprise one to ten D amino acids. The cCPP can comprise 1, 2, 3, or 4 D amino acids. The cCPP can comprise 2, 3, 4,

5, 6, 7, or 8 contiguous amino acids having alternating D and L chirality. The cCPP can comprise three contiguous amino acids having the same chirality. The cCPP can comprise two contiguous amino acids having the same chirality. At least two of the amino acids can have the opposite chirality. The at least two amino acids having the opposite chirality can be adjacent to each other.

5 At least three amino acids can have alternating stereochemistry relative to each other. The at least three amino acids having the alternating chirality relative to each other can be adjacent to each other. At least four amino acids have alternating stereochemistry relative to each other. The at least four amino acids having the alternating chirality relative to each other can be adjacent to each other. At least two of the amino acids can have the same chirality. At least two amino acids having

10 the same chirality can be adjacent to each other. At least two amino acids have the same chirality and at least two amino acids have the opposite chirality. The at least two amino acids having the opposite chirality can be adjacent to the at least two amino acids having the same chirality. Accordingly, adjacent amino acids in the cCPP can have any of the following sequences: D-L; L-D; D-L-L-D; L-D-D-L; L-D-L-L-D; D-L-D-D-L; D-L-L-D-L; or L-D-D-L-D. The amino acid

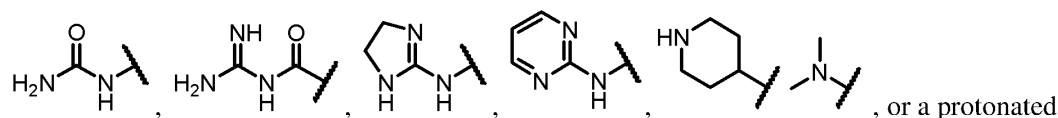
15 residues that form the cCPP can all be L-amino acids. The amino acid residues that form the cCPP can all be D-amino acids.

[0094] At least two of the amino acids can have a different chirality. At least two amino acids having a different chirality can be adjacent to each other. At least three amino acids can have different chirality relative to an adjacent amino acid. At least four amino acids can have different

20 chirality relative to an adjacent amino acid. At least two amino acids have the same chirality and at least two amino acids have a different chirality. One or more amino acid residues that form the cCPP can be achiral. The cCPP can comprise a motif of 3, 4, or 5 amino acids, wherein two amino acids having the same chirality can be separated by an achiral amino acid. The cCPPs can comprise the following sequences: D-X-D; D-X-D-X; D-X-D-X-D; L-X-L; L-X-L-X; or L-X-L-X-L,

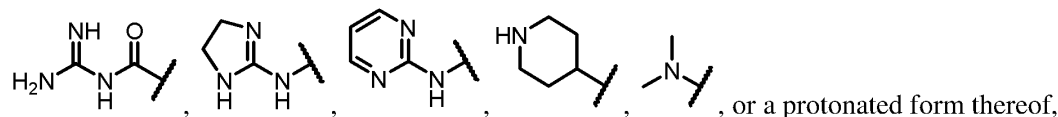
25 wherein X is an achiral amino acid. The achiral amino acid can be glycine.

[0095] An amino acid having a side chain comprising:



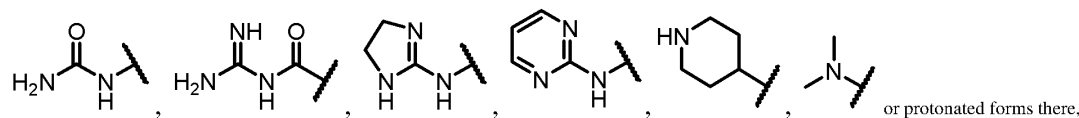
form thereof, can be adjacent to an amino acid having a side chain comprising an aromatic or

heteroaromatic group. An amino acid having a side chain comprising:



can be adjacent to at least one amino acid having a side chain comprising a guanidine or protonated form thereof. An amino acid having a side chain comprising a guanidine or

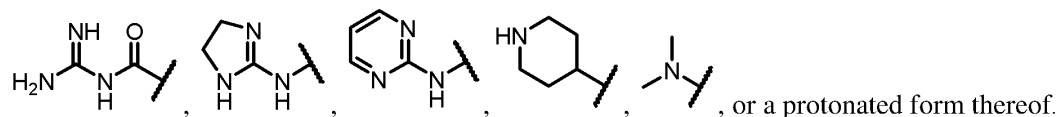
- 5 protonated form thereof can be adjacent to an amino acid having a side chain comprising an aromatic or heteroaromatic group. Two amino acids having a side chain comprising:



can be adjacent to each other. Two amino acids having a side chain comprising a guanidine or protonated form thereof are adjacent to each other. The cCPPs can comprise at least two

- 10 contiguous amino acids having a side chain can comprise an aromatic or heteroaromatic group

and at least two non-adjacent amino acids having a side chain comprising:



The cCPPs can comprise at least two contiguous amino acids having a side chain comprising an aromatic or heteroaromatic group and at least two non-adjacent amino acids having a side chain

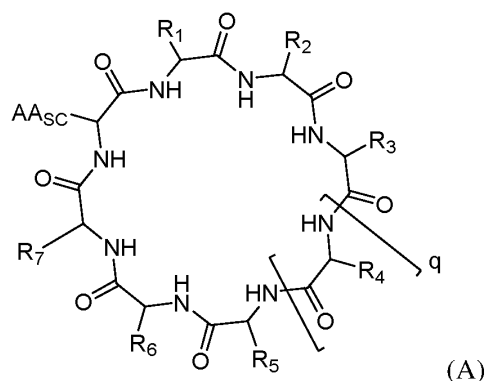
- 15 comprising $\text{H}_2\text{N}-\text{C}(=\text{NH})-\text{NH}-$, or a protonated form thereof. The adjacent amino acids can have the same chirality. The adjacent amino acids can have the opposite chirality. Other combinations of amino acids can have any arrangement of D and L amino acids, e.g., any of the sequences described in the preceding paragraph.

[0096] At least two amino acids having a side chain comprising:

- 20 $\text{H}_2\text{N}-\text{C}(=\text{NH})-\text{NH}-$, $\text{H}_2\text{N}-\text{C}(=\text{NH})-\text{NH}-$ (cyclic), $\text{H}_2\text{N}-\text{C}(=\text{NH})-\text{NH}-$ (pyridine), $\text{HN}-$ (piperidine), $\text{N}-$ (methyl), or a protonated

form thereof, are alternating with at least two amino acids having a side chain comprising a guanidine group or protonated form thereof.

[0097] The cCPP can comprise the structure of Formula (A):



- 5 or a protonated form thereof,
wherein:

R₁, R₂, and R₃ are each independently H or an aromatic or heteroaromatic side chain of an amino acid;

at least one of R₁, R₂, and R₃ is an aromatic or heteroaromatic side chain of an amino acid;

- 10 R₄, R₅, R₆, R₇ are independently H or an amino acid side chain;

at least one of R₄, R₅, R₆, R₇ is the side chain of 3-guanidino-2-aminopropionic acid, 4-guanidino-2-aminobutanoic acid, arginine, homoarginine, N-methylarginine, N,N-dimethylarginine, 2,3-diaminopropionic acid, 2,4-diaminobutanoic acid, lysine, N-methyllysine, N,N-dimethyllysine, N-ethyllysine, N,N,N-trimethyllysine, 4-guanidinophenylalanine,

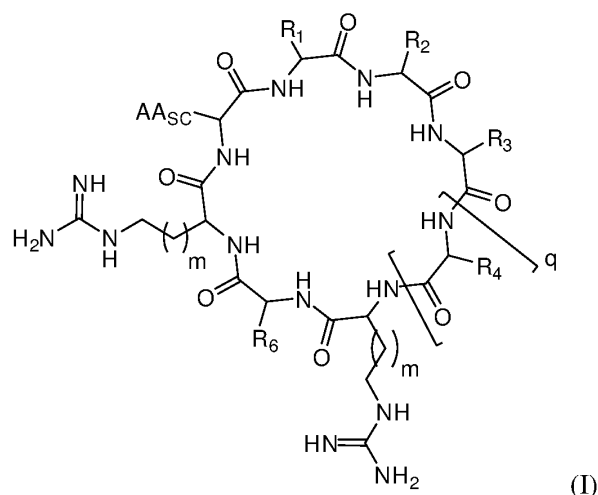
- 15 citrulline, N,N-dimethyllysine, β-homoarginine, 3-(1-piperidinyl)alanine;

AA_{SC} is an amino acid side chain; and

q is 1, 2, 3 or 4;

wherein the cyclic peptide of Formula (A) is not FfΦRrRrQ.

[0098] The cCPP can comprise the structure of Formula (I):



or a protonated form thereof,

wherein:

R_1 , R_2 , and R_3 can each independently be H or an amino acid residue having a side chain comprising an aromatic group;

at least one of R_1 , R_2 , and R_3 is an aromatic or heteroaromatic side chain of an amino acid;

R_4 and R_6 are independently H or an amino acid side chain;

AA_{SC} is an amino acid side chain;

q is 1, 2, 3 or 4; and

each m is independently an integer 0, 1, 2, or 3.

[0099] R_1 , R_2 , and R_3 can each independently be H, -alkylene-aryl, or -alkylene-heteroaryl. R_1 , R_2 , and R_3 can each independently be H, - C_{1-3} alkylene-aryl, or - C_{1-3} alkylene-heteroaryl. R_1 , R_2 , and R_3 can each independently be H or -alkylene-aryl. R_1 , R_2 , and R_3 can each independently be H or - C_{1-3} alkylene-aryl. C_{1-3} alkylene can be methylene. Aryl can be a 6- to 14-membered aryl. Heteroaryl can be a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S. Aryl can be selected from phenyl, naphthyl, or anthracenyl. Aryl can be phenyl or naphthyl. Aryl can be phenyl. Heteroaryl can be pyridyl, quinolyl, and isoquinolyl. R_1 , R_2 , and R_3 can each independently be H, - C_{1-3} alkylene-Ph or - C_{1-3} alkylene-Naphthyl. R_1 , R_2 , and R_3 can each independently be H, - CH_2 Ph, or - CH_2 Naphthyl. R_1 , R_2 , and R_3 can each independently be H or - CH_2 Ph.

[0100] R₁, R₂, and R₃ can each independently be the side chain of tyrosine, phenylalanine, 1-naphthylalanine, 2-naphthylalanine, tryptophan, 3-benzothienylalanine, 4-phenylphenylalanine, 3,4-difluorophenylalanine, 4-trifluoromethylphenylalanine, 2,3,4,5,6-pentafluorophenylalanine, homophenylalanine, β -homophenylalanine, 4-tert-butyl-phenylalanine, 4-pyridinylalanine, 3-pyridinylalanine, 4-methylphenylalanine, 4-fluorophenylalanine, 4-chlorophenylalanine, 3-(9-anthryl)-alanine.

[0101] R₁ can be the side chain of tyrosine. R₁ can be the side chain of phenylalanine. R₁ can be the side chain of 1-naphthylalanine. R₁ can be the side chain of 2-naphthylalanine. R₁ can be the side chain of tryptophan. R₁ can be the side chain of 3-benzothienylalanine. R₁ can be the side chain of 4-phenylphenylalanine. R₁ can be the side chain of 3,4-difluorophenylalanine. R₁ can be the side chain of 4-trifluoromethylphenylalanine. R₁ can be the side chain of 2,3,4,5,6-pentafluorophenylalanine. R₁ can be the side chain of homophenylalanine. R₁ can be the side chain of β -homophenylalanine. R₁ can be the side chain of 4-tert-butyl-phenylalanine. R₁ can be the side chain of 4-pyridinylalanine. R₁ can be the side chain of 3-pyridinylalanine. R₁ can be the side chain of 4-methylphenylalanine. R₁ can be the side chain of 4-fluorophenylalanine. R₁ can be the side chain of 4-chlorophenylalanine. R₁ can be the side chain of 3-(9-anthryl)-alanine.

[0102] R₂ can be the side chain of tyrosine. R₂ can be the side chain of phenylalanine. R₂ can be the side chain of 1-naphthylalanine. R₂ can be the side chain of 2-naphthylalanine. R₂ can be the side chain of tryptophan. R₂ can be the side chain of 3-benzothienylalanine. R₂ can be the side chain of 4-phenylphenylalanine. R₂ can be the side chain of 3,4-difluorophenylalanine. R₂ can be the side chain of 4-trifluoromethylphenylalanine. R₂ can be the side chain of 2,3,4,5,6-pentafluorophenylalanine. R₂ can be the side chain of homophenylalanine. R₂ can be the side chain of β -homophenylalanine. R₂ can be the side chain of 4-tert-butyl-phenylalanine. R₂ can be the side chain of 4-pyridinylalanine. R₂ can be the side chain of 3-pyridinylalanine. R₂ can be the side chain of 4-methylphenylalanine. R₂ can be the side chain of 4-fluorophenylalanine. R₂ can be the side chain of 4-chlorophenylalanine. R₂ can be the side chain of 3-(9-anthryl)-alanine.

[0103] R₃ can be the side chain of tyrosine. R₃ can be the side chain of phenylalanine. R₃ can be the side chain of 1-naphthylalanine. R₃ can be the side chain of 2-naphthylalanine. R₃ can be the side chain of tryptophan. R₃ can be the side chain of 3-benzothienylalanine. R₃ can be the side chain of 4-phenylphenylalanine. R₃ can be the side chain of 3,4-difluorophenylalanine. R₃ can be the side chain of 4-trifluoromethylphenylalanine. R₃ can be the side chain of 2,3,4,5,6-

pentafluorophenylalanine. R₃ can be the side chain of homophenylalanine. R₃ can be the side chain of β-homophenylalanine. R₃ can be the side chain of 4-tert-butyl-phenylalanine. R₃ can be the side chain of 4-pyridinylalanine. R₃ can be the side chain of 3-pyridinylalanine. R₃ can be the side chain of 4-methylphenylalanine. R₃ can be the side chain of 4-fluorophenylalanine. R₃ can be the side chain of 4-chlorophenylalanine. R₃ can be the side chain of 3-(9-anthryl)-alanine.

[0104] R₄ can be H, -alkylene-aryl, -alkylene-heteroaryl. R₄ can be H, -C₁₋₃alkylene-aryl, or -C₁₋₃alkylene-heteroaryl. R₄ can be H or -alkylene-aryl. R₄ can be H or -C₁₋₃alkylene-aryl. C₁₋₃alkylene can be a methylene. Aryl can be a 6- to 14-membered aryl. Heteroaryl can be a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S. Aryl can be selected from phenyl, naphthyl, or anthracenyl. Aryl can be phenyl or naphthyl. Aryl can phenyl. Heteroaryl can be pyridyl, quinolyl, and isoquinolyl. R₄ can be H, -C₁₋₃alkylene-Ph or -C₁₋₃alkylene-Naphthyl. R₄ can be H or the side chain of an amino acid in Table 1 or Table 3. R₄ can be H or an amino acid residue having a side chain comprising an aromatic group. R₄ can be H, -CH₂Ph, or -CH₂Naphthyl. R₄ can be H or -CH₂Ph.

[0105] R₅ can be H, -alkylene-aryl, -alkylene-heteroaryl. R₅ can be H, -C₁₋₃alkylene-aryl, or -C₁₋₃alkylene-heteroaryl. R₅ can be H or -alkylene-aryl. R₅ can be H or -C₁₋₃alkylene-aryl. C₁₋₃alkylene can be a methylene. Aryl can be a 6- to 14-membered aryl. Heteroaryl can be a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S. Aryl can be selected from phenyl, naphthyl, or anthracenyl. Aryl can be phenyl or naphthyl. Aryl can phenyl. Heteroaryl can be pyridyl, quinolyl, and isoquinolyl. R₅ can be H, -C₁₋₃alkylene-Ph or -C₁₋₃alkylene-Naphthyl. R₅ can be H or the side chain of an amino acid in **Table 1** or **Table 3**. R₄ can be H or an amino acid residue having a side chain comprising an aromatic group. R₅ can be H, -CH₂Ph, or -CH₂Naphthyl. R₄ can be H or -CH₂Ph.

[0106] R₆ can be H, -alkylene-aryl, -alkylene-heteroaryl. R₆ can be H, -C₁₋₃alkylene-aryl, or -C₁₋₃alkylene-heteroaryl. R₆ can be H or -alkylene-aryl. R₆ can be H or -C₁₋₃alkylene-aryl. C₁₋₃alkylene can be a methylene. Aryl can be a 6- to 14-membered aryl. Heteroaryl can be a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S. Aryl can be selected from phenyl, naphthyl, or anthracenyl. Aryl can be phenyl or naphthyl. Aryl can phenyl. Heteroaryl can be pyridyl, quinolyl, and isoquinolyl. R₆ can be H, -C₁₋₃alkylene-Ph or -C₁₋₃alkylene-Naphthyl. R₆ can be H or the side chain of an amino acid in **Table 1** or **Table 3**. R₆ can be H or an amino acid

residue having a side chain comprising an aromatic group. R₆ can be H, -CH₂Ph, or -CH₂Naphthyl. R₆ can be H or -CH₂Ph.

[0107] R₇ can be H, -alkylene-aryl, -alkylene-heteroaryl. R₇ can be H, -C₁₋₃alkylene-aryl, or -C₁₋₃alkylene-heteroaryl. R₇ can be H or -alkylene-aryl. R₇ can be H or -C₁₋₃alkylene-aryl. C₁₋₃alkylene can be a methylene. Aryl can be a 6- to 14-membered aryl. Heteroaryl can be a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, and S. Aryl can be selected from phenyl, naphthyl, or anthracenyl. Aryl can be phenyl or naphthyl. Aryl can be phenyl. Heteroaryl can be pyridyl, quinolyl, and isoquinolyl. R₇ can be H, -C₁₋₃alkylene-Ph or -C₁₋₃alkylene-Naphthyl. R₇ can be H or the side chain of an amino acid in Table 1 or Table 3. R₇ can be H or an amino acid residue having a side chain comprising an aromatic group. R₇ can be H, -CH₂Ph, or -CH₂Naphthyl. R₇ can be H or -CH₂Ph.

[0108] One, two or three of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph. One of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph. Two of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph. Three of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph. At least one of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph. No more than four of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph.

[0109] One, two or three of R₁, R₂, R₃, and R₄ are -CH₂Ph. One of R₁, R₂, R₃, and R₄ is -CH₂Ph. Two of R₁, R₂, R₃, and R₄ are -CH₂Ph. Three of R₁, R₂, R₃, and R₄ are -CH₂Ph. At least one of R₁, R₂, R₃, and R₄ is -CH₂Ph.

[0110] One, two or three of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be H. One of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be H. Two of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ are H. Three of R₁, R₂, R₃, R₅, R₆, and R₇ can be H. At least one of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be H. No more than three of R₁, R₂, R₃, R₄, R₅, R₆, and R₇ can be -CH₂Ph.

[0111] One, two or three of R₁, R₂, R₃, and R₄ are H. One of R₁, R₂, R₃, and R₄ is H. Two of R₁, R₂, R₃, and R₄ are H. Three of R₁, R₂, R₃, and R₄ are H. At least one of R₁, R₂, R₃, and R₄ is H.

[0112] At least one of R₄, R₅, R₆, and R₇ can be side chain of 3-guanidino-2-aminopropionic acid. At least one of R₄, R₅, R₆, and R₇ can be side chain of 4-guanidino-2-aminobutanoic acid. At least one of R₄, R₅, R₆, and R₇ can be side chain of arginine. At least one of R₄, R₅, R₆, and R₇ can be side chain of homoarginine. At least one of R₄, R₅, R₆, and R₇ can be side chain of N-methylarginine. At least one of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethylarginine. At least one of R₄, R₅, R₆, and R₇ can be side chain of 2,3-diaminopropionic acid. At least one of R₄, R₅, R₆, and R₇ can be side chain of 2,4-diaminobutanoic acid, lysine. At least one of R₄, R₅, R₆,

and R₇ can be side chain of N-methyllysine. At least one of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine. At least one of R₄, R₅, R₆, and R₇ can be side chain of N-ethyllysine. At least one of R₄, R₅, R₆, and R₇ can be side chain of N,N,N-trimethyllysine, 4-guanidinophenylalanine. At least one of R₄, R₅, R₆, and R₇ can be side chain of citrulline. At least one of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine, , β-homoarginine. At least one of R₄, R₅, R₆, and R₇ can be side chain of 3-(1-piperidinyl)alanine.

[0113] At least two of R₄, R₅, R₆, and R₇ can be side chain of 3-guanidino-2-aminopropionic acid. At least two of R₄, R₅, R₆, and R₇ can be side chain of 4-guanidino-2-aminobutanoic acid. At least two of R₄, R₅, R₆, and R₇ can be side chain of arginine. At least two of R₄, R₅, R₆, and R₇ can be side chain of homoarginine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N-methylarginine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethylarginine. At least two of R₄, R₅, R₆, and R₇ can be side chain of 2,3-diaminopropionic acid. At least two of R₄, R₅, R₆, and R₇ can be side chain of 2,4-diaminobutanoic acid, lysine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N-methyllysine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N-ethyllysine. At least two of R₄, R₅, R₆, and R₇ can be side chain of N,N,N-trimethyllysine, 4-guanidinophenylalanine. At least two of R₄, R₅, R₆, and R₇ can be side chain of citrulline. At least two of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine, , β-homoarginine. At least two of R₄, R₅, R₆, and R₇ can be side chain of 3-(1-piperidinyl)alanine.

[0114] At least three of R₄, R₅, R₆, and R₇ can be side chain of 3-guanidino-2-aminopropionic acid. At least three of R₄, R₅, R₆, and R₇ can be side chain of 4-guanidino-2-aminobutanoic acid. At least three of R₄, R₅, R₆, and R₇ can be side chain of arginine. At least three of R₄, R₅, R₆, and R₇ can be side chain of homoarginine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N-methylarginine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethylarginine. At least three of R₄, R₅, R₆, and R₇ can be side chain of 2,3-diaminopropionic acid. At least three of R₄, R₅, R₆, and R₇ can be side chain of 2,4-diaminobutanoic acid, lysine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N-methyllysine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N-ethyllysine. At least three of R₄, R₅, R₆, and R₇ can be side chain of N,N,N-trimethyllysine, 4-guanidinophenylalanine. At least three of R₄, R₅, R₆, and R₇ can be side chain of citrulline, . At

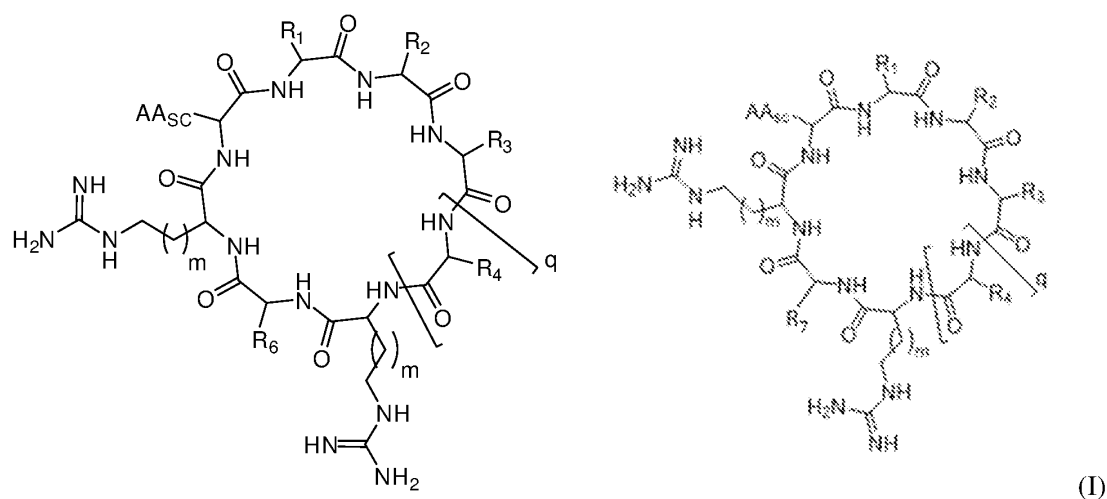
least three of R₄, R₅, R₆, and R₇ can be side chain of N,N-dimethyllysine, , β-homoarginine. At least three of R₄, R₅, R₆, and R₇ can be side chain of 3-(1-piperidiny)alanine.

[0115] AA_{SC} can be a side chain of a residue of asparagine, glutamine, or homoglutamine. AA_{SC} can be a side chain of a residue of glutamine. The cCPP can further comprise a linker conjugated the AA_{SC}, e.g., the residue of asparagine, glutamine, or homoglutamine. Hence, the cCPP can further comprise a linker conjugated to the asparagine, glutamine, or homoglutamine residue. The cCPP can further comprise a linker conjugated to the glutamine residue.

[0116] q can be 1, 2, or 3. q can be 1 or 2. q can be 1. q can be 2. q can be 3. q can be 4.

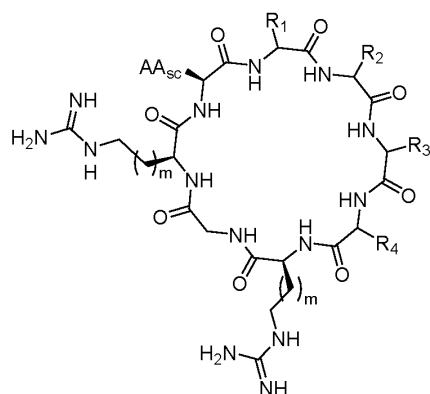
[0117] m can be 1-3. m can be 1 or 2. m can be 0. m can be 1. m can be 2. m can be 3.

10 [0118] The cCPP of Formula (A) can comprise the structure of Formula (I)

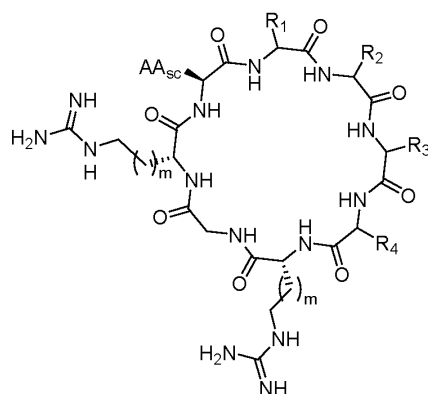


or protonated form thereof, wherein AA_{SC}, R₁, R₂, R₃, R₄, R₆, m and q are as defined herein

[0119] The cCPP of Formula (A) can comprise the structure of Formula (I-a) or Formula (I-b):



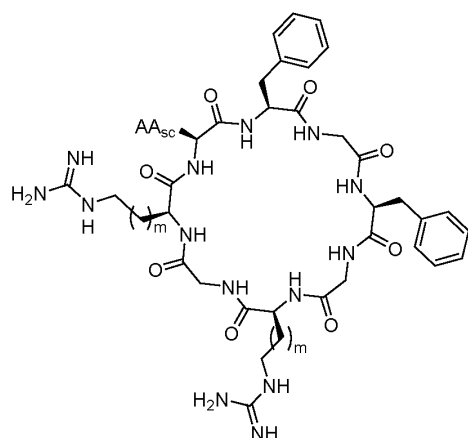
(I-a),



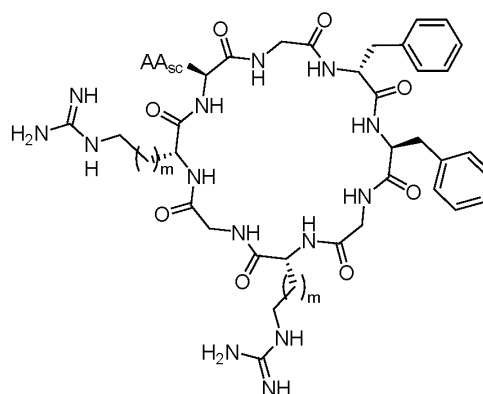
(I-b),

or protonated form thereof, wherein AA_{sc}, R₁, R₂, R₃, R₄, and m are as defined herein.

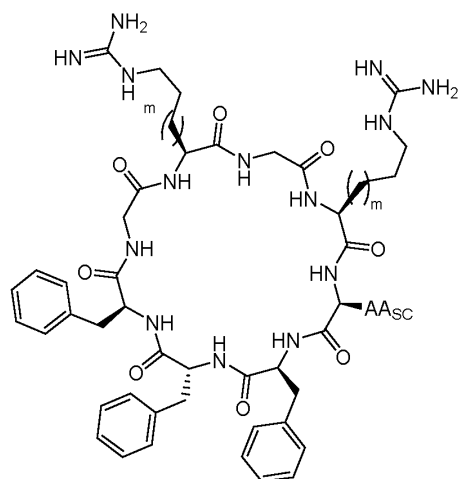
[0120] The cCPP of Formula (A) can comprise the structure of Formula (I-1), (I-2), (I-3) or (I-4):



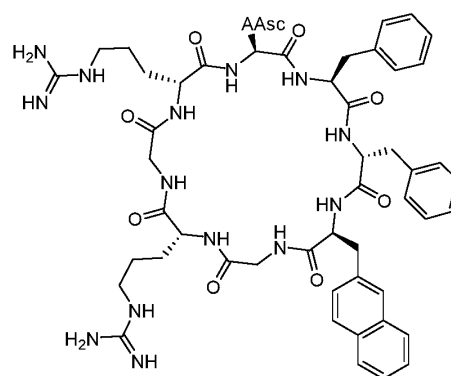
(I-1),



(I-2),



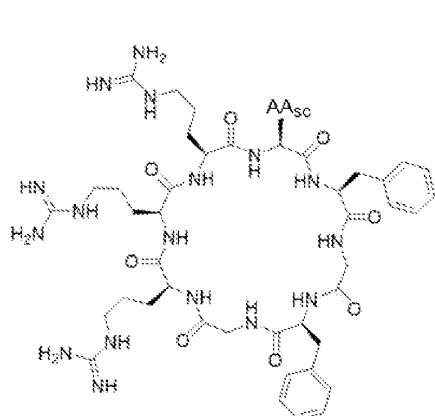
(I-3),



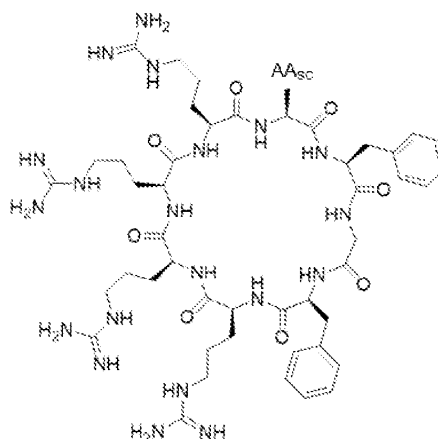
(I-4)

or protonated form thereof, wherein AA_{sc} and m are as defined herein.

[0121] The cCPP of Formula (A) can comprise the structure of Formula (I-5) or (I-6):



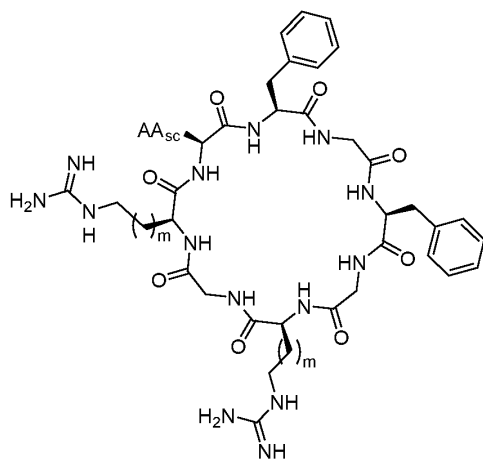
(I-5), or



(I-6)

or protonated form thereof, wherein AA_{sc} is as defined herein.

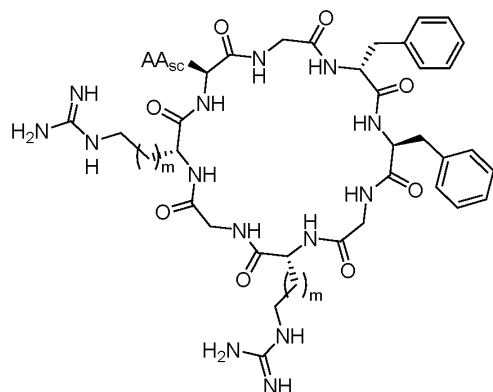
[0122] The cCPP of Formula (A) can comprise the structure of Formula (I-1):



(I-1), or a protonated form thereof,

wherein AA_{sc} and m are as defined herein.

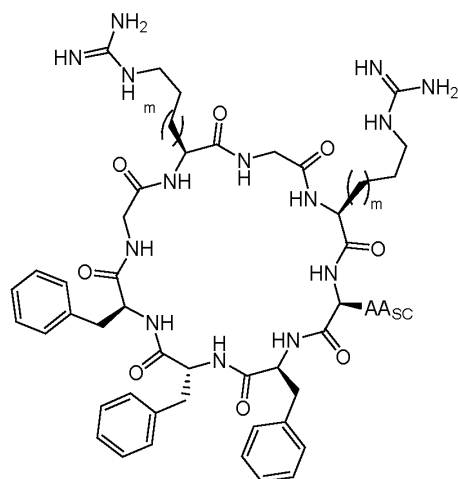
[0123] The cCPP of Formula (A) can comprise the structure of Formula (I-2):



(I-2), or a protonated form thereof,

wherein AA_{Sc} and m are as defined herein.

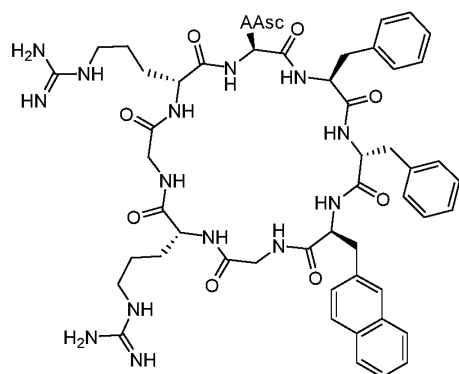
[0124] The cCPP of Formula (A) can comprise the structure of Formula (I-3):



(I-3), or a protonated form thereof,

5 wherein AA_{Sc} and m are as defined herein.

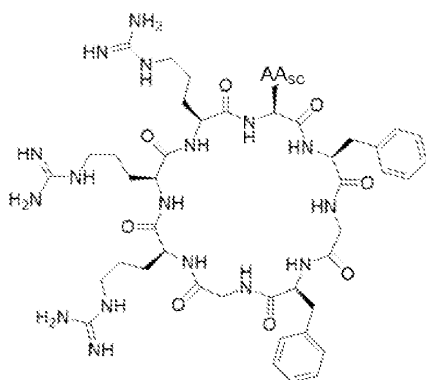
[0125] The cCPP of Formula (A) can comprise the structure of Formula (I-4):



(I-4), or a protonated form thereof,

wherein AAAsc and m are as defined herein.

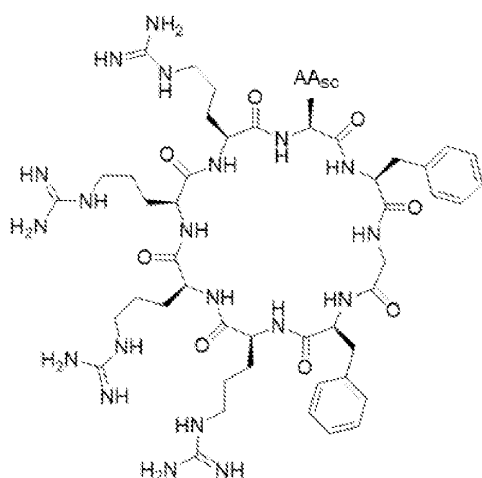
[0126] The cCPP of Formula (A) can comprise the structure of Formula (I-5):



(I-5), or a protonated form thereof,

5 wherein AAAsc and m are as defined herein.

[0127] The cCPP of Formula (A) can comprise the structure of Formula (I-6):

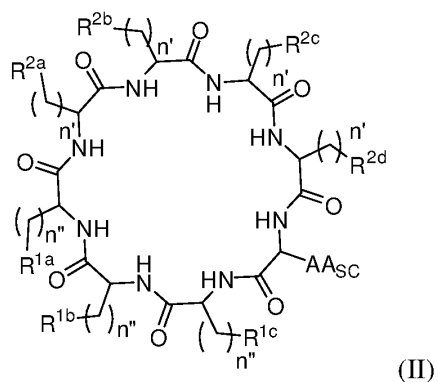


(I-6), or a protonated form thereof, wherein AAsc

and m are as defined herein.

- [0128] The cCPP can comprise one of the following sequences: FGFGRR; GfFGrr, FfΦGRGR; FfFGRR; or FfΦGrGr. The cCPP can have one of the following sequences: FGFΦ; 5 GfFGrrQ, FfΦGRGRQ; FfFGRRQ; or FfΦGrGrQ.

[0129] The disclosure also relates to a cCPP having the structure of Formula (II):



wherein:

AAsc is an amino acid side chain;

- 10 R^{1a}, R^{1b}, and R^{1c} are each independently a 6- to 14-membered aryl or a 6- to 14-membered heteroaryl;

R^{2a}, R^{2b}, R^{2c} and R^{2d} are independently an amino acid side chain;

at least one of R^{2a}, R^{2b}, R^{2c} and R^{2d} is













, or a protonated form thereof;

at least one of R^{2a} , R^{2b} , R^{2c} and R^{2d} is guanidine or a protonated form thereof;

each n'' is independently an integer 0, 1, 2, 3, 4, or 5;

5 each n' is independently an integer from 0, 1, 2, or 3; and

if n' is 0 then R^{2a} , R^{2b} , R^{2b} or R^{2d} is absent.

[0130] At least two of R^{2a}, R^{2b}, R^{2c} and R^{2d} can be



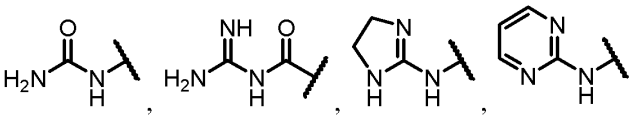
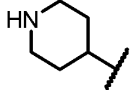
, or a protonated form thereof. Two or three of R^{2a}, R^{2b}, R^{2c} and

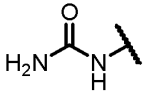
10 R^{2d} can be , or a protonated form thereof. One of R^{2a}, R^{2b}, R^{2c} and R^{2d} can be .

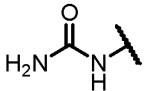

, or a protonated form thereof. At least one of R^{2a}

R^{2b}, R^{2c} and R^{2d} can be , or a protonated form thereof, and the remaining of R^{2a}, R^{2b}, R^{2c} and R^{2d} can be guanidine or a protonated form thereof. At least two of R^{2a}, R^{2b}, R^{2c} and R^{2d}

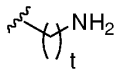
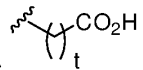
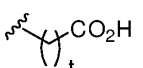
15 can be , or a protonated form thereof, and the remaining of R^{2a}, R^{2b}, R^{2c} and R^{2d} can be guanidine, or a protonated form thereof.

[0131] All of R^{2a} , R^{2b} , R^{2c} and R^{2d} can be , , , or a protonated form thereof. At least one of R^{2a} , R^{2b} , R^{2c} and R^{2d} can be

, or a protonated form thereof, and the remaining of R^{2a} , R^{2b} , R^{2c} and R^{2d} can be guanidine or a protonated form thereof. At least two R^{2a} , R^{2b} , R^{2c} and R^{2d} groups can be

5 , or a protonated form thereof, and the remaining of R^{2a} , R^{2b} , R^{2c} and R^{2d} are guanidine, or a protonated form thereof.

[0132] Each of R^{2a} , R^{2b} , R^{2c} and R^{2d} can independently be 2,3-diaminopropionic acid, 2,4-diaminobutyric acid, the side chains of ornithine, lysine, methyllysine, dimethyllysine, trimethyllysine, homo-lysine, serine, homo-serine, threonine, allo-threonine, histidine, 1-methylhistidine, 2-aminobutanedioic acid, aspartic acid, glutamic acid, or homo-glutamic acid.

[0133] AA_{SC} can be  or , wherein t can be an integer from 0 to 5. AA_{SC} can be , wherein t can be an integer from 0 to 5. t can be 1 to 5. t is 2 or 3. t can be 2. t can be 3.

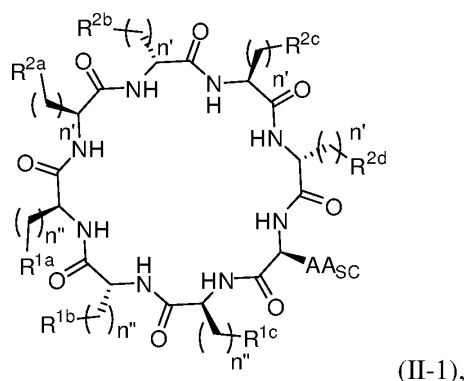
[0134] R^{1a} , R^{1b} , and R^{1c} can each independently be 6- to 14-membered aryl. R^{1a} , R^{1b} , and R^{1c} can be each independently a 6- to 14-membered heteroaryl having one or more heteroatoms selected from N, O, or S. R^{1a} , R^{1b} , and R^{1c} can each be independently selected from phenyl, naphthyl, anthracenyl, pyridyl, quinolyl, or isoquinolyl. R^{1a} , R^{1b} , and R^{1c} can each be independently selected from phenyl, naphthyl, or anthracenyl. R^{1a} , R^{1b} , and R^{1c} can each be independently phenyl or naphthyl. R^{1a} , R^{1b} , and R^{1c} can each be independently selected pyridyl, quinolyl, or isoquinolyl.

[0135] Each n' can independently be 1 or 2. Each n' can be 1. Each n' can be 2. At least one n' can be 0. At least one n' can be 1. At least one n' can be 2. At least one n' can be 3. At least one n' can be 4. At least one n' can be 5.

[0136] Each n'' can independently be an integer from 1 to 3. Each n'' can independently be 2 or 3. Each n'' can be 2. Each n'' can be 3. At least one n'' can be 0. At least one n'' can be 1. At least one n'' can be 2. At least one n'' can be 3.

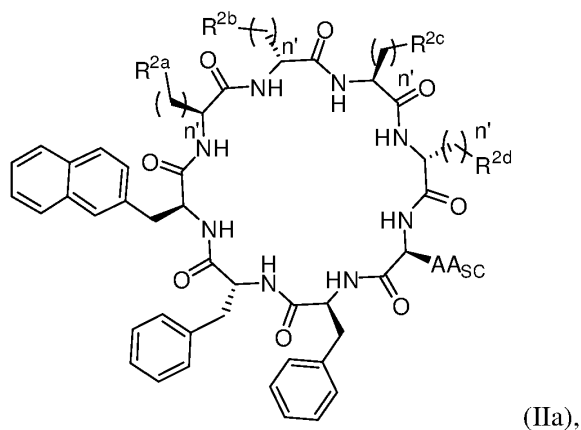
[0137] Each n'' can independently be 1 or 2 and each n' can independently be 2 or 3. Each n'' can be 1 and each n' can independently be 2 or 3. Each n'' can be 1 and each n' can be 2. Each n'' is 1 and each n' is 3.

[0138] The cCPP of Formula (II) can have the structure of Formula (II-1):



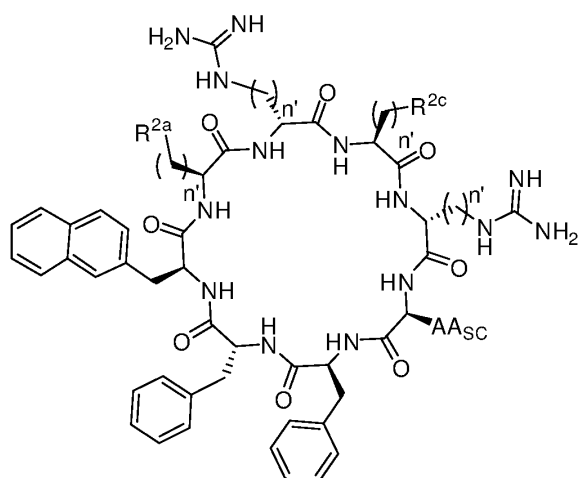
wherein R^{1a} , R^{1b} , R^{1c} , R^{2a} , R^{2b} , R^{2c} , R^{2d} , AA_{SC} , n' and n'' are as defined herein.

[0139] The cCPP of Formula (II) can have the structure of Formula (IIa):



wherein R^{1a} , R^{1b} , R^{1c} , R^{2a} , R^{2b} , R^{2c} , R^{2d} , AA_{SC} and n' are as defined herein.

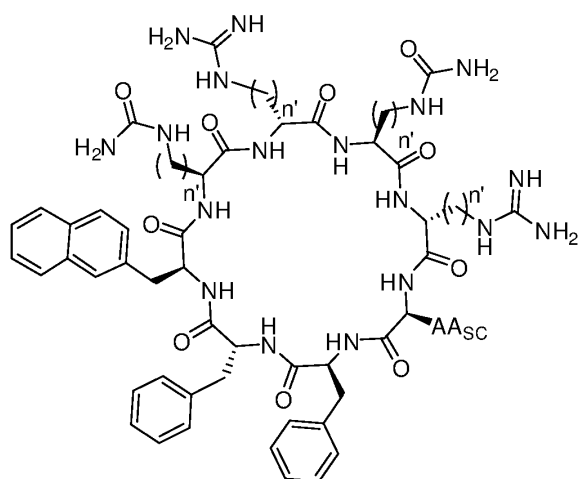
[0140] The cCPP of formula (II) can have the structure of Formula (IIb):



(IIb),

wherein R^{2a} , R^{2b} , AA_{Sc} , and n' are as defined herein.

[0141] The cCPP can have the structure of Formula (IIb):

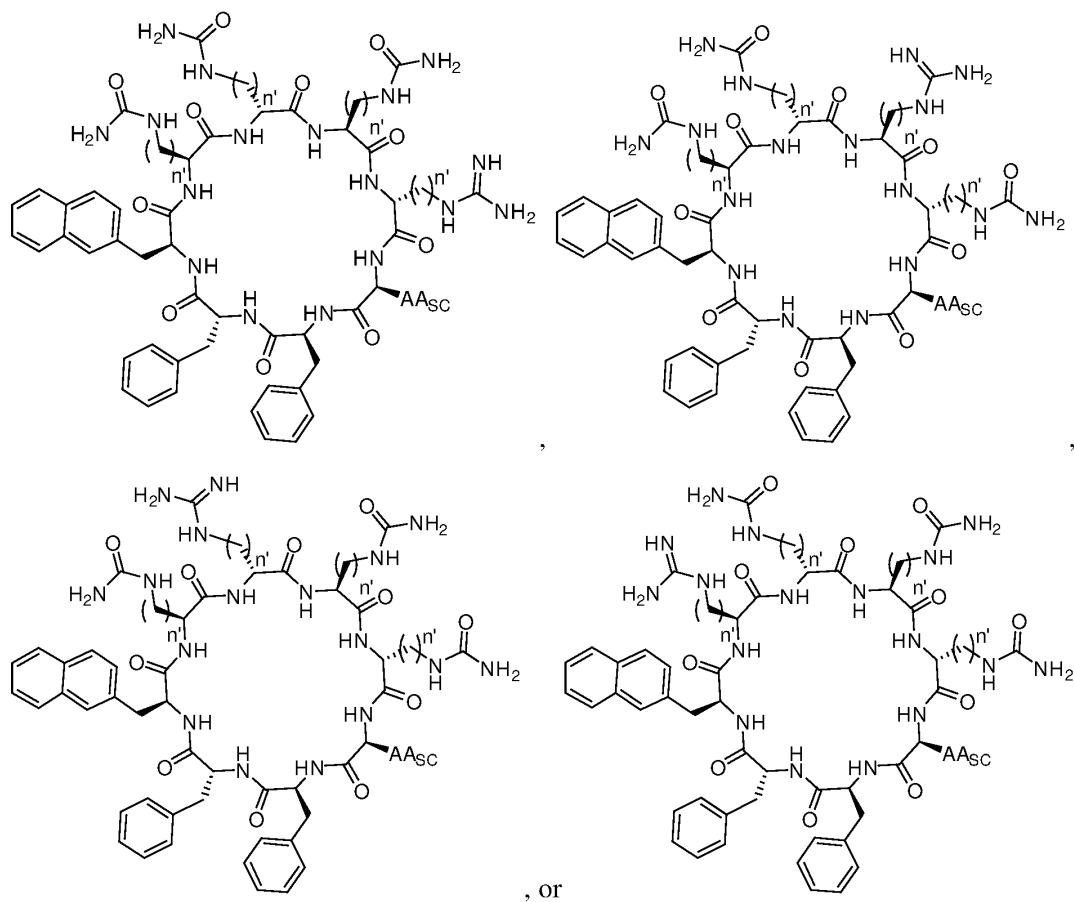


(IIc), or a protonated form thereof,

5 wherein:

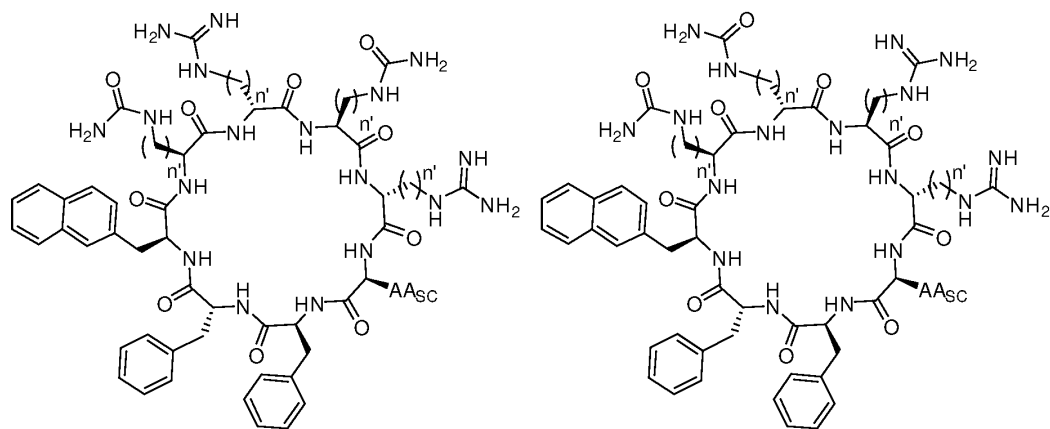
AA_{Sc} and n' are as defined herein.

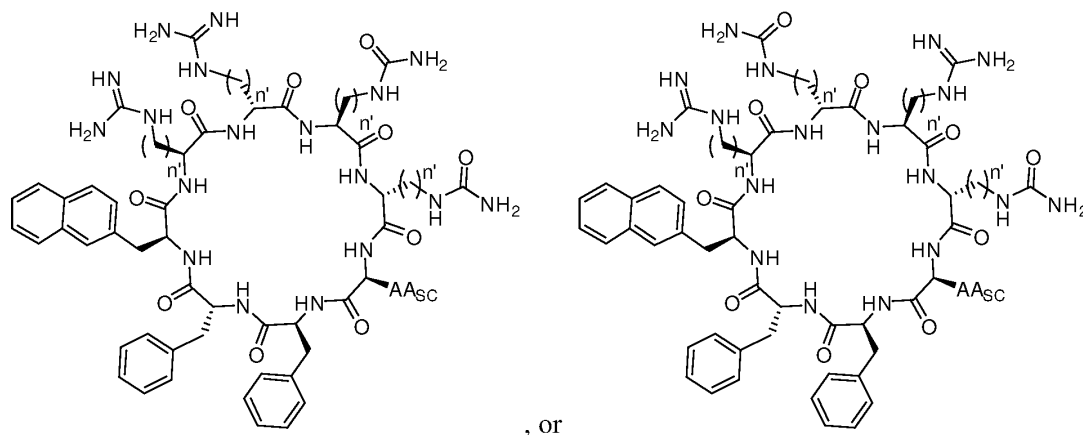
[0142] The cCPP of Formula (IIa) has one of the following structures:



wherein AA_{Sc} and n are as defined herein.

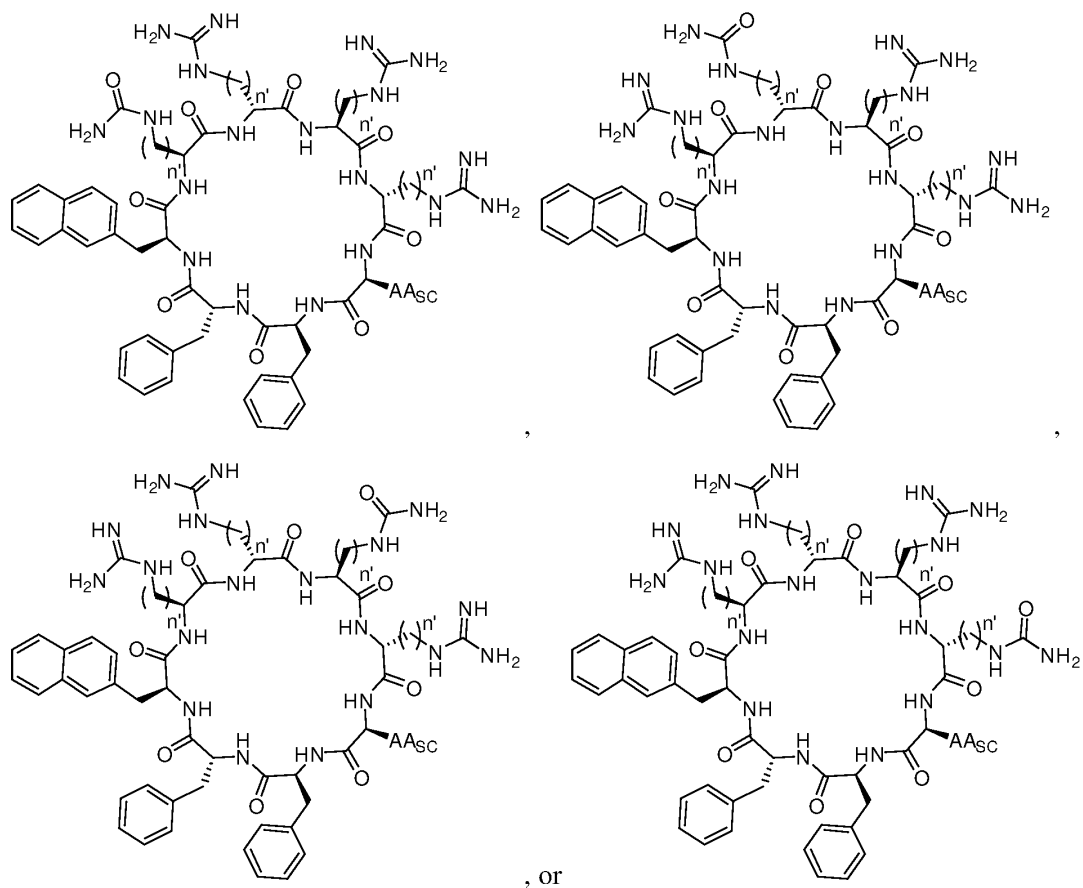
[0143] The cCPP of Formula (IIa) has one of the following structures:





wherein AA_{Sc} and n are as defined herein

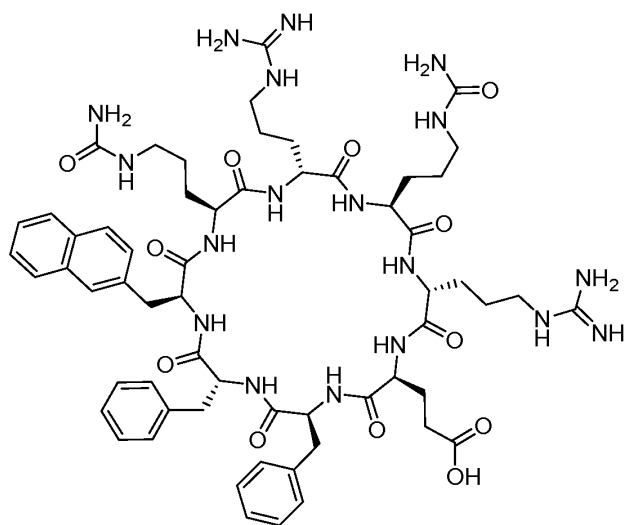
[0144] The cCPP of Formula (IIa) has one of the following structures:



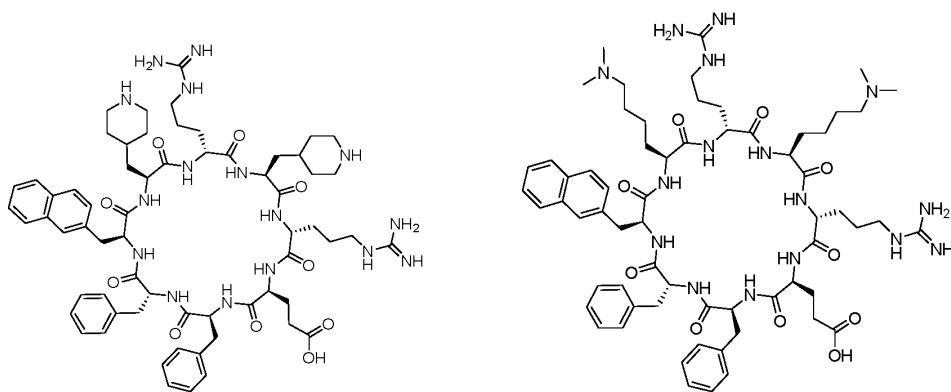
5

wherein AA_{Sc} and n are as defined herein.

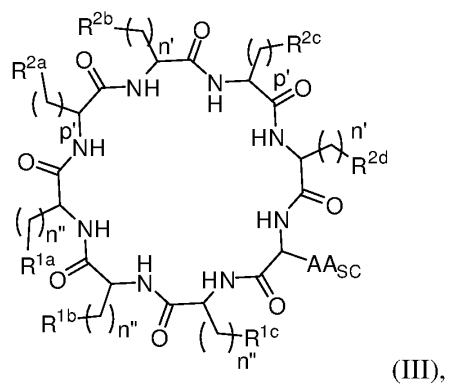
[0145] The cCPP of Formula (II) can have the structure:



[0146] The cCPP of Formula (II) can have the structure:



5 [0147] The cCPP can have the structure of Formula (III):

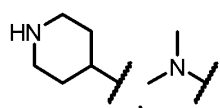
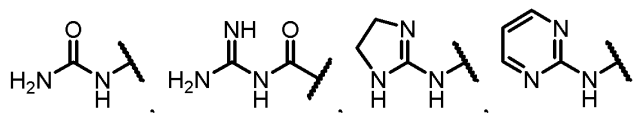


wherein:

AA_{SC} is an amino acid side chain;

R^{1a}, R^{1b}, and R^{1c} are each independently a 6- to 14-membered aryl or a 6- to 14-membered heteroaryl;

5 R^{2a} and R^{2c} are each independently H,



, or a protonated form thereof;

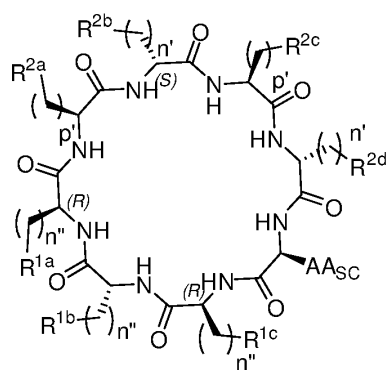
R^{2b} and R^{2d} are each independently guanidine or a protonated form thereof;

each n'' is independently an integer from 1 to 3;

each n' is independently an integer from 1 to 5; and

10 each p' is independently an integer from 0 to 5.

[0148] The cCPP of Formula (III) can have the structure of Formula (III-1):

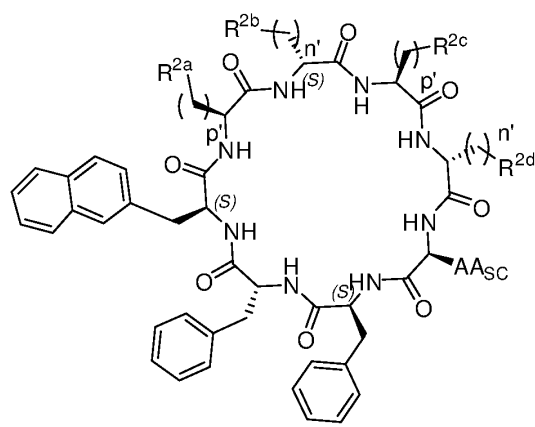


(III-1),

wherein:

AA_{SC}, R^{1a}, R^{1b}, R^{1c}, R^{2a}, R^{2c}, R^{2b}, R^{2d}, n', n'', and p' are as defined herein.

15 [0149] The cCPP of Formula (III) can have the structure of Formula (IIIa):



wherein:

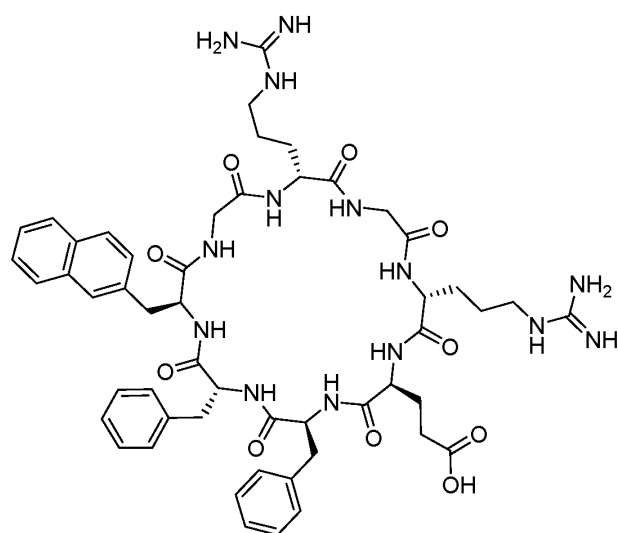
AA_{SC}, R^{2a}, R^{2c}, R^{2b}, R^{2d}, n', n'', and p' are as defined herein.

[0150] In Formulas (III), (III-1), and (IIIa), R^a and R^c can be H. R^a and R^c can be H and R^b and R^d can each independently be guanidine or protonated form thereof. R^a can be H. R^b can be H. p' can be 0. R^a and R^c can be H and each p' can be 0.

[0151] In Formulas (III), (III-1), and (IIIa), R^a and R^c can be H, R^b and R^d can each independently be guanidine or protonated form thereof, n'' can be 2 or 3, and each p' can be 0.

[0152] p' can 0. p' can 1. p' can 2. p' can 3. p' can 4. p' can be 5.

10 [0153] The cCPP can have the structure:



[0154] The cCPP of Formula (A) can be selected from:

CPP Sequence
(FfΦRrRrQ)
(FfΦCit-r-Cit-rQ)
(FfΦGrGrQ)
(FfFGRGRQ)
(FGFGRGRQ)
(GfFGrGrQ)
(FGFGRRRQ) or
(FGRRRRRQ)

[0155] The cCPP of Formula (A) can be selected from:

CPP Sequence
FΦRRRRQ
fΦRrRrQ
FfΦRrRrQ
FfΦCit-r-Cit-rQ
FfΦGrGrQ
FfΦRGRGQ
FfFGRGRQ
FGFGRGRQ
GfFGrGrQ
FGFGRRRQ or
FGRRRRRQ

[0156] AAsC can be conjugated to a linker.

5 **[0157]** In embodiments, the cCPP is selected from:

CPP sequence
FΦRRRQ
FΦRRRC
FΦRRRU
RRRΦFQ
RRRRΦF
FΦRRRR
FφrRrRq
FφrRrRQ
FΦRRRRQ
fΦRrRrQ
RRFRΦRQ
FRRRRΦQ
rRFRΦRQ
RRΦFRRQ
CRRRRFWQ
FfΦRrRrQ
FFΦRRRRQ

RFRFR Φ RQ
URRRRFWQ
CRRRRFWQ
F Φ RRRRQK
F Φ RRRRQC
f Φ RrRrRQ
F Φ RRRRRQ
RRRR Φ FD Ω C
F Φ RRR
FWRRR
RRR Φ F
RRRWF

Φ = L-naphthylalanine; ϕ = D-naphthylalanine; Ω = L-norleucine

[0158] In embodiments, the cCPP is not selected from:

CPP sequence
F Φ RRRQ
F Φ RRRC
F Φ RRRU
RRR Φ FQ
RRRR Φ F
F Φ RRRR
F ϕ rRrRq
F ϕ rRrRQ
F Φ RRRRQ
f Φ RrRrQ
RRFR Φ RQ
FRRRR Φ Q
rRFR Φ RQ
RR Φ FRRQ
CRRRRFWQ
Ff Φ RrRrQ
FF Φ RRRRQ
RFRFR Φ RQ
URRRRFWQ
CRRRRFWQ
F Φ RRRRQK
F Φ RRRRQC
f Φ RrRrRQ
F Φ RRRRRQ

RRRRΦFDΩC
FΦRRR
FWRRR
RRRΦF
RRRWF

Φ = L-naphthylalanine; φ = D-naphthylalanine; Ω = L-norleucine

Linker

[0159] The cCPP of the disclosure can be conjugated to a linker. The linker can link a cargo to the cCPP. The linker can be attached to the side chain of an amino acid of the cCPP, and the cargo can be attached at a suitable position on linker.

[0160] The linker can be any appropriate moiety which can conjugate a cCPP to one or more additional moieties, e.g., an exocyclic peptide (EP) and/or a cargo. Prior to conjugation to the cCPP and one or more additional moieties, the linker has two or more functional groups, each of which are independently capable of forming a covalent bond to the cCPP and one or more additional moieties. If the cargo is an oligonucleotide, the linker can be covalently bound to the 5' end of the cargo or the 3' end of the cargo. The linker can be covalently bound to the 5' end of the cargo. The linker can be covalently bound to the 3' end of the cargo. If the cargo is a peptide, the linker can be covalently bound to the N-terminus or the C-terminus of the cargo. The linker can be covalently bound to the backbone of the oligonucleotide or peptide cargo. The linker can be any appropriate moiety which conjugates a cCPP described herein to a cargo such as an oligonucleotide, peptide or small molecule.

[0161] The linker can comprise hydrocarbon linker.

[0162] The linker can comprise a cleavage site. The cleavage site can be a disulfide, or caspase-cleavage site (e.g., Val-Cit-PABC).

[0163] The linker can comprise: (i) one or more D or L amino acids, each of which is optionally substituted; (ii) optionally substituted alkylene; (iii) optionally substituted alkenylene; (iv) optionally substituted alkynylene; (v) optionally substituted carbocyclyl; (vi) optionally substituted heterocyclyl; (vii) one or more $-(R^1-J-R^2)z''$ - subunits, wherein each of R^1 and R^2 , at each instance, are independently selected from alkylene, alkenylene, alkynylene, carbocyclyl, and heterocyclyl, each J is independently C, NR^3 , $-NR^3C(O)-$, S, and O, wherein R^3 is independently selected from H, alkyl, alkenyl, alkynyl, carbocyclyl, and heterocyclyl, each of which is optionally substituted, and z'' is an integer from 1 to 50; (viii) $-(R^1-J)z''$ - or $-(J-R^1)z''$ -, wherein each of R^1 , at

each instance, is independently alkylene, alkenylene, alkynylene, carbocyclyl, or heterocyclyl, each J is independently C, NR^3 , $-\text{NR}^3\text{C}(\text{O})-$, S, or O, wherein R^3 is H, alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, each of which is optionally substituted, and z'' is an integer from 1 to 50; or (ix) the linker can comprise one or more of (i) through (x).

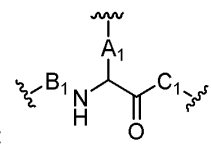
- 5 **[0164]** The linker can comprise one or more D or L amino acids and/or $-(\text{R}^1\text{-J-R}^2)z''$ -, wherein each of R^1 and R^2 , at each instance, are independently alkylene, each J is independently C, NR^3 , $-\text{NR}^3\text{C}(\text{O})-$, S, and O, wherein R^4 is independently selected from H and alkyl, and z'' is an integer from 1 to 50; or combinations thereof.

[0165] The linker can comprise a $-(\text{OCH}_2\text{CH}_2)_{z'}$ - (e.g., as a spacer), wherein z' is an integer from 10 1 to 23, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, or 23. “ $-(\text{OCH}_2\text{CH}_2)_{z'}$ ” can also be referred to as polyethylene glycol (PEG).

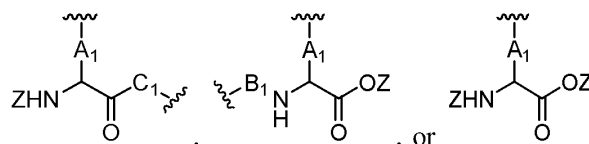
[0166] The linker can comprise one or more amino acids. The linker can comprise a peptide. The linker can comprise a $-(\text{OCH}_2\text{CH}_2)_{z'}$ -, wherein z' is an integer from 1 to 23, and a peptide. The peptide can comprise from 2 to 10 amino acids. The linker can further comprise a functional group 15 (FG) capable of reacting through click chemistry. FG can be an azide or alkyne, and a triazole is formed when the cargo is conjugated to the linker.

[0167] The linker can comprises (i) a β alanine residue and lysine residue; (ii) $-(\text{J-R}^1)z''$; or (iii) a combination thereof. Each R^1 can independently be alkylene, alkenylene, alkynylene, carbocyclyl, or heterocyclyl, each J is independently C, NR^3 , $-\text{NR}^3\text{C}(\text{O})-$, S, or O, wherein R^3 is H, alkyl, 20 alkenyl, alkynyl, carbocyclyl, or heterocyclyl, each of which is optionally substituted, and z'' can be an integer from 1 to 50. Each R^1 can be alkylene and each J can be O.

[0168] The linker can comprise (i) residues of β -alanine, glycine, lysine, 4-aminobutyric acid, 5-aminopentanoic acid, 6-aminohexanoic acid or combinations thereof; and (ii) $-(\text{R}^1\text{-J})z''$ - or $-(\text{J-R}^1)z''$. Each R^1 can independently be alkylene, alkenylene, alkynylene, carbocyclyl, or 25 heterocyclyl, each J is independently C, NR^3 , $-\text{NR}^3\text{C}(\text{O})-$, S, or O, wherein R^3 is H, alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, each of which is optionally substituted, and z'' can be an integer from 1 to 50. Each R^1 can be alkylene and each J can be O. The linker can comprise glycine, beta-alanine, 4-aminobutyric acid, 5-aminopentanoic acid, 6-aminohexanoic acid, or a combination thereof.



[0169] The linker can be a trivalent linker. The linker can have the structure:



, wherein A_1 , B_1 , and C_1 , can independently be a hydrocarbon linker (e.g., $\text{NRH}-(\text{CH}_2)_n-\text{COOH}$), a PEG linker (e.g., $\text{NRH}-(\text{CH}_2\text{O})_n-\text{COOH}$, wherein R is H , methyl or ethyl) or one or more amino acid residue, and Z is independently a protecting group. The linker can also incorporate a cleavage site, including a disulfide [$\text{NH}_2-(\text{CH}_2\text{O})_n-\text{S}-\text{S}-(\text{CH}_2\text{O})_n-\text{COOH}$], or caspase-cleavage site (Val-Cit-PABC).

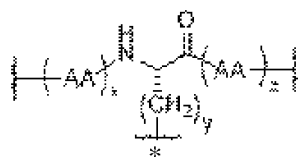
[0170] The hydrocarbon can be a residue of glycine or beta-alanine.

[0171] The linker can be bivalent and link the cCPP to a cargo. The linker can be bivalent and link the cCPP to an exocyclic peptide (EP).

10 [0172] The linker can be trivalent and link the cCPP to a cargo and to an EP.

[0173] The linker can be a bivalent or trivalent $\text{C}_1\text{-C}_{50}$ alkylene, wherein 1-25 methylene groups are optionally and independently replaced by $-\text{N}(\text{H})-$, $-\text{N}(\text{C}_1\text{-C}_4 \text{ alkyl})-$, $-\text{N}(\text{cycloalkyl})-$, $-\text{O}-$, $-\text{C}(\text{O})-$, $-\text{C}(\text{O})\text{O}-$, $-\text{S}-$, $-\text{S}(\text{O})-$, $-\text{S}(\text{O})_2-$, $-\text{S}(\text{O})_2\text{N}(\text{C}_1\text{-C}_4 \text{ alkyl})-$, $-\text{S}(\text{O})_2\text{N}(\text{cycloalkyl})-$, $-\text{N}(\text{H})\text{C}(\text{O})-$, $-\text{N}(\text{C}_1\text{-C}_4 \text{ alkyl})\text{C}(\text{O})-$, $-\text{N}(\text{cycloalkyl})\text{C}(\text{O})-$, $-\text{C}(\text{O})\text{N}(\text{H})-$, $-\text{C}(\text{O})\text{N}(\text{C}_1\text{-C}_4 \text{ alkyl})-$, $-\text{C}(\text{O})\text{N}(\text{cycloalkyl})$, aryl, heterocyclyl, heteroaryl, cycloalkyl, or cycloalkenyl. The linker can be a bivalent or trivalent $\text{C}_1\text{-C}_{50}$ alkylene, wherein 1-25 methylene groups are optionally and independently replaced by $-\text{N}(\text{H})-$, $-\text{O}-$, $-\text{C}(\text{O})\text{N}(\text{H})-$, or a combination thereof.

[0174] The linker can have the structure:

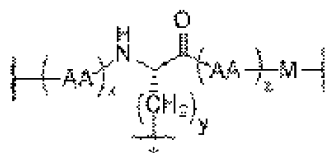


20 , wherein: each AA is independently an amino acid residue; * is the point of attachment to the AA_{SC} , and AA_{SC} is side chain of an amino acid residue of the cCPP ; x is an integer from 1-10; y is an integer from 1-5; and z is an integer from 1-10. x can be an integer from 1-5. x can be an integer from 1-3. x can be 1. y can be an integer from 2-4. y can be 4. z can be an integer from 1-5. z can be an integer from 1-3. z can be 1. Each AA can

independently be selected from glycine, β -alanine, 4-aminobutyric acid, 5-aminopentanoic acid, and 6-aminohexanoic acid.

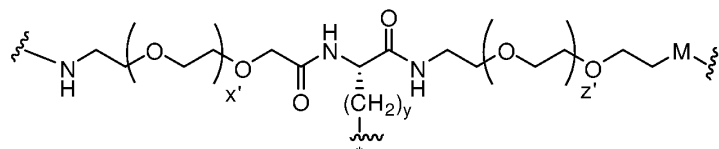
[0175] The cCPP can be attached to the cargo through a linker ("L"). The linker can be conjugated to the cargo through a bonding group ("M").

5 [0176] The linker can have the structure:



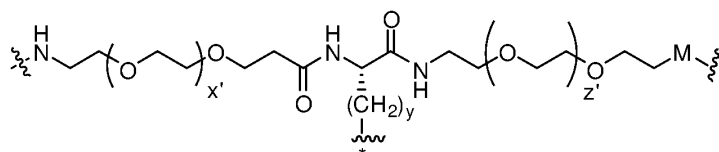
, wherein: x is an integer from 1-10; y is an integer from 1-5; z is an integer from 1-10; each AA is independently an amino acid residue; * is the point of attachment to the AA_{sc}, and AA_{sc} is side chain of an amino acid residue of the cCPP; and M is a bonding group defined herein.

10 [0177] The linker can have the structure:



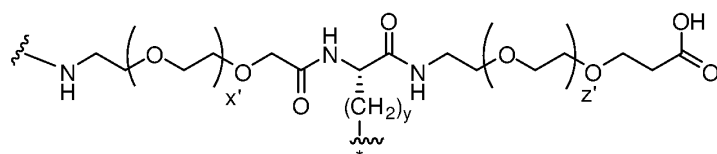
wherein: x' is an integer from 1-23; y is an integer from 1-5; z' is an integer from 1-23; * is the point of attachment to the AA_{sc}, and AA_{sc} is a side chain of an amino acid residue of the cCPP; and M is a bonding group defined herein.

15 [0178] The linker can have the structure:



wherein: x' is an integer from 1-23; y is an integer from 1-5; z' is an integer from 1-23; * is the point of attachment to the AA_{sc}, and AA_{sc} is a side chain of an amino acid residue of the cCPP; and M is a bonding group defined herein.

20 [0179] The linker can have the structure:



wherein: x' is an integer from 1-23; y is an integer from 1-5; and z' is an integer from 1-23; $*$ is the point of attachment to the AA_{SC}, and AA_{SC} is a side chain of an amino acid residue of the cCPP.

5 [0180] x can be an integer from 1-10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, inclusive of all ranges and subranges therebetween.

[0181] x' can be an integer from 1-23, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, or 23, inclusive of all ranges and subranges therebetween. x' can be an integer from 5-15. x' can be an integer from 9-13. x' can be an integer from 1-5. x' can be 1.

10 [0182] y can be an integer from 1-5, e.g., 1, 2, 3, 4, or 5, inclusive of all ranges and subranges therebetween. y can be an integer from 2-5. y can be an integer from 3-5. y can be 3 or 4. y can be 4 or 5. y can be 3. y can be 4. y can be 5.

[0183] z can be an integer from 1-10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10, inclusive of all ranges and subranges therebetween.

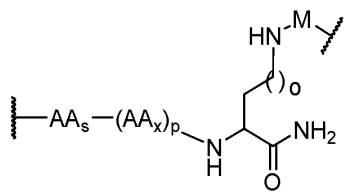
15 [0184] z' can be an integer from 1-23, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, or 23, inclusive of all ranges and subranges therebetween. z' can be an integer from 5-15. z' can be an integer from 9-13. z' can be 11.

[0185] As discussed above, the linker or M (wherein M is part of the linker) can be covalently bound to cargo at any suitable location on the cargo. The linker or M (wherein M is part of the linker) can be covalently bound to the 3' end of oligonucleotide cargo or the 5' end of an oligonucleotide cargo. The linker or M (wherein M is part of the linker) can be covalently bound to the N-terminus or the C-terminus of a peptide cargo. The linker or M (wherein M is part of the linker) can be covalently bound to the backbone of an oligonucleotide or a peptide cargo.

25 [0186] The linker can be bound to the side chain of aspartic acid, glutamic acid, glutamine, asparagine, or lysine, or a modified side chain of glutamine or asparagine (e.g., a reduced side chain having an amino group), on the cCPP. The linker can be bound to the side chain of lysine on the cCPP.

[0187] The linker can be bound to the side chain of aspartic acid, glutamic acid, glutamine, asparagine, or lysine, or a modified side chain of glutamine or asparagine (e.g., a reduced side chain having an amino group), on a peptide cargo. The linker can be bound to the side chain of lysine on the peptide cargo.

5 [0188] The linker can have a structure:



wherein

M is a group that conjugates L to a cargo, for example, an oligonucleotide;

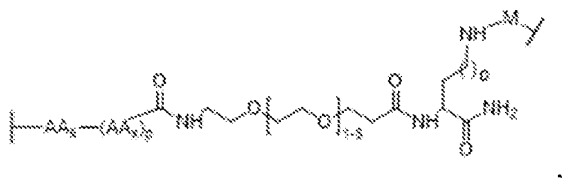
AA_s is a side chain or terminus of an amino acid on the cCPP;

10 each AA_x is independently an amino acid residue;

o is an integer from 0 to 10; and

p is an integer from 0 to 5.

[0189] The linker can have a structure:



15 wherein

M is a group that conjugates L to a cargo, for example, an oligonucleotide;

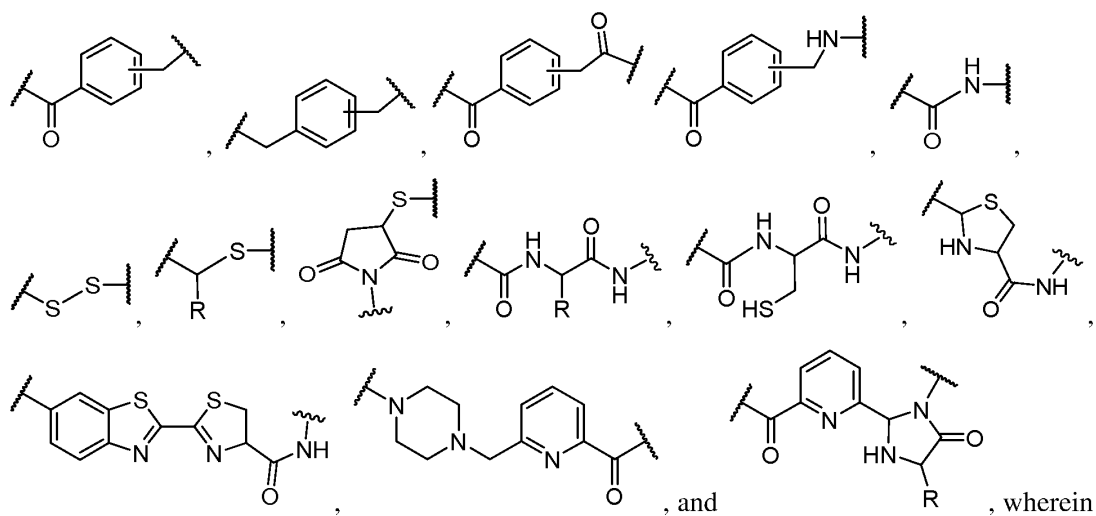
AA_s is a side chain or terminus of an amino acid on the cCPP;

each AA_x is independently an amino acid residue;

o is an integer from 0 to 10; and

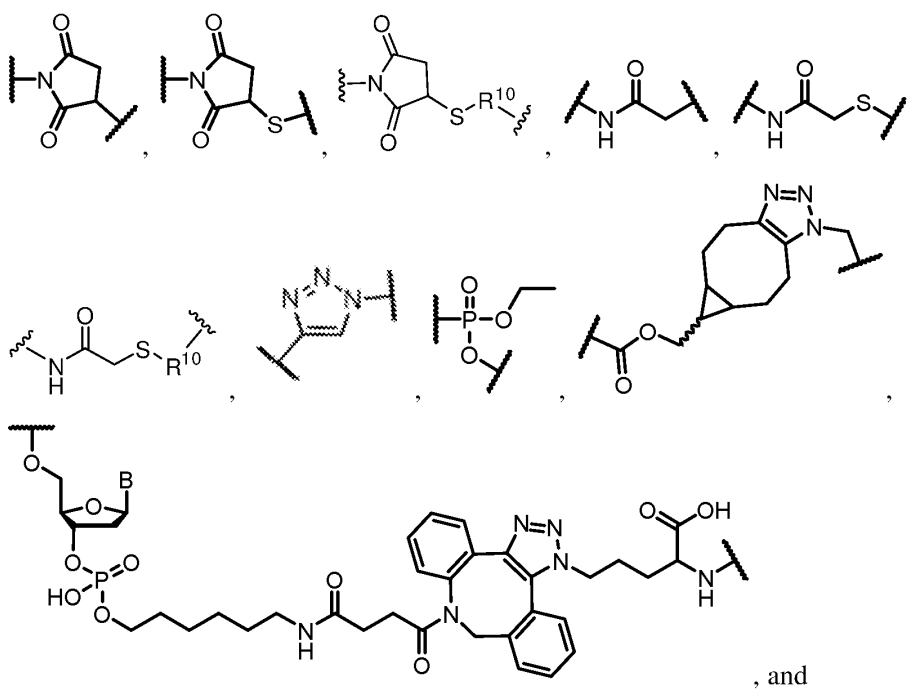
20 p is an integer from 0 to 5.

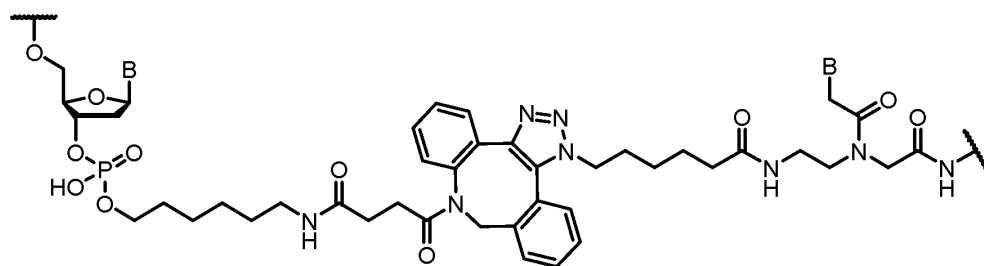
[0190] M can comprise an alkylene, alkenylene, alkynylene, carbocyclyl, or heterocyclyl, each of which is optionally substituted. M can be selected from:

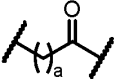


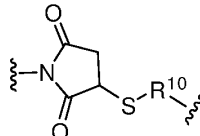
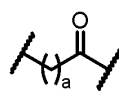
R is alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl.

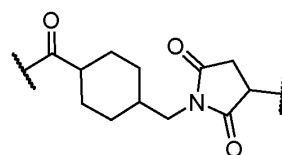
5 **[0191]** M can be selected from:

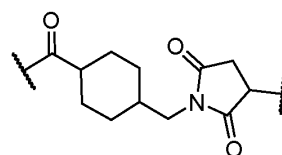




wherein: R^{10} is alkylene, cycloalkyl, or , wherein a is 0 to 10.

[0192] M can be , R^{10} can be , and a is 0 to 10. M can be .



[0193] M can be a heterobifunctional crosslinker, e.g., , which is disclosed in Williams et al. *Curr. Protoc Nucleic Acid Chem.* **2010**, 42, 4.41.1-4.41.20, incorporated herein by reference its entirety.

[0194] M can be -C(O)-.

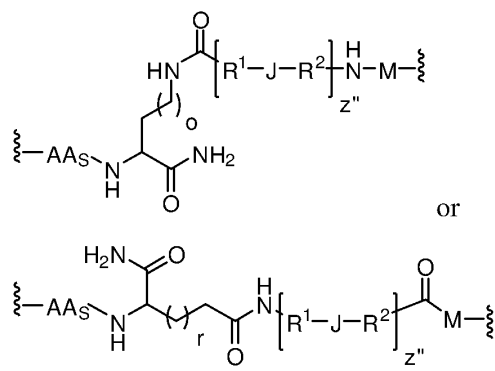
[0195] AA_s can be a side chain or terminus of an amino acid on the cCPP. Non-limiting examples of AA_s include aspartic acid, glutamic acid, glutamine, asparagine, or lysine, or a modified side chain of glutamine or asparagine (e.g., a reduced side chain having an amino group). AA_s can be an AA_{SC} as defined herein.

[0196] Each AA_x is independently a natural or non-natural amino acid. One or more AA_x can be a natural amino acid. One or more AA_x can be a non-natural amino acid. One or more AA_x can be a β -amino acid. The β -amino acid can be β -alanine.

[0197] o can be an integer from 0 to 10, e.g., 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. o can be 0, 1, 2, or 3. o can be 0. o can be 1. o can be 2. o can be 3.

[0198] p can be 0 to 5, e.g., 0, 1, 2, 3, 4, or 5. p can be 0. p can be 1. p can be 2. p can be 3. p can be 4. p can be 5.

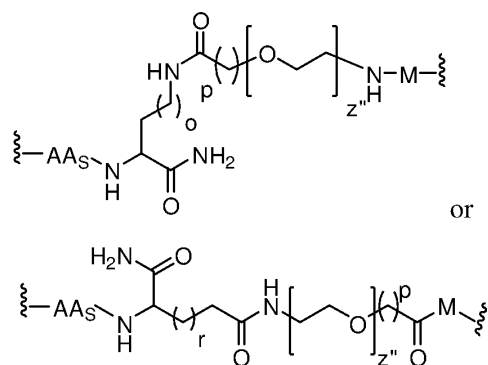
[0199] The linker can have the structure:



wherein M, AA_s, each $-(R^1-J-R^2)_{z''}$, o and z'' are defined herein; r can be 0 or 1.

[0200] r can be 0. r can be 1.

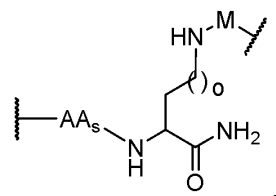
5 **[0201]** The linker can have the structure:



wherein each of M, AA_s, o, p, q, r and z'' can be as defined herein.

[0202] z'' can be an integer from 1 to 50, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, and 50, inclusive of all ranges and values therebetween. z'' can be an integer from 5-20. z'' can be an integer from 10-15.

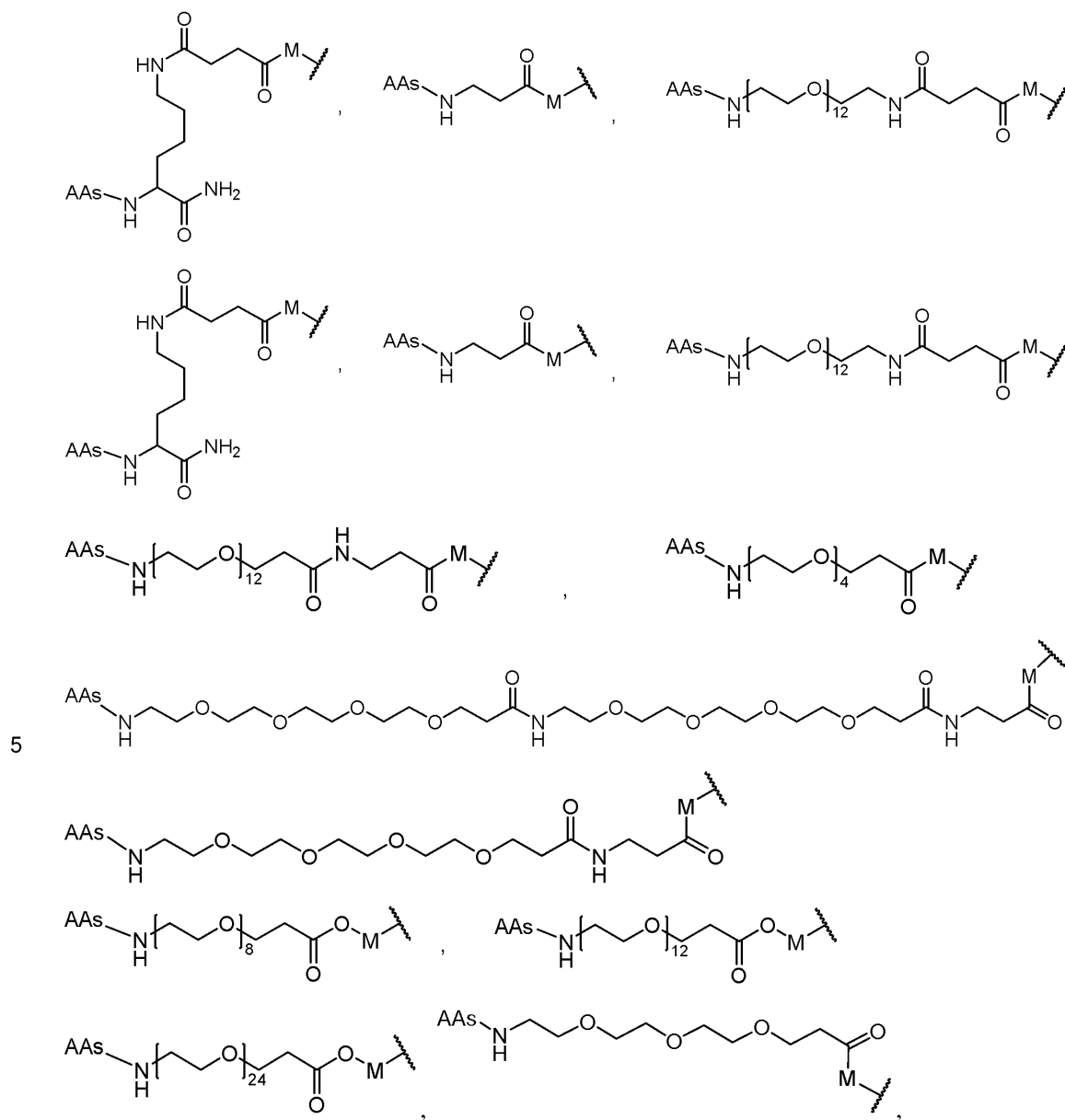
[0203] The linker can have the structure:

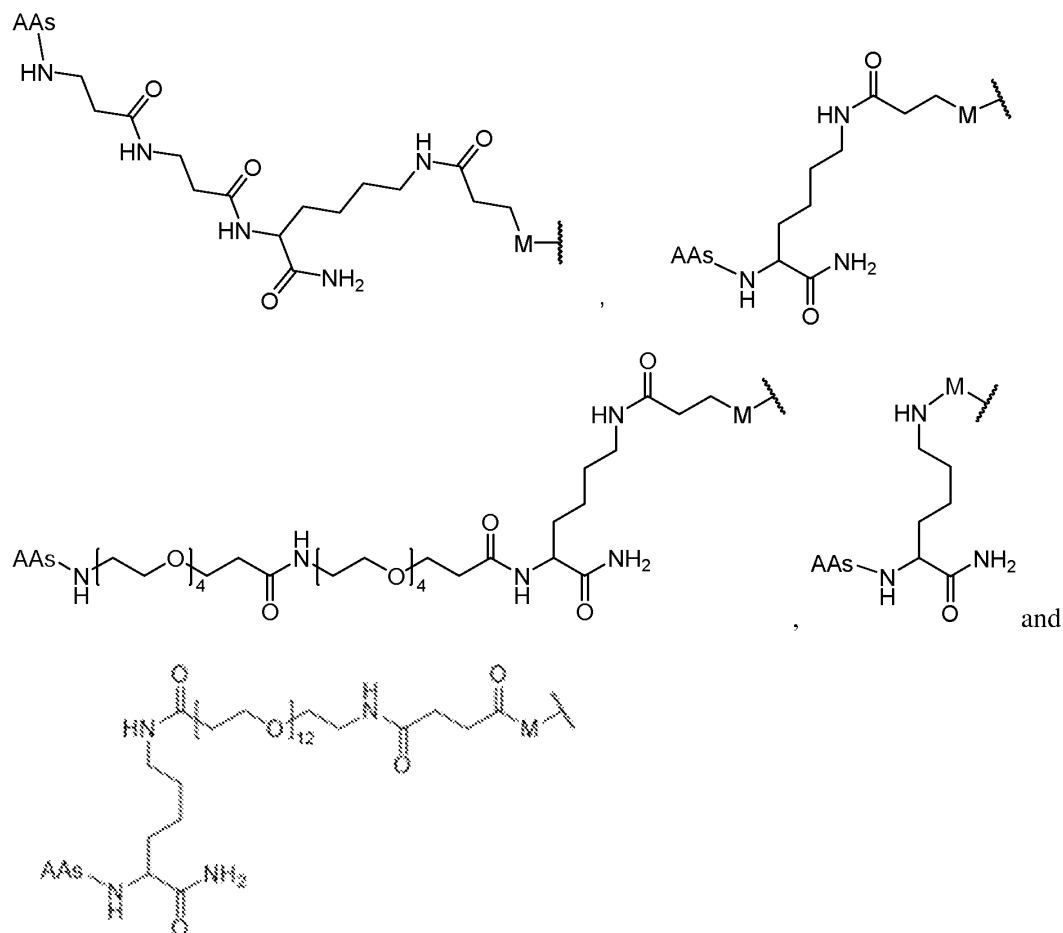


15 wherein:

M, AA_s and o are as defined herein.

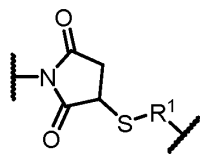
[0204] Other non-limiting examples of suitable linkers include:





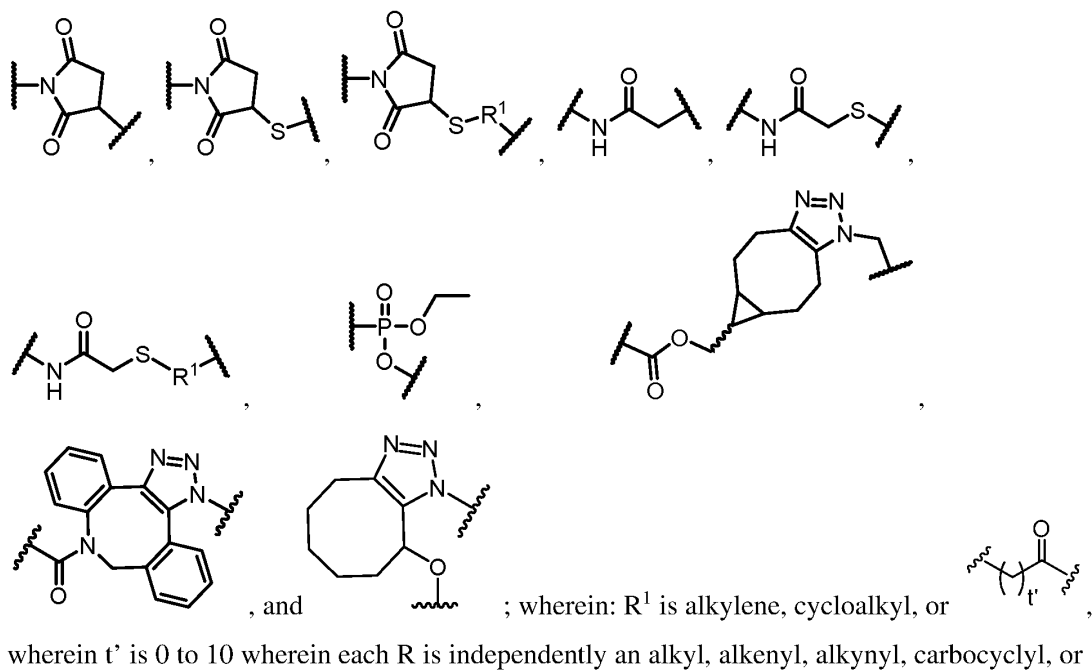
wherein M and AAs are as defined herein.

- 5 **[0205]** Provided herein is a compound comprising a cCPP and an AC that is complementary to a target in a pre-mRNA sequence further comprising L, wherein the linker is conjugated to the AC



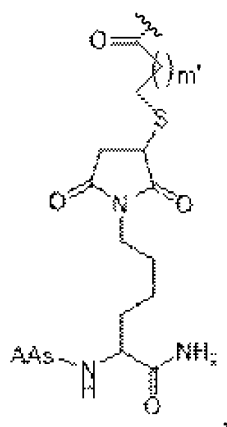
through a bonding group (M), wherein M is

- [0206]** Provided herein is a compound comprising a cCPP and a cargo that comprises an antisense compound (AC), for example, an antisense oligonucleotide, that is complementary to a target in a pre-mRNA sequence, wherein the compound further comprises L, wherein the linker is conjugated to the AC through a bonding group (M), wherein M is selected from:
- 10



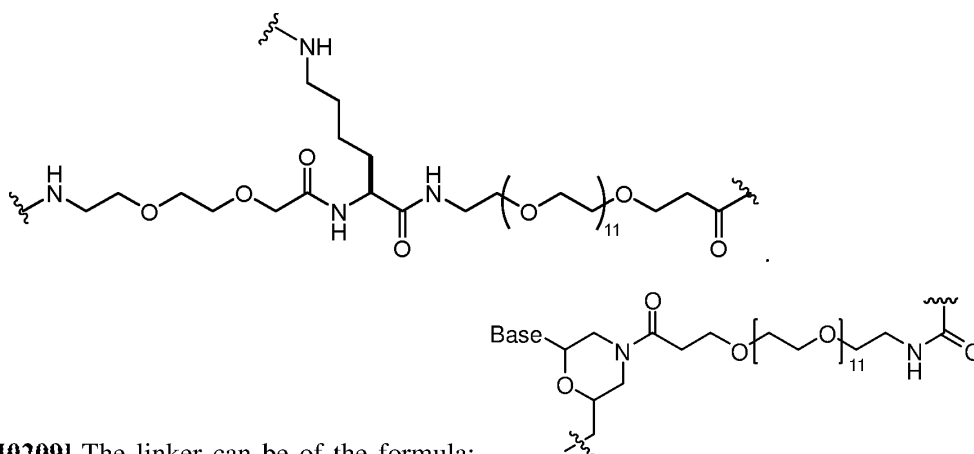
5 heterocyclyl, wherein R¹ is  , and t' is 2.

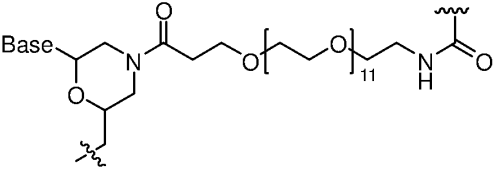
[0207] The linker can have the structure:



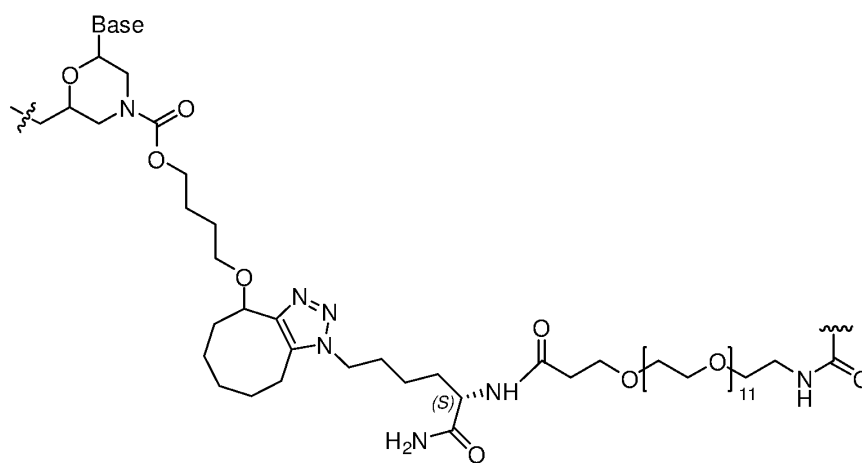
wherein AA_s is as defined herein, and m' is 0-10.

[0208] The linker can be of the formula:



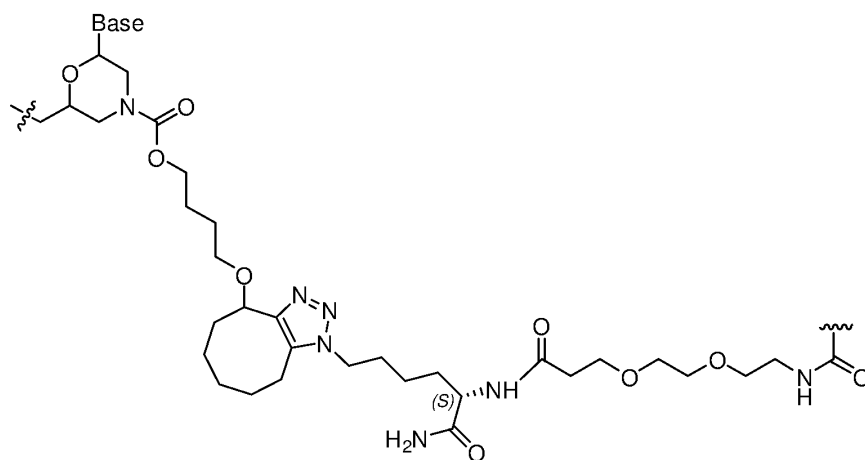
[0209] The linker can be of the formula: , wherein "base" corresponds to a nucleobase at the 3' end of a cargo phosphorodiamidate morpholino oligomer.

5 [0210] The linker can be of the formula:



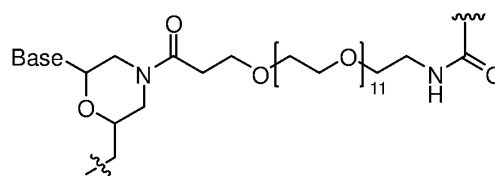
, wherein "base" corresponds to a nucleobase at the 3' end of a cargo phosphorodiamidate morpholino oligomer.

[0211] The linker can be of the formula:



, wherein “base”

corresponds to a nucleobase at the 3' end of a cargo phosphorodiamidate morpholino oligomer..



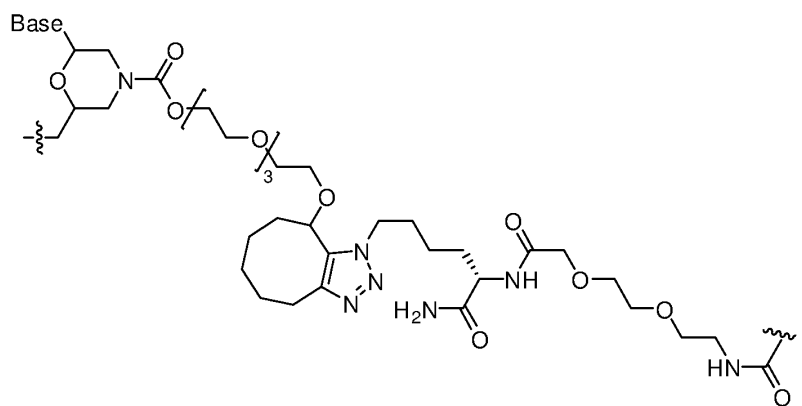
[0212] The linker can be of the formula:

, wherein

“base” corresponds to a nucleobase at the 3' end of a cargo phosphorodiamidate morpholino oligomer.

5

[0213] The linker can be of the formula:



[0214] The linker can be covalently bound to a cargo at any suitable location on the cargo. The linker is covalently bound to the 3' end of cargo or the 5' end of an oligonucleotide cargo. The linker can be covalently bound to the backbone of a cargo.

10

[0215] The linker can be bound to the side chain of aspartic acid, glutamic acid, glutamine, asparagine, or lysine, or a modified side chain of glutamine or asparagine (e.g., a reduced side chain having an amino group), on the cCPP. The linker can be bound to the side chain of lysine on the cCPP.

5 *cCPP-linker conjugates*

[0216] The cCPP can be conjugated to a linker defined herein. The linker can be conjugated to an AA_{SC} of the cCPP as defined herein.

[0217] The linker can comprise a $-(\text{OCH}_2\text{CH}_2)_{z'}$ - subunit (e.g., as a spacer), wherein z' is an integer from 1 to 23, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23. “-
10 $(\text{OCH}_2\text{CH}_2)_{z'}$ ” is also referred to as PEG. The cCPP-linker conjugate can have a structure selected from **Table 4**:

Table 4: Examples of cCPP linker conjugates

<i>cyclo</i> (FfΦ-4gp-r-4gp-rQ)-PEG ₄ -K-NH ₂
<i>cyclo</i> (FfΦ-Cit-r-Cit-rQ)-PEG ₄ -K-NH ₂
<i>cyclo</i> (FfΦ-Pia-r-Pia-rQ)-PEG ₄ -K-NH ₂
<i>cyclo</i> (FfΦ-Dml-r-Dml-rQ)-PEG ₄ -K-NH ₂
<i>cyclo</i> (FfΦ-Cit-r-Cit-rQ)-PEG ₁₂ -OH
<i>cyclo</i> (fΦR-Cit-R-Cit-Q)-PEG ₁₂ -OH

[0218] The linker can comprise a $-(\text{OCH}_2\text{CH}_2)_{z'}$ - subunit, wherein z' is an integer from 1 to 23, and a peptide subunit. The peptide subunit can comprise from 2 to 10 amino acids. The cCPP-
15 linker conjugate can have a structure selected from **Table 5**:

Table 5: Examples of cCPP-linker conjugates

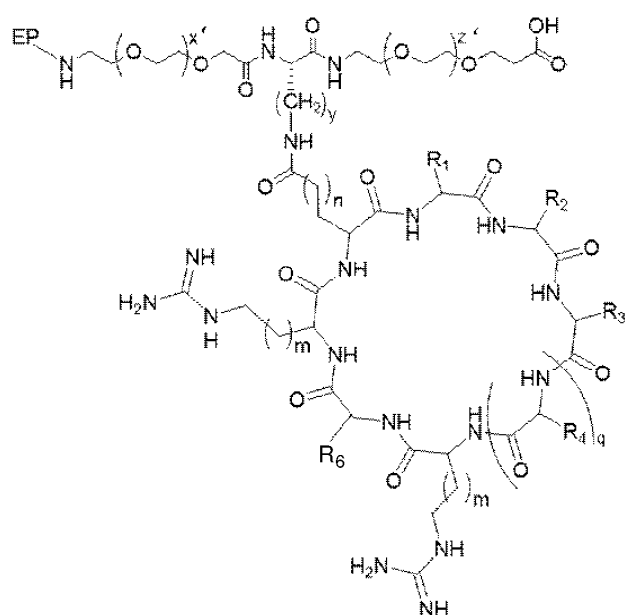
Ac-PKKKRKV-Lys(<i>cyclo</i> [FfΦ-R-r-Cit-rQ])-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-Lys(<i>cyclo</i> [FfΦ-Cit-r-R-rQ])-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-K(<i>cyclo</i> (FfΦR-cit-R-cit-Q))-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-B-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG2-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG4-k(N ₃)-NH ₂
Ac-PKKKRKV-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂
Ac-pkkkrkv-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂
Ac-rrv-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-OH
Ac-PKKKRKV-PEG2-Lys(<i>cyclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂

Ac-PKKK-Cit-KV-PEG2-Lys(cyclo[FfΦ-Cit-r-Cit-r-Q])-PEG12-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(cyclo[FfΦ-Cit-r-Cit-r-Q])-PEG12-K(N ₃)-NH ₂

[0219] The cCPP-linker conjugate can have a structure shown in Figure 1 (e.g., Compound 1a, Compound 1b, Compound 2a, or Compound 3a) or sequence listed in **Table 4**.

[0220] The cCPP-linker conjugate can have a sequence as listed in **Table 5**.

- 5 [0221] The cCPP-linker conjugate can be Ac-PKKKRKV-K(cyclo[FfΦGrGrQ])-PEG12-K(N₃)-NH₂. EEVs comprising a cyclic cell penetrating peptide (cCPP), linker and exocyclic peptide (EP) are provided. An EEV can comprise the structure of Formula (B):



(B), or a protonated form thereof,

wherein:

- 10 R₁, R₂, and R₃ are each independently H or an aromatic or heteroaromatic side chain of an amino acid;
- R₄ and R₆ are independently H or an amino acid side chain;
- EP is an exocyclic peptide as defined herein;
- each m is independently an integer from 0-3;
- 15 n is an integer from 0-2;
- x' is an integer from 1-20;
- y is an integer from 1-5;

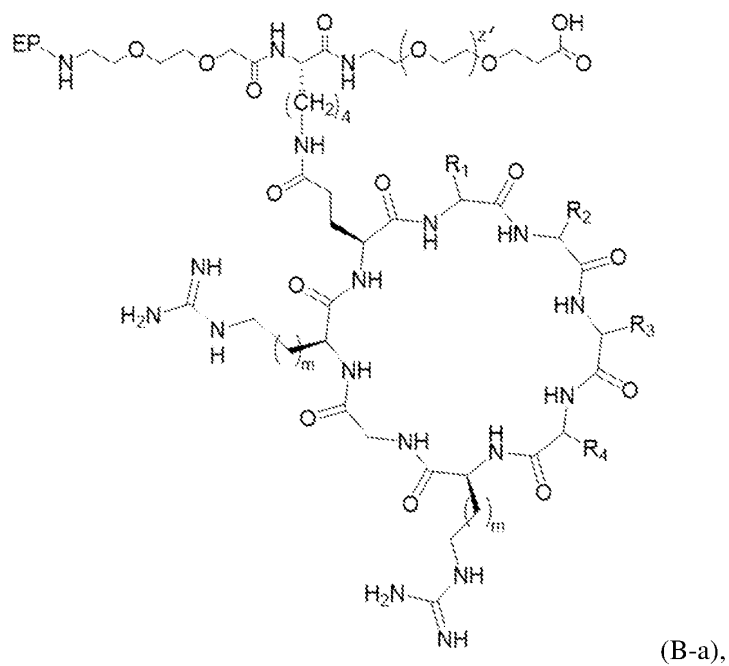
q is 1-4; and

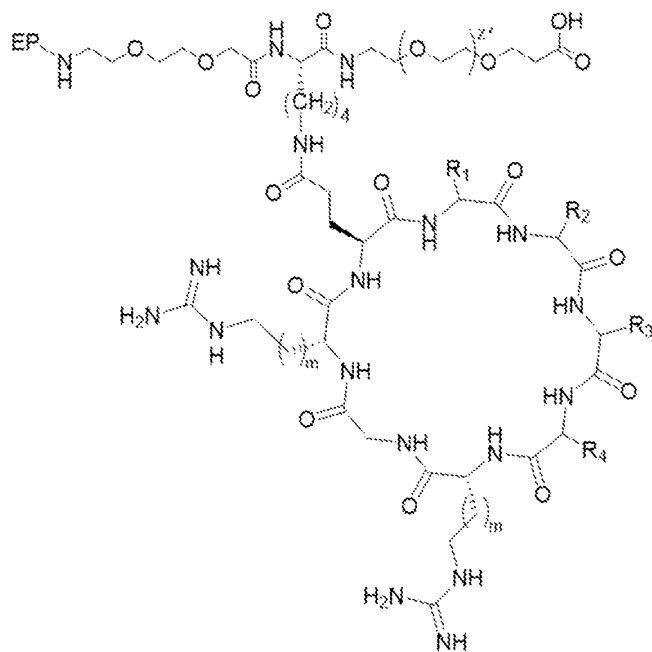
z' is an integer from 1-23.

[0222] R₁, R₂, R₃, R₄, R₇, EP, m, q, y, x', z' are as described herein.

[0223] n can be 0. n can be 1. n can be 2.

5 [0224] The EEV can comprise the structure of Formula (B-a) or (B-b):

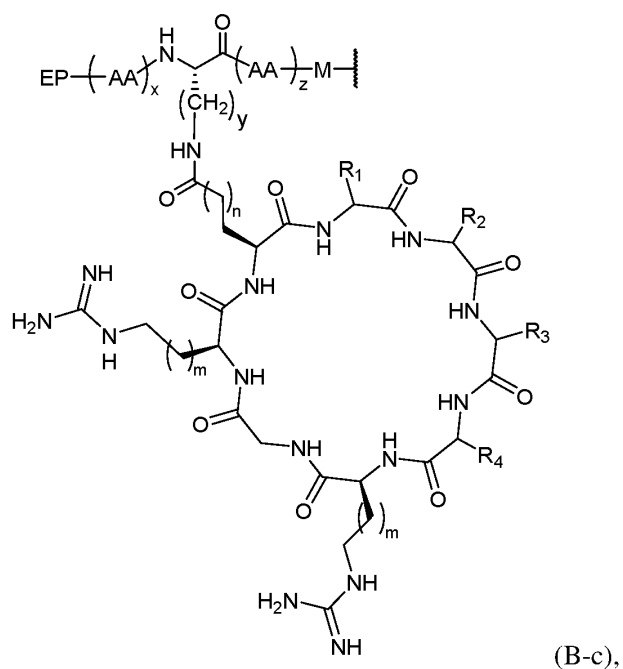




(B-b), or a protonated form thereof,

wherein EP, R¹, R², R³, R⁴, m and z' are as defined above in Formula (B).

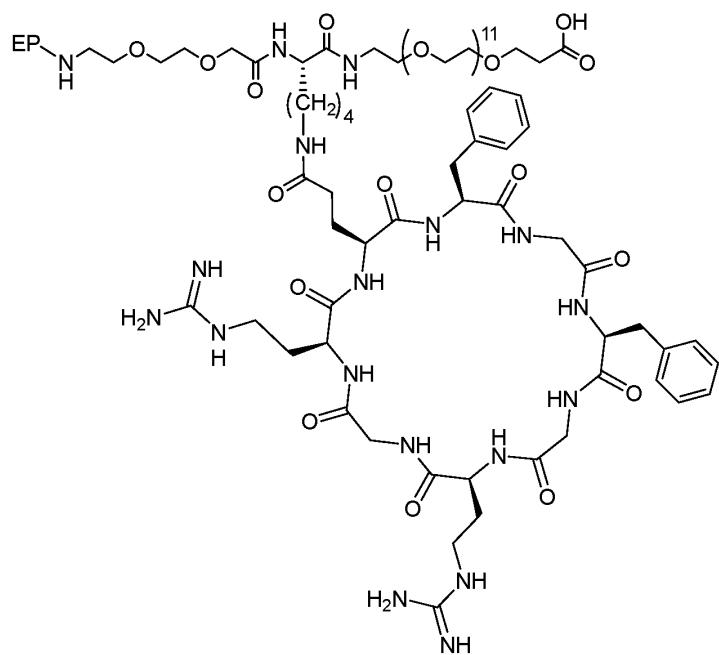
[0225] The EEV can comprises the structure of Formula (B-c):



(B-c),

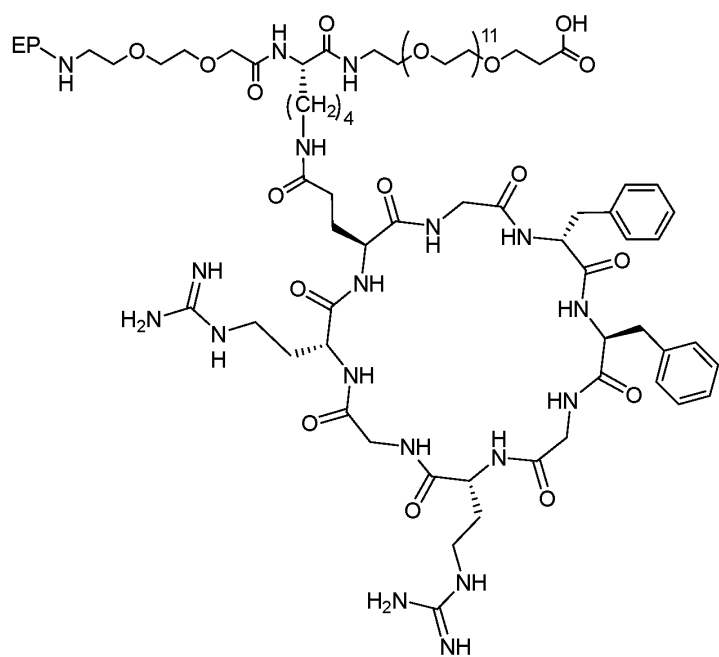
or a protonated form thereof, wherein EP, R¹, R², R³, R⁴, and m are as defined above in Formula (B); AA is an amino acid as defined herein; M is as defined herein; n is an integer from 0-2; x is an integer from 1-10; y is an integer from 1-5; and z is an integer from 1-10.

[0226] The EEV can have the structure of Formula (B-1), (B-2), (B-3), or (B-4):

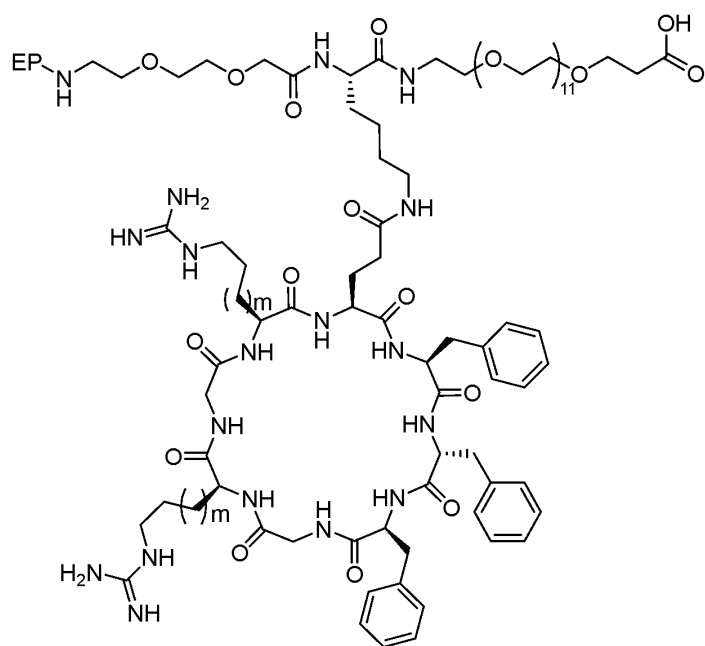


5

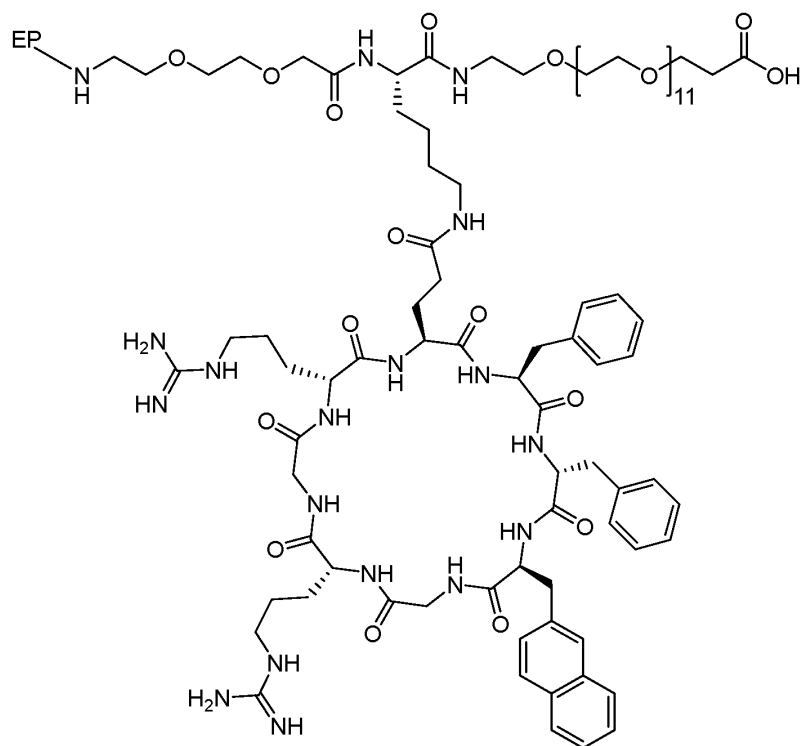
(B-1),



(B-2),



(B-3),



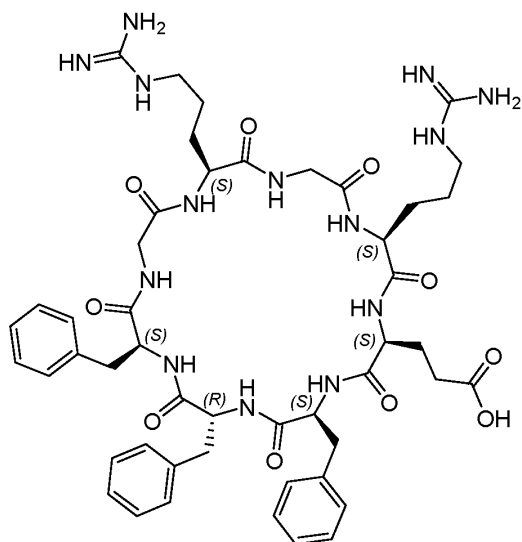
(B-4)

or a protonated form thereof, wherein EP is as defined above in Formula (B).

[0227] The EEV can comprise Formula (B) and can have the structure: Ac-PKKKRKVAEEA-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH or Ac-PK-KKR-KV-AEEA-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH.

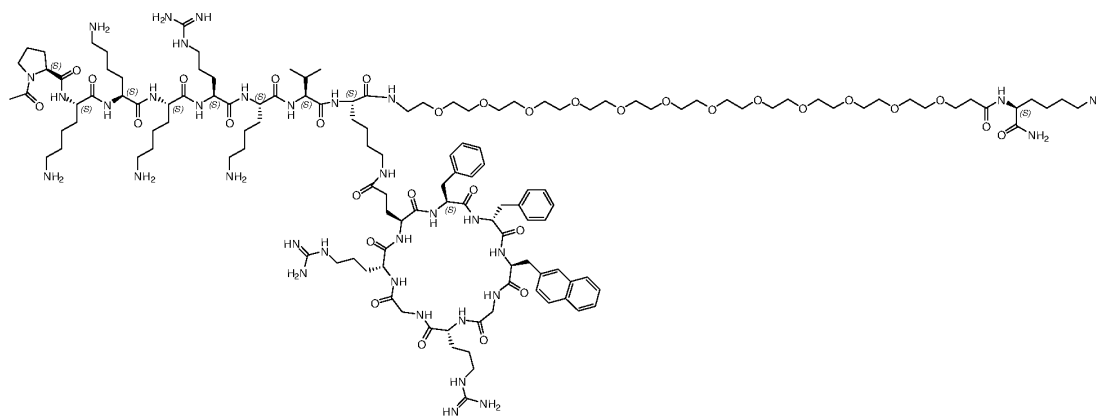
5

[0228] The EEV can comprise a cCPP of formula:



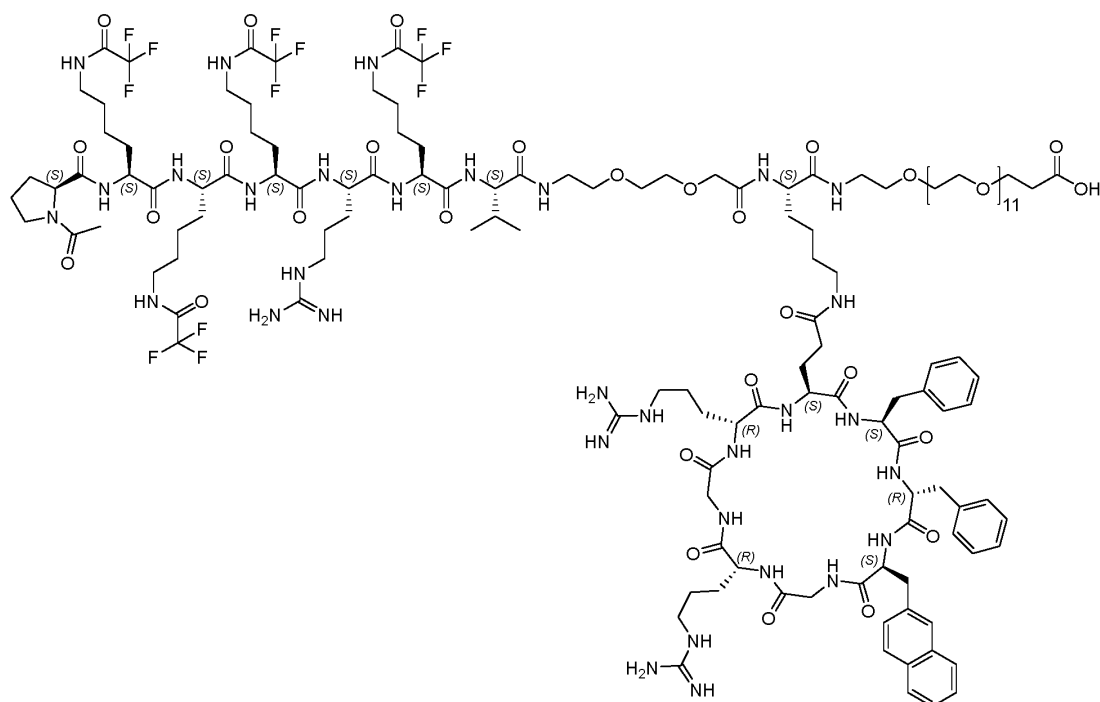
[0229] The EEV can comprise formula: Ac-PKKKRV-miniPEG2-Lys(cyclo(FfFGRGRQ))-miniPEG2-K(N3).

[0230] The EEV can be:



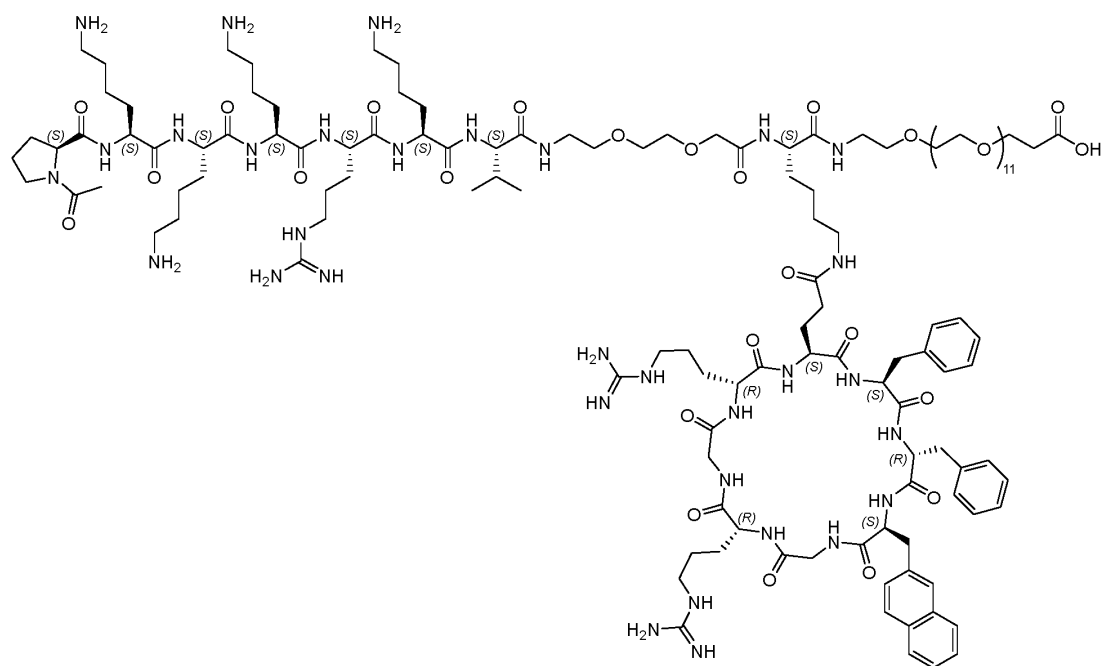
5

[0231] The EEV can be



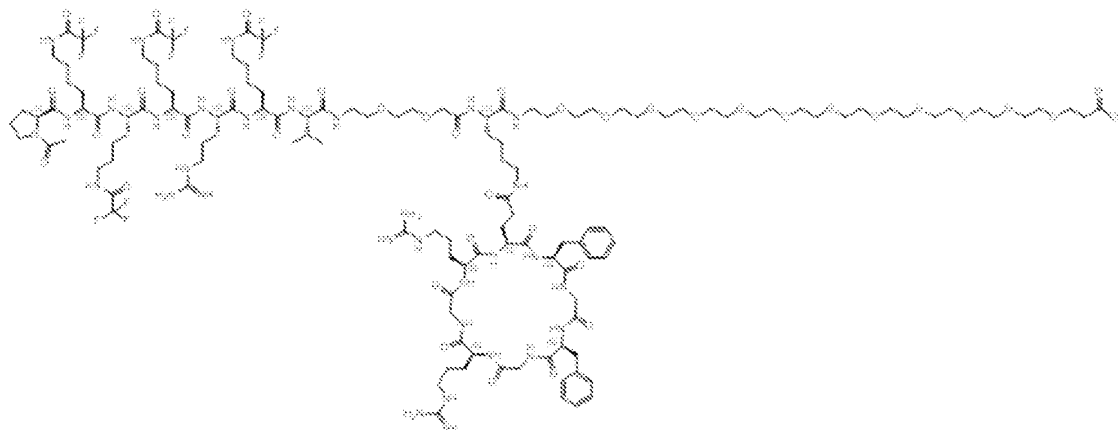
[0232] The EEV can be Ac-P-K(Tfa)-K(Tfa)-K(Tfa)-R-K(Tfa)-V-miniPEG-K(*cyclo*(Ff-Nal-GrGrQ))-PEG12-OH.

5 [0233] The EEV can be



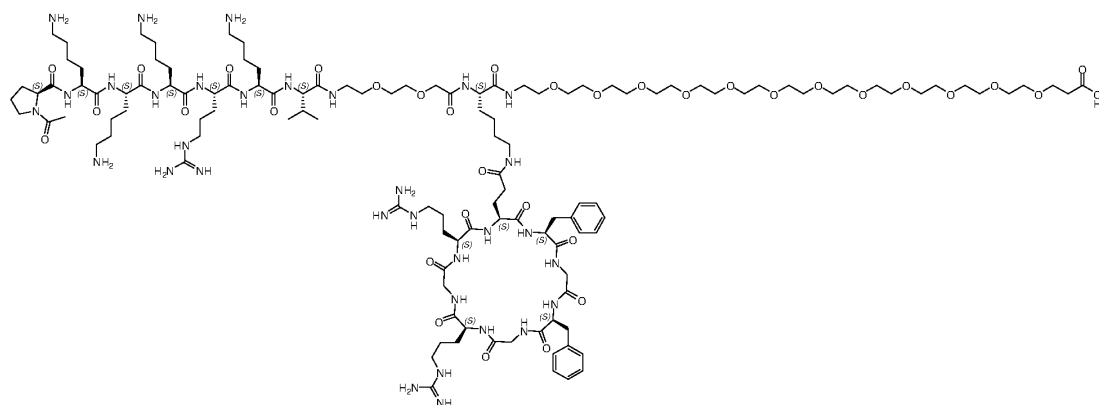
[0234] The EEV can be Ac-P-K-K-K-R-K-V-miniPEG-K(*cyclo*(Ff-Nal-GrGrQ))-PEG12-OH.

[0235] The EEV can be

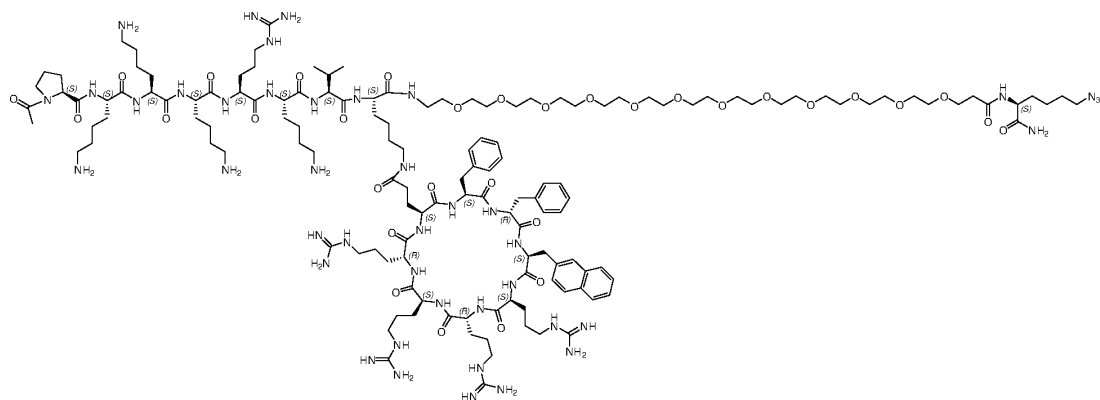


5

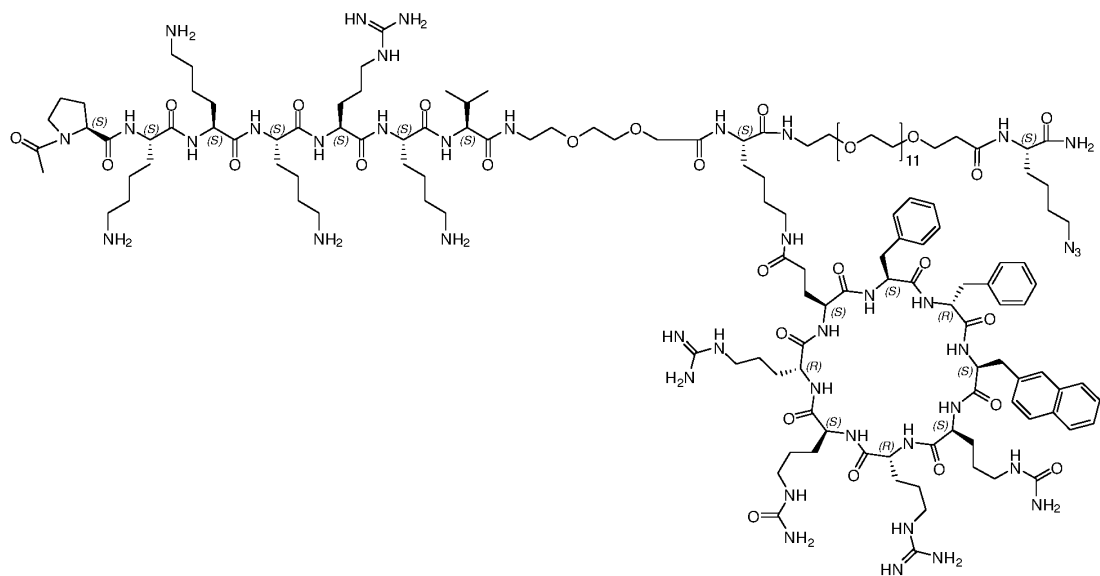
[0236] The EEV can be



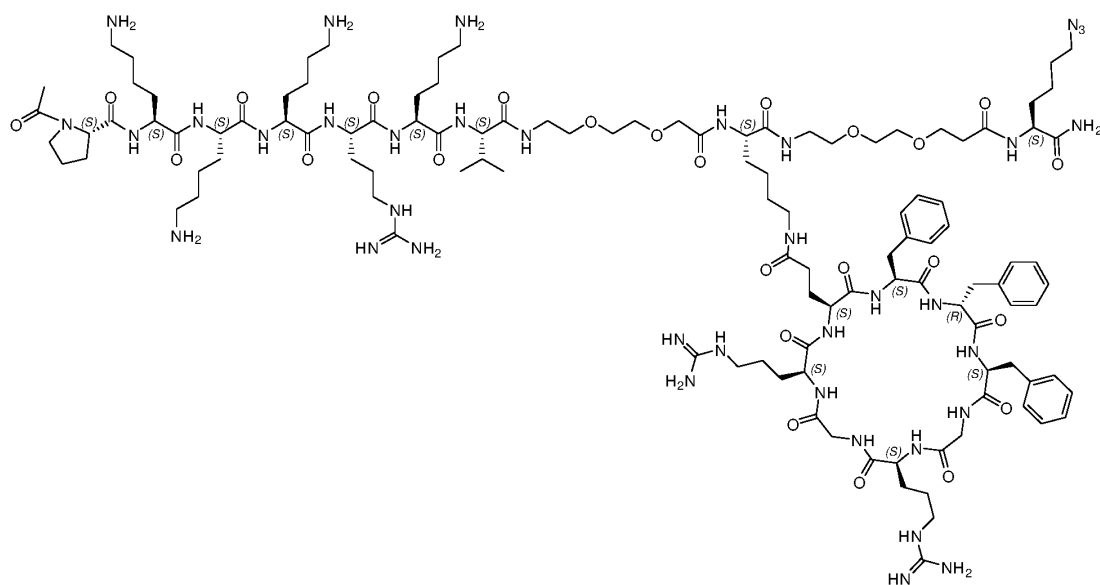
[0237] The EEV can be



[0238] The EEV can be

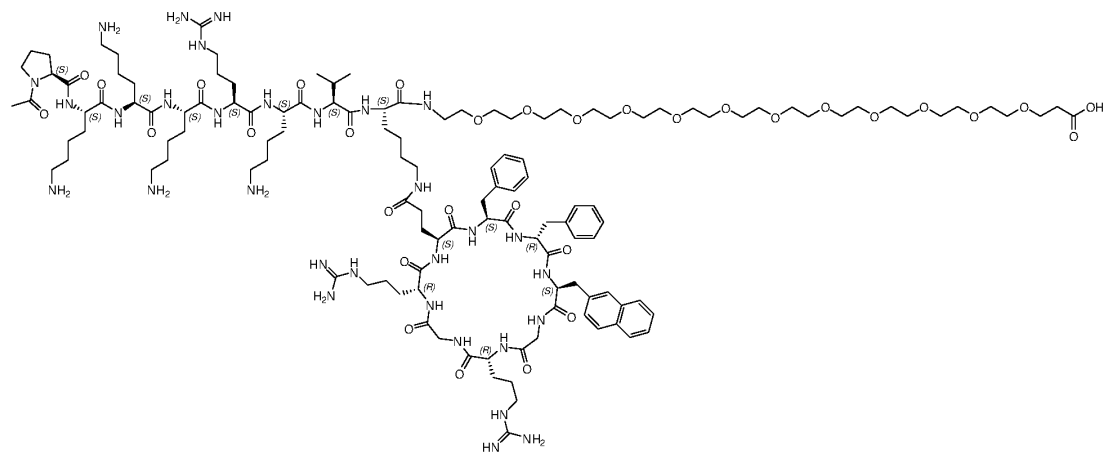


[0239] The EEV can be

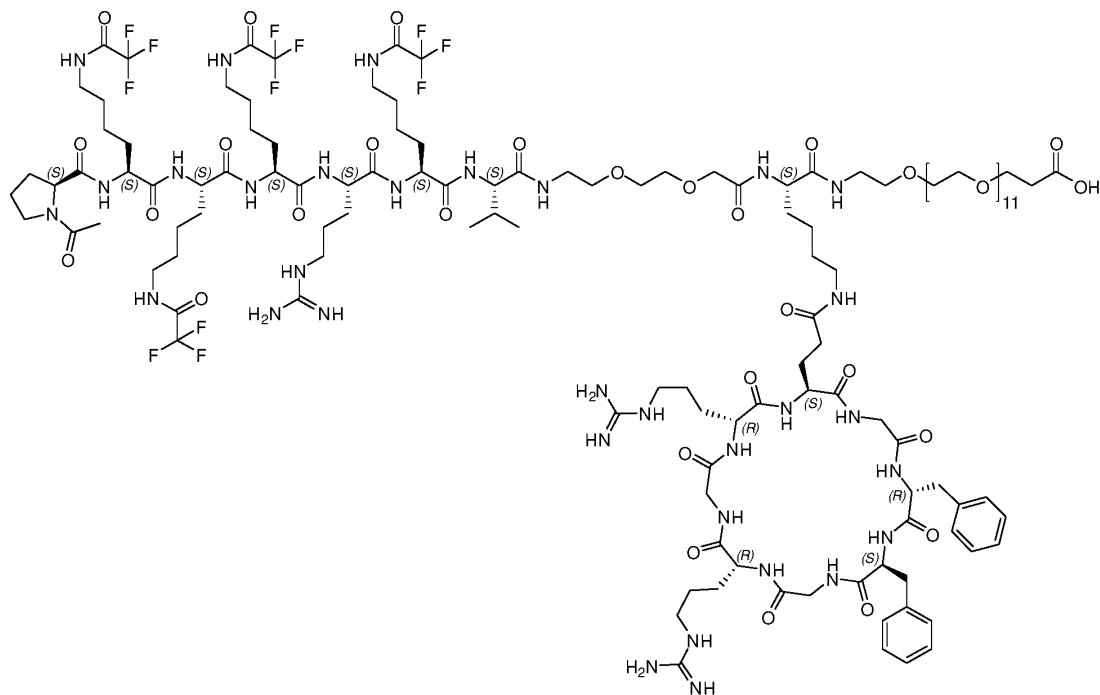


5 .

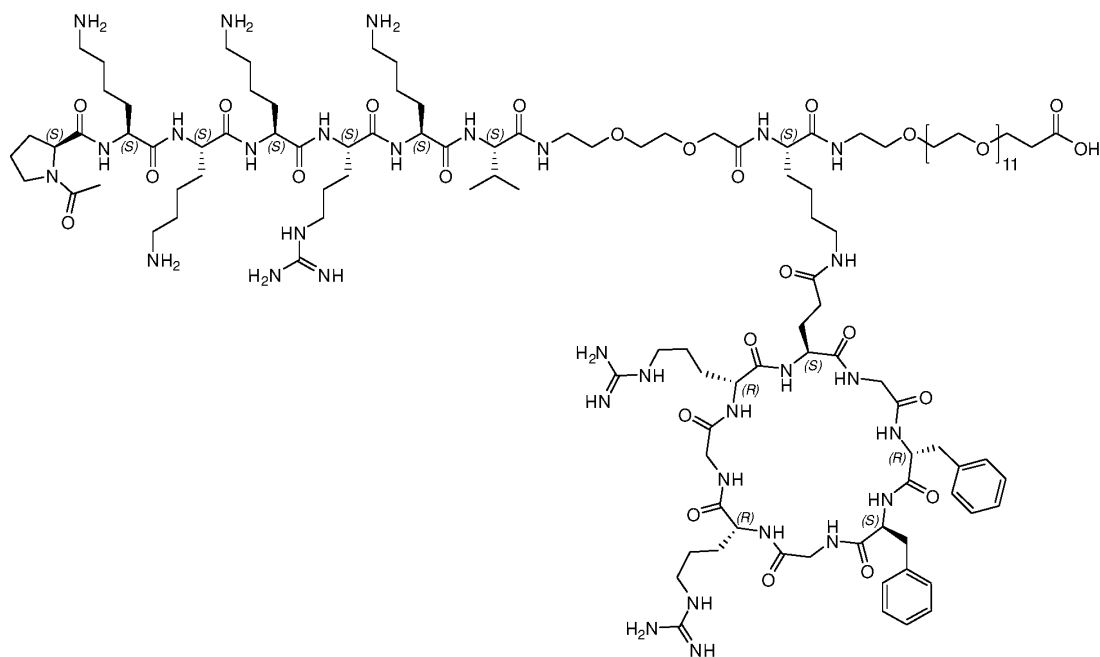
[0240] The EEV can be



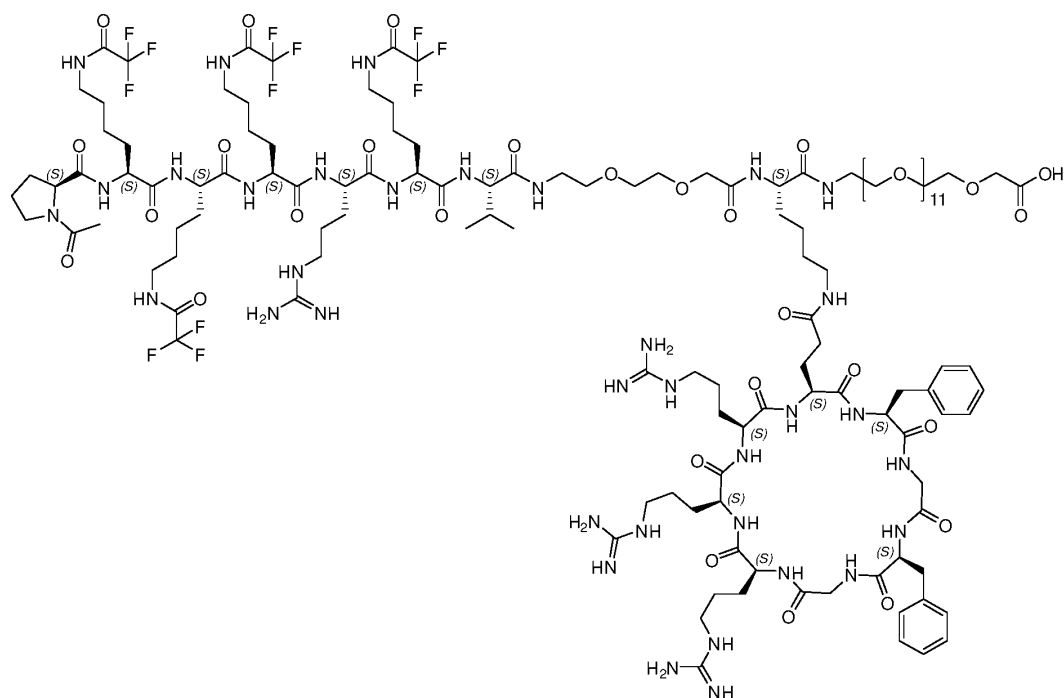
[0241] The EEV can be:



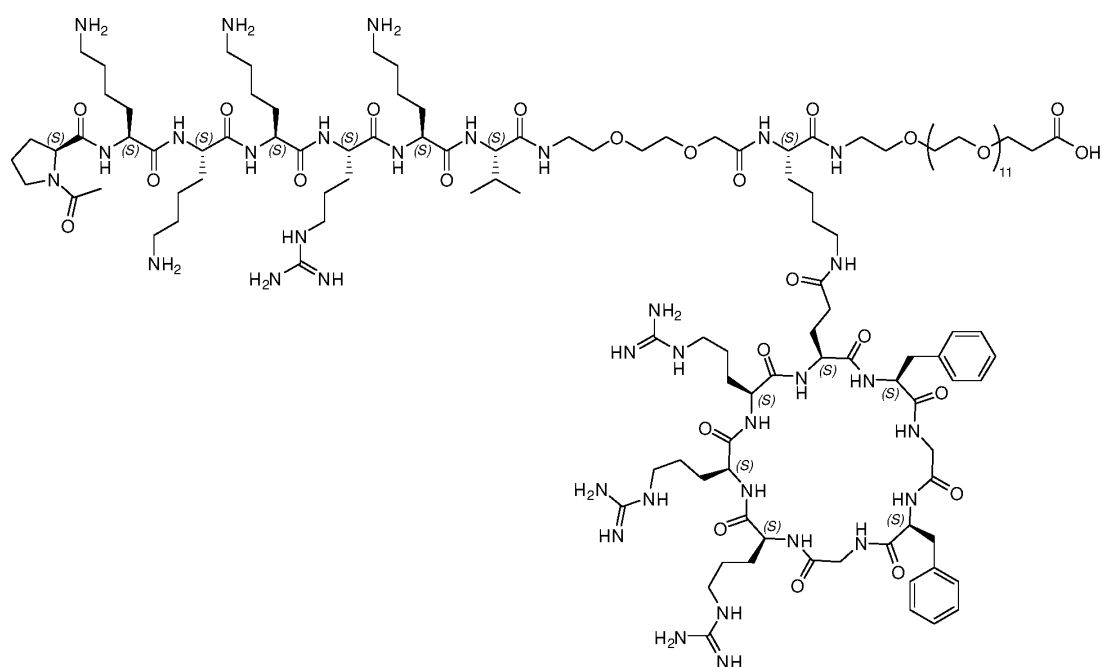
[0242] The EEV can be



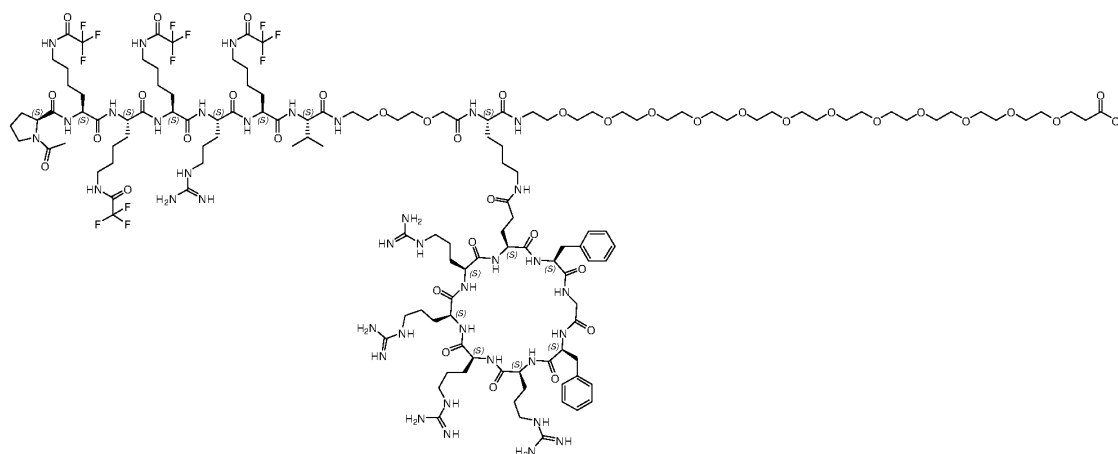
[0243] The EEV can be



[0244] The EEV can be



[0245] The EEV can be



[0246] The EEV can be selected from

Ac-rr-miniPEG2-Dap[cyclo(FfΦ-Cit-r-Cit-rQ)]-PEG12-OH
 Ac-frr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rfr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rbfr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rrr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rbr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rbrbr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-hh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-hbh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-hbhb-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rbhb-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-hbrb-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH
 Ac-rr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-frr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rfr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rbfr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rrr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rbr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rbrbr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH

Ac-hh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-hbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-hbhbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-rbhbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-hbrbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH
 Ac-KKKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KGKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KKKG-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KGK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KBK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KBKBK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KR-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KBR-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PGKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKGKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKGRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKKGKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKKRGV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-PKKKRKG-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KKKRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2
 Ac-KKRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2 and
 Ac-KRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2.

[0247] The EEV can be selected from

Ac-PKKKRKV-Lys(cyclo[Ff-G-r-G-rQ])-PEG12-K(N3)-NH ₂
Ac-PKKKRKV-miniPEG2-Lys(cyclo[Ff-Nal-GrGrQ])-miniPEG2-K(N3)-NH ₂
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFGRGRQ])-miniPEG2-K(N3)-NH ₂
Ac-KR-PEG2-K(cyclo[FGFGRGRQ])-PEG2-K(N3)-NH ₂
Ac-PKKKGKV-PEG2-K(cyclo[FGFGRGRQ])-PEG2-K(N3)-NH ₂
Ac-PKKKRKG-PEG2-K(cyclo[FGFGRGRQ])-PEG2-K(N3)-NH ₂
Ac-KKKRK-PEG2-K(cyclo[FGFGRGRQ])-PEG2-K(N3)-NH ₂
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FFØGRGRQ])-miniPEG2-K(N3)-NH ₂

[0248] The EEV can be selected from

Ac-PKKKRKV-miniPEG2-Lys(cyclo[GfF-GrGrQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFKRKQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFRGRQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFGRGRGRQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFGRrRQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFRRRQ])-PEG12-OH
Ac-PKKKRKV-miniPEG2-Lys(cyclo[FGFRRRRQ])-PEG12-OH

[0249] The EEV can be selected from

Ac-K-K-K-R-K-G-miniPEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-K-K-K-R-K-miniPEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-K-K-R-K-K-PEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-K-R-K-K-K-PEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-K-K-K-K-R-PEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-R-K-K-K-K-PEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH
Ac-K-K-K-R-K-PEG ₂ -K(cyclo[FGFGRGRQ])-PEG ₁₂ -OH

[0250] The EEV can be selected from

Ac-PKKKKRKV-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂
 Ac-PKKKKRKV-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
 Ac-PKKKKRKV-PEG₂-K(*cyclo*[GfFGGrGrQ])-PEG₂-K(N₃)-NH₂
 Ac-PKKKKRKV-PEG₂-K(*cyclo*[GfFGGrGrQ])-PEG₁₂-OH

[0251] The cargo can be a protein and the EEV can be selected from:

- Ac-PKKKKRKV-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
- 5 Ac-PKKKKRKV-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
- Ac-PKKKKRKV-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
- Ac-PKKKKRKV-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
- Ac-PKKKKRKV-PEG₂-K(*cyclo*[GfFGGrGrQ])-PEG₁₂-OH
- Ac-PKKKKRKV-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
- 10 Ac-PKKKKRKV-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
- 15 Ac-rr-PEG₂-K(*cyclo*[GfFGGrGrQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
- Ac-rr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- Ac-rrr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
- 20 Ac-rrr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
- Ac-rrr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
- Ac-rrr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
- Ac-rrr-PEG₂-K(*cyclo*[GfFGGrGrQ])-PEG₁₂-OH
- Ac-rrr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
- 25 Ac-rrr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- Ac-rhr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
- Ac-rhr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH

- Ac-rhr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
Ac-rhr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
Ac-rhr-PEG₂-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH
Ac-rhr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
5 Ac-rhr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- Ac-rbr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
Ac-rbr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
Ac-rbr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
10 Ac-rbr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
Ac-rbr-PEG₂-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH
Ac-rbr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
Ac-rbr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- 15 Ac-rbrbr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
Ac-rbrbr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
Ac-rbrbr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
Ac-rbrbr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
Ac-rbrbr-PEG₂-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH
20 Ac-rbrbr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
Ac-rbrbr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
- Ac-rbhbr-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH
Ac-rbhbr-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH
25 Ac-rbhbr-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH
Ac-rbhbr-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH
Ac-rbhbr-PEG₂-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH
Ac-rbhbr-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH
Ac-rbhbr-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH
30 Ac-hbrbh-PEG₂-K(*cyclo*[Ff-Nal-GrGrQ])-PEG₁₂-OH

Ac-hbrbh-PEG₂-K(*cyclo*[Ff-Nal-Cit-r-Cit-rQ])-PEG₁₂-OH

Ac-hbrbh-PEG₂-K(*cyclo*[FfF-GRGRQ])-PEG₁₂-OH

Ac-hbrbh-PEG₂-K(*cyclo*[FGFGRGRQ])-PEG₁₂-OH

Ac-hbrbh-PEG₂-K(*cyclo*[GfFGrGrQ])-PEG₁₂-OH

5 Ac-hbrbh-PEG₂-K(*cyclo*[FGFGRRRQ])-PEG₁₂-OH and

Ac-hbrbh-PEG₂-K(*cyclo*[FGFRRRRQ])-PEG₁₂-OH

wherein b is beta-alanine, and the exocyclic sequence can be D or L stereochemistry.

[0252] In embodiments, provided herein are TMs that are conjugated to two CPPs. Non-limiting
 10 examples of the structures of TMs that are conjugated to two CPPs are provided below. For
 illustrative purposes only, the TM in the structures shown is an AC. Other TM, for example,
 therapeutic polypeptides, can also be used. Underlining represents the illustrative antisense
 oligonucleotide. The antisense oligonucleotide sequences shown below are for illustrative
 purposes only, and can be substituted for another antisense oligonucleotide sequence depending
 15 on the target of interest. In embodiments, provided herein are TMs that are conjugated to three
 CPPs. Non-limiting examples of the structures of TMs that are conjugated to three CPPs are
 provided below. For illustrative purposes only, the TM in the structures shown is an AC. Other
 TM, for example, therapeutic polypeptides, can also be used. Underlining represents the antisense
 oligonucleotide. The antisense oligonucleotide sequences shown below are for illustrative
 20 purposes only, and can be substituted for another antisense oligonucleotide sequence depending
 on the target of interest.

Cargo

[0253] The cell penetrating peptide (CPP), such as a cyclic cell penetrating peptide (e.g., cCPP),
 25 can be conjugated to a cargo. The cargo can be a therapeutic moiety (TM). The cargo can be
 conjugated to a terminal carbonyl group of a linker. At least one atom of the cyclic peptide can be
 replaced by a cargo or at least one lone pair can form a bond to a cargo. The cargo can be
 conjugated to the cCPP by a linker. The cargo can be conjugated to an AA_{SC} by a linker. At least
 one atom of the cCPP can be replaced by a therapeutic moiety or at least one lone pair of the cCPP
 30 forms a bond to a therapeutic moiety. A hydroxyl group on an amino acid side chain of the cCPP
 can be replaced by a bond to the cargo. A hydroxyl group on a glutamine side chain of the cCPP

can be replaced by a bond to the cargo. The cargo can be conjugated to the cCPP by a linker. The cargo can be conjugated to an AA_{Sc} by a linker.

[0254] The cargo can comprise one or more detectable moieties, one or more therapeutic moieties, one or more targeting moieties, or any combination thereof. The cargo can be a peptide, oligonucleotide, or small molecule. The cargo can be a peptide sequence or a non-peptidyl therapeutic agent. The cargo can be an antibody or an antigen binding fragment thereof, including, but not limited to an scFv or nanobody.

[0255] The cargo can comprise one or more additional amino acids (e.g., K, UK, TRV); a linker (e.g., bifunctional linker LC-SMCC); coenzyme A; phosphocoumaryl amino propionic acid (pCAP); 8-amino-3,6-dioxaoctanoic acid (miniPEG); L-2,3-diaminopropionic acid (Dap or J); L-β-naphthylalanine; L-pipecolic acid (Pip); sarcosine; trimesic acid; 7-amino-4-methylcoumarin (Amc); fluorescein isothiocyanate (FITC); L-2-naphthylalanine; norleucine; 2-aminobutyric acid; Rhodamine B (Rho); Dexamethasone (DEX); or combinations thereof.

[0256] The cargo can comprise any of those listed in Table 6, or derivatives or combinations thereof.

Table 6. Example cargo moieties

Abbreviation	Sequence*
R ₅	RRRRR
A ₅	AAAAA
F ₄	FFFF
PCP	DE(pCAP)LI
A ₇	AAAAAAA
	RARAR
	DADAD
	DΩUD
	UTRV
	D-pThr-Pip-Nal

*pCAP, phosphocoumaryl amino propionic acid; Ω, norleucine; U, 2-aminobutyric acid; D-pThr is D-phosphothreonine, Pip is L-piperidine-2-carboxylate.

20 Detectable moiety

[0257] The compound can include a detectable moiety. The detectable moiety can be attached to a cell penetrating peptide (CPP) at the amino group, the carboxylate group, or the side chain of any of the amino acids of the CPP (e.g., at the amino group, the carboxylate group, or the side chain of any amino acid in the cCPP). The detectable moiety can be attached to a cyclic cell penetrating

peptide (cCPP) at the side chain of any amino acid in the cCPP. The cargo can include a detectable moiety. The cargo can include a therapeutic agent and a detectable moiety. The detectable moiety can include any detectable label. Examples of suitable detectable labels include, but are not limited to, a UV-Vis label, a near-infrared label, a luminescent group, a phosphorescent group, a magnetic spin resonance label, a photosensitizer, a photocleavable moiety, a chelating center, a heavy atom, a radioactive isotope, an isotope detectable spin resonance label, a paramagnetic moiety, a chromophore, or any combination thereof. The label can be detectable without the addition of further reagents.

[0258] The detectable moiety can be a biocompatible detectable moiety, such that the compounds can be suitable for use in a variety of biological applications. “Biocompatible” and “biologically compatible”, as used herein, generally refer to compounds that are, along with any metabolites or degradation products thereof, generally non-toxic to cells and tissues, and which do not cause any significant adverse effects to cells and tissues when cells and tissues are incubated (*e.g.*, cultured) in their presence.

[0259] The detectable moiety can contain a luminophore such as a fluorescent label or near-infrared label. Examples of suitable luminophores include, but are not limited to, metal porphyrins; benzoporphyrins; azabenzoporphyrine; naphthoporphyrin; phthalocyanine; polycyclic aromatic hydrocarbons such as diimine, pyrenes; azo dyes; xanthene dyes; boron dipyrromethene, aza-boron dipyrromethene, cyanine dyes, metal-ligand complex such as bipyridine, bipyridyls, phenanthroline, coumarin, and acetylacetonates of ruthenium and iridium; acridine, oxazine derivatives such as benzophenoxazine; aza-annulene, squaraine; 8-hydroxyquinoline, polymethines, luminescent producing nanoparticle, such as quantum dots, nanocrystals; carbostyryl; terbium complex; inorganic phosphor; ionophore such as crown ethers affiliated or derivatized dyes; or combinations thereof. Specific examples of suitable luminophores include, but are not limited to, Pd (II) octaethylporphyrin; Pt (II)-octaethylporphyrin; Pd (II) tetraphenylporphyrin; Pt (II) tetraphenylporphyrin; Pd (II) meso-tetraphenylporphyrin tetrabenzoporphine; Pt (II) meso-tetraphenyl methylbenzoporphyrin; Pd (II) octaethylporphyrin ketone; Pt (II) octaethylporphyrin ketone; Pd (II) meso-tetra(pentafluorophenyl)porphyrin; Pt (II) meso-tetra (pentafluorophenyl) porphyrin; Ru (II) tris(4,7-diphenyl-1,10-phenanthroline) (Ru (dpp)₃); Ru (II) tris(1,10-phenanthroline) (Ru(phen)₃), tris(2,2'-bipyridine)ruthenium (II) chloride hexahydrate (Ru(bpy)₃); erythrosine B; fluorescein; fluorescein isothiocyanate (FITC); eosin;

iridium (III) ((N-methyl-benzimidazol-2-yl)-7-(diethylamino)-coumarin)); 8-enzothiazole ((benzothiazol-2-yl)-7-(diethylamino)-coumarin))-2-(acetylacetonate); Lumogen dyes; Macroflex fluorescent red; Macrolex fluorescent yellow; Texas Red; rhodamine B; rhodamine 6G; sulfur rhodamine; m-cresol; thymol blue; xylenol blue; cresol red; chlorophenol blue; bromocresol green; bromocresol red; bromothymol blue; Cy2; a Cy3; a Cy5; a Cy5.5; Cy7; 4-nitrophenol; alizarin; phenolphthalein; o-cresolphthalein; chlorophenol red; calmagite; bromo-xylenol; phenol red; neutral red; nitrazine; 3,4,5,6-tetrabromophenolphthalein; congo red; fluor'sc'in; eosin; 2',7'-dichlorofluorescein; 5(6)-carboxy-fluorescein; carboxynaphthofluorescein; 8-hydroxypyrene-1,3,6-trisulfonic acid; semi-naphthorhodafluor; semi-naphthofluorescein; tris (4,7-diphenyl-1,10-phenanthroline) ruthenium (II) dichloride; (4,7-diphenyl-1,10-phenanthroline) ruthenium (II) tetraphenylboron; platinum (II) octaethylporphyrin; dialkylcarbocyanine; dioctadecylcycloxacarbocyanine; fluorenylmethyloxycarbonyl chloride; 7-amino-4-methylcoumarin (Amc); green fluorescent protein (GFP); and derivatives or combinations thereof.

[0260] The detectable moiety can include Rhodamine B (Rho), fluorescein isothiocyanate (FITC), 7-amino-4-methylcoumarin (Amc), green fluorescent protein (GFP), or derivatives or combinations thereof.

[0261] The detectable moiety can be attached to a cell penetrating peptide (CPP) at the amino group, the carboxylate group, or the side chain of any of the amino acids of the cell penetrating peptide (e.g., at the amino group, the carboxylate group, or the side chain of any amino acid in the cCPP).

Therapeutic moiety

[0262] The disclosed compounds can comprise a therapeutic moiety. The cargo can comprise a therapeutic moiety. The detectable moiety can be linked to a therapeutic moiety or a detectable moiety can also serve as the therapeutic moiety. Therapeutic moiety refers to a group that when administered to a subject will reduce one or more symptoms of a disease or disorder. The therapeutic moiety can comprise a peptide, protein (e.g., enzyme, antibody or fragment thereof), small molecule, or oligonucleotide.

[0263] The therapeutic moiety can comprise a wide variety of drugs, including antagonists, for example enzyme inhibitors, and agonists, for example a transcription factor which results in an increase in the expression of a desirable gene product (although as will be appreciated by those in

the art, antagonistic transcription factors can also be used), are all included. In addition, therapeutic moiety includes those agents capable of direct toxicity and/or capable of inducing toxicity towards healthy and/or unhealthy cells in the body. Also, the therapeutic moiety can be capable of inducing and/or priming the immune system against potential pathogens.

- 5 **[0264]** The therapeutic moiety can, for example, comprise an anticancer agent, antiviral agent, antimicrobial agent, anti-inflammatory agent, immunosuppressive agent, anesthetics, or any combination thereof.

- [0265]** The therapeutic moiety can comprise an anticancer agent. Example anticancer agents include 13-cis-Retinoic Acid, 2-Amino-6-Mercaptopurine, 2-CdA, 2-Chlorodeoxyadenosine, 5-
 10 fluorouracil, 6-Thioguanine, 6-Mercaptopurine, Accutane, Actinomycin-D, Adriamycin, Adrucil, Agrylin, Ala-Cort, Aldesleukin, Alemtuzumab, Alitretinoin, Alkaban-AQ, Alkeran, All-transretinoic acid, Alpha interferon, Altretamine, Amethopterin, Amifostine, Aminoglutethimide, Anagrelide, Anandron, Anastrozole, Arabinosylcytosine, Aranesp, Aredia, Arimidex, Aromasin, Arsenic trioxide, Asparaginase, ATRA, Avastin, BCG, BCNU, Bevacizumab, Bexarotene,
 15 Bicalutamide, BiCNU, Blenoxane, Bleomycin, Bortezomib, Busulfan, Busulfex, C225, Calcium Leucovorin, Campath, Camptosar, Camptothecin-11, Capecitabine, Carac, Carboplatin, Carmustine, Carmustine wafer, Casodex, CCNU, CDDP, CeeNU, Cerubidine, cetuximab, Chlorambucil, Cisplatin, Citrovorum Factor, Cladribine, Cortisone, Cosmegen, CPT-11, Cyclophosphamide, Cytadren, Cytarabine, Cytarabine liposomal, Cytosar-U, Cytosan,
 20 Dacarbazine, Dactinomycin, Darbepoetin alfa, Daunomycin, Daunorubicin, Daunorubicin hydrochloride, Daunorubicin liposomal, DaunoXome, Decadron, Delta-Cortef, Deltasone, Denileukin diftitox, DepoCyt, Dexamethasone, Dexamethasone acetate, Dexamethasone sodium phosphate, Dexasone, Dexrazoxane, DHAD, DIC, Diodex, Docetaxel, Doxil, Doxorubicin, Doxorubicin liposomal, Droxia, DTIC, DTIC-Dome, Duralone, Efudex, Eligard, Ellence,
 25 Eloxatin, Elspar, Emcyt, Epirubicin, Epoetin alfa, Erbitux, Erwinia L-asparaginase, Estramustine, Ethiol, Etopophos, Etoposide, Etoposide phosphate, Eulexin, Evista, Exemestane, Fareston, Faslodex, Femara, Filgrastim, Floxuridine, Fludara, Fludarabine, Fluoroplex, Fluorouracil, Fluorouracil (cream), Fluoxymesterone, Flutamide, Folinic Acid, FUDR, Fulvestrant, G-CSF, Gefitinib, Gemcitabine, Gemtuzumab ozogamicin, Gemzar, Gleevec, Lupron, Lupron Depot,
 30 Matulane, Maxidex, Mechlorethamine, -Mechlorethamine Hydrochlorine, Medralone, Medrol, Megace, Megestrol, Megestrol Acetate, Melphalan, Mercaptopurine, Mesna, Mesnex,

Methotrexate, Methotrexate Sodium, Methylprednisolone, Mylocel, Letrozole, Neosar, Neulasta, Neumega, Neupogen, Nilandron, Nilutamide, Nitrogen Mustard, Novaldex, Novantrone, Octreotide, Octreotide acetate, Oncospar, Oncovin, Ontak, Onxal, Oprevelkin, Orapred, Orasone, Oxaliplatin, Paclitaxel, Pamidronate, Panretin, Paraplatin, Pediapred, PEG Interferon, 5 Pegaspargase, Pegfilgrastim, PEG-INTRON, PEG-L-asparaginase, Phenylalanine Mustard, Platinol, Platinol-AQ, Prednisolone, Prednisone, Prelone, Procarbazine, PROCIT, Proleukin, Prolifeprospan 20 with Carmustine implant, Purinethol, Raloxifene, Rheumatrex, Rituxan, Rituximab, Roveron-A (interferon alfa-2a), Rubex, Rubidomycin hydrochloride, Sandostatin, Sandostatin LAR, Sargramostim, Solu-Cortef, Solu-Medrol, STI-571, Streptozocin, Tamoxifen, 10 Targretin, Taxol, Taxotere, Temodar, Temozolomide, Teniposide, TESP, Thalidomide, Thalomid, TheraCys, Thioguanine, Thioguanine Tabloid, Thiophosphoamide, Thioplex, Thiotepa, TICE, Toposar, Topotecan, Toremifene, Trastuzumab, Tretinoin, Trexall, Trisenox, TSPA, VCR, Velban, Velcade, VePesid, Vesanoide, Viadur, Vinblastine, Vinblastine Sulfate, Vincasar Pfs, Vincristine, Vinorelbine, Vinorelbine tartrate, VLB, VP-16, Vumon, Xeloda, Zanosar, Zevalin, 15 Zinecard, Zoladex, Zoledronic acid, Zometa, Gliadel wafer, Glivec, GM-CSF, Goserelin, granulocyte colony stimulating factor, Halotestin, Herceptin, Hexadrol, Hexalen, Hexamethylmelamine, HMM, Hycamtin, Hydrea, Hydrocort Acetate, Hydrocortisone, Hydrocortisone sodium phosphate, Hydrocortisone sodium succinate, Hydrocortone phosphate, Hydroxyurea, Ibritumomab, Ibritumomab Tiuxetan, Idamycin, Idarubicin, Ifex, IFN-alpha, 20 Ifosfamide, IL 2, IL-11, Imatinib mesylate, Imidazole Carboxamide, Interferon alfa, Interferon Alfa-2b (PEG conjugate), Interleukin 2, Interleukin-11, Intron A (interferon alfa-2b), Leucovorin, Leukeran, Leukine, Leuprolide, Leurocristine, Leustatin, Liposomal Ara-C, Liquid Pred, Lomustine, L-PAM, L-Sarcosine, Meticorten, Mitomycin, Mitomycin-C, Mitoxantrone, M-Prednisol, MTC, MTX, Mustargen, Mustine, Mutamycin, Myleran, Iressa, Irinotecan, Isotretinoin, 25 Kidrolase, Lanacort, L-asparaginase, and LCR. The therapeutic moiety can also comprise a biopharmaceutical such as, for example, an antibody.

[0266] The therapeutic moiety can comprise an antiviral agent, such as ganciclovir, azidothymidine (AZT), lamivudine (3TC), *etc.*

[0267] The therapeutic moiety can comprise an antibacterial agent, such as acedapsone; 30 acetosulfone sodium; alamecin; alexidine; amdinocillin; amdinocillin pivoxil; amicycline; amifloxacin; amifloxacin mesylate; amikacin; amikacin sulfate; aminosalicic acid;

aminosalicylate sodium; amoxicillin; amphomycin; ampicillin; ampicillin sodium; apalcillin sodium; apramycin; aspartocin; astromicin sulfate; avilamycin; avoparcin; azithromycin; azlocillin; azlocillin sodium; bacampicillin hydrochloride; bacitracin; bacitracin methylene disalicylate; bacitracin zinc; bambermycins; benzoylpas calcium; berythromycin; betamicin sulfate; biapenem; biniramycin; biphenamine hydrochloride; bispyrithione magsulfex; butikacin; butirosin sulfate; capreomycin sulfate; carbadox; carbenicillin disodium; carbenicillin indanyl sodium; carbenicillin phenyl sodium; carbenicillin potassium; carumonam sodium; cefaclor; cefadroxil; cefamandole; cefamandole nafate; cefamandole sodium; cefaparole; cefatrizine; cefazaflur sodium; cefazolin; cefazolin sodium; cefbuperazone; cefdinir; cefepime; cefepime hydrochloride; cefetecol; cefixime; cefmenoxime hydrochloride; cefmetazole; cefmetazole sodium; cefonicid monosodium; cefonicid sodium; cefoperazone sodium; ceforanide; cefotaxime sodium; cefotetan; cefotetan disodium; cefotiam hydrochloride; cefoxitin; cefoxitin sodium; cefpimizole; cefpimizole sodium; cefpiramide; cefpiramide sodium; cefpirome sulfate; cefpodoxime proxetil; cefprozil; cefroxadine; cefsulodin sodium; ceftazidime; ceftibuten; ceftizoxime sodium; ceftriaxone sodium; cefuroxime; cefuroxime axetil; cefuroxime pivoxetil; cefuroxime sodium; cephaetrile sodium; cephalixin; cephalixin hydrochloride; cephaloglycin; cephaloridine; cephalothin sodium; cephapirin sodium; cephradine; cetocycline hydrochloride; cetophenicol; chloramphenicol; chloramphenicol palmitate; chloramphenicol pantothenate complex; chloramphenicol sodium succinate; chlorhexidine phosphanilate; chloroxylenol; chlortetracycline bisulfate; chlortetracycline hydrochloride; cinoxacin; ciprofloxacin; ciprofloxacin hydrochloride; cirolemycin; clarithromycin; clinafloxacin hydrochloride; clindamycin; clindamycin hydrochloride; clindamycin palmitate hydrochloride; clindamycin phosphate; clofazimine; cloxacillin benzathine; cloxacillin sodium; cloxyquin; colistimethate sodium; colistin sulfate; coumermycin; coumermycin sodium; cyclacillin; cycloserine; dalfopristin; dapsone; daptomycin; demeclocycline; demeclocycline hydrochloride; demecycline; denofungin; diaveridine; dicloxacillin; dicloxacillin sodium; dihydrostreptomycin sulfate; dipyrithione; dirithromycin; doxycycline; doxycycline calcium; doxycycline fosfatex; doxycycline hyclate; droxacin sodium; enoxacin; epicillin; epitetracycline hydrochloride; erythromycin; erythromycin acistrate; erythromycin estolate; erythromycin ethylsuccinate; erythromycin gluceptate; erythromycin lactobionate; erythromycin propionate; erythromycin stearate; ethambutol hydrochloride; ethionamide; fleroxacin; floxacillin; fludalanine; flumequine;

fosfomycin; fosfomycin tromethamine; fumoxicillin; furazolidium chloride; furazolidium tartrate; fusidate sodium; fusidic acid; gentamicin sulfate; gloximonam; gramicidin; haloprogin; hetacillin; hetacillin potassium; hexedine; ibafloxacin; imipenem; isoconazole; isepamicin; isoniazid; josamycin; kanamycin sulfate; kitasamycin; levofuraltadone; levopropylcillin potassium; 5 lexithromycin; lincomycin; lincomycin hydrochloride; lomefloxacin; Lomefloxacin hydrochloride; lomefloxacin mesylate; loracarbef; mafenide; meclocycline; meclocycline sulfosalicylate; megalomicin potassium phosphate; mequidox; meropenem; methacycline; methacycline hydrochloride; methenamine; methenamine hippurate; methenamine mandelate; methicillin sodium; metioprime; metronidazole hydrochloride; metronidazole phosphate; 10 mezlocillin; mezlocillin sodium; minocycline; minocycline hydrochloride; mirincamycin hydrochloride; monensin; monensin sodium; nafcillin sodium; nalidixate sodium; nalidixic acid; natacycin; nebramycin; neomycin palmitate; neomycin sulfate; neomycin undecylenate; netilmicin sulfate; neutramycin; nifuradene; nifuraldehyde; nifuratel; nifuratrone; nifurazil; nifurimide; nifiupirinol; nifurquinazol; nifurthiazole; nitrocyline; nitrofurantoin; nitromide; 15 norfloxacin; novobiocin sodium; ofloxacin; onnetoprim; oxacillin; oxacillin sodium; oximonam; oximonam sodium; oxolinic acid; oxytetracycline; oxytetracycline calcium; oxytetracycline hydrochloride; paldimycin; parachlorophenol; paulomycin; pefloxacin; pefloxacin mesylate; penamcillin; penicillin G benzathine; penicillin G potassium; penicillin G procaine; penicillin G sodium; penicillin V; penicillin V benzathine; penicillin V hydrabamine; penicillin V potassium; 20 pentizidone sodium; phenyl aminosalicylate; piperacillin sodium; pirbenicillin sodium; piridicillin sodium; pirlimycin hydrochloride; pivampicillin hydrochloride; pivampicillin pamoate; pivampicillin probenate; polymyxin B sulfate; porfiromycin; propikacin; pyrazinamide; pyrrhione zinc; quindecamine acetate; quinupristin; racephenicol; ramoplanin; ranimycin; relomycin; repromicin; rifabutin; rifamethane; rifamexil; rifamide; rifampin; rifapentine; rifaximin; 25 rolitetracycline; rolitetracycline nitrate; rosaramicin; rosaramicin butyrate; rosaramicin propionate; rosaramicin sodium phosphate; rosaramicin stearate; rosoxacin; roxarsone; roxithromycin; sancycline; sanfetrinem sodium; sarmoxicillin; sarpicillin; scopafungin; sisomicin; sisomicin sulfate; sparfloxacin; spectinomycin hydrochloride; spiramycin; stallimycin hydrochloride; steffimycin; streptomycin sulfate; streptonicozid; sulfabenz; sulfabenzamide; 30 sulfacetamide; sulfacetamide sodium; sulfacycline; sulfadiazine; sulfadiazine sodium; sulfadoxine; sulfalene; sulfamerazine; sulfameter; sulfamethazine; sulfamethazole; sulfamethoxazole;

sulfamonomethoxine; sulfamoxole; sulfanilate zinc; sulfanitran; sulfasalazine; sulfasomizole; sulfathiazole; sulfazamet; sulfisoxazole; sulfisoxazole acetyl; sulfisboxazole diolamine; sulfomyxin; sulopenem; sultamricillin; suncillin sodium; talampicillin hydrochloride; teicoplanin; temafloxacin hydrochloride; temocillin; tetracycline; tetracycline hydrochloride; tetracycline phosphate complex; tetroxoprim; thiamphenicol; thiphencillin potassium; ticarcillin cresyl sodium; ticarcillin disodium; ticarcillin monosodium; ticlatone; tiodonium chloride; tobramycin; tobramycin sulfate; tosufloxacin; trimethoprim; trimethoprim sulfate; trisulfapyrimidines; troleandomycin; trospectomycin sulfate; tyrothricin; vancomycin; vancomycin hydrochloride; virginiamycin; or zorbamycin.

10 [0268] The therapeutic moiety can comprise an anti-inflammatory agent.

[0269] The therapeutic moiety can comprise dexamethasone (Dex).

[0270] The therapeutic moiety can comprise a therapeutic protein. For example, some people have defects in certain enzymes (e.g., lysosomal storage disease). Such enzymes/proteins can be delivered to human cells by linking the enzyme/protein to a cyclic cell penetrating peptide (cCPP) disclosed herein. The disclosed cCPP have been tested with proteins (e.g., GFP, PTP1B, actin, calmodulin, troponin C) and shown to work.

[0271] The therapeutic moiety can be an anti-infective agent. The term “anti-infective agent” refers to agents that are capable of killing, inhibiting, or otherwise slowing the growth of an infectious agent. The term “infectious agent” refers to pathogenic microorganisms, such as bacteria, viruses, fungi, and intracellular or extracellular parasites. The anti-infective agent can be used to treat an infectious disease, as infectious diseases are caused by infectious agents.

[0272] The infectious agent can be a Gram-negative bacteria. The Gram-negative bacteria can be of a genus selected from Escherichia, Proteus, Salmonella, Klebsiella, Providencia, Enterobacter, Burkholderia, Pseudomonas, Acinetobacter, Aeromonas, Haemophilus, Yersinia, Neisseria, Erwinia, Rhodopseudomonas and Burkholderia. The infectious agent can be a Gram-positive bacteria. The Gram-positive bacteria can be of a genus selected from Lactobacillus, Azorhizobium, Streptococcus, Pediococcus, Photobacterium, Bacillus, Enterococcus, Staphylococcus, Clostridium, Butyrivibrio, Sphingomonas, Rhodococcus and Streptomyces. The infectious agent can be an acid-fast bacteria of the Mycobacterium genus, such as Mycobacterium tuberculosis, Mycobacterium bovis. Mycobacterium avium and Mycobacterium leprae. The infectious agent

can be of the genus *Nocardia*. The infectious agent can be selected from any one of the following species *Nocardia asteroides*, *Nocardia brasiliensis* and *Nocardia caviae*.

[0273] The infectious agent can be a fungus. The fungus can be from the genus *Mucor*. The fungus can be from the genus *Cryptococcus*. The fungus can be from the genus *Candida*. The fungus can be selected from any one of *Mucor racemosus*, *Candida albicans*, *Cryptococcus neoformans*, or *Aspergillus fumigatus*.

[0274] The infectious agent can be a protozoa. The protozoa can be of the genus *Plasmodium* (e.g., *P. falciparum*, *P. vivax*, *P. ovale*, or *P. malariae*). The protozoa causes malaria.

[0275] Illustrative organisms include *Bacillus*, *Bartonella*, *Bordetella*, *Borrelia*, *Brucella*, *Campylobacter*, *Chlamydia*, *Chlamydophila*, *Clostridium*, *Corynebacterium*, *Enterococcus*, *Escherichia*, *Francisella*, *Haemophilus*, *Helicobacter*, *Legionella*, *Leptospira*, *Listeria*, *Mycobacterium*, *Mycoplasma*, *Neisseria*, *Pseudomonas*, *Rickettsia*, *Salmonella*, *Shigella*, *Staphylococcus*, *Streptococcus*, *Treponema*, *Ureaplasma*, *Vibrio*, and *Yersinia*.

[0276] The infectious agent can be a parasite. The parasite can be cryptosporidium. The parasite can be an endoparasite. The endoparasite can be heartworm, tapeworm, or flatworm. The parasite can be an epiparasite. The parasite causes a disease selected from acanthamoebiasis, babesiosis, balantidiasis, blastocystosis, coccidiosis, amoebiasis, giardiasis, isosporiasis, cystosporiasis, leishmaniasis, primary amoebic meningoencephalitis, malaria, rhinosporidiosis, toxoplasmosis, trichomoniasis, trypanomiasis, Chagas disease, or scabies.

[0277] The infectious agent can be a virus. Non-limiting examples of viruses include sudden acute respiratory coronavirus 2 (SARS-CoV-2), sudden acute respiratory coronavirus (SARS-CoV), Middle East Respiratory virus (MERS), influenza, Hepatitis C virus, Dengue virus, West Nile virus, Ebola virus, Hepatitis B, Human immunodeficiency virus (HIV), herpes simplex, Herpes zoster, and Lassa virus.

[0278] The anti-infective agent can be an antiviral agent. Non-limiting examples of antiviral agents include nucleoside or nucleotide reverse transcriptase inhibitors, such as zidovudine (AZT), didanosine (ddI), zalcitabine (ddC), stavudine (d4T), lamivudine (3TC), emtricitabine, abacavir succinate, elvucitabine, adefovir dipivoxil, lobucavir (BMS-180194) lodenosine (FddA) and tenofovir including tenofovir disoproxil and tenofovir disoproxil fumarate salt, non-nucleoside reverse transcriptase inhibitors, such as nevirapine, delaviradine, efavirenz, etravirine and rilpivirine, protease inhibitors, such as ritonavir, tipranavir, saquinavir, nelfinavir, indinavir,

amprenavir, fosamprenavir, atazanavir, lopinavir, darunavir (TMC-114), lasinavir and brecanavir (VX-385), cellular entry inhibitors, such as CCR5 antagonists (e.g., maraviroc, vicriviroc, INCB9471 and TAK-652) and CXCR4 antagonists (AMD-11070), fusion inhibitors, such as enfuvirtide, integrase inhibitors, such as raltegravir, BMS-707035, and elvitegravir, Tat inhibitors, such as didehydro-cortistatin A (dCA), maturation inhibitors, such as berivimat, immunomodulating agents, such as levamisole, and other antiviral agents, such as hydroxyurea, ribavirin, interleukin 2 (IL-2), interleukin 12 (IL-12), pensafuside, peramivir, zanamivir, oseltamivir phosphate, baloxavir marboxil,

[0279] The anti-infective agent can be an antibiotic. Non-limiting examples of antibiotics include aminoglycosides, such as amikacin, gentamicin, kanamycin, neomycin, netilmicin, streptomycin and tobramycin; cabecephems, such as loracarbef; carbapenems, such as ertapenem, imipenem/cilastatin and meropenem; cephalosporins, such as cefadroxil, cefazolin, cephalexin, cefaclor, cefamandole, cephalexin, cefoxitin, cefprozil, cefuroxime, cefixime, cefdinir, cefditoren, cefoperazone, cefotaxime, cefpodoxime, ceftazidime, ceftibuten, ceftizoxime, ceftriaxone and cefepime; macrolides, such as azithromycin, clarithromycin, dirithromycin, erythromycin and troleandomycin; monobactam; penicillins, such as amoxicillin, ampicillin, carbenicillin, cloxacillin, dicloxacillin, nafcillin, oxacillin, penicillin G, penicillin V, piperacillin and ticarcillin; polypeptides, such as bacitracin, colistin and polymyxin B; quinolones, such as ciprofloxacin, enoxacin, gatifloxacin, levofloxacin, lomefloxacin, moxifloxacin, norfloxacin, ofloxacin and trovafloxacin; sulfonamides, such as mafenide, sulfacetamide, sulfamethizole, sulfasalazine, sulfisoxazole and trimethoprim-sulfamethoxazole; tetracyclines, such as demeclocycline, doxycycline, minocycline, oxytetracycline and tetracycline; and vancomycin. The anti-infective agent can be a steroidal anti-inflammatory agent. Non-limiting examples of steroidal anti-inflammatory agents include fluocinolone, triamcinolone, triamcinoline acetonide, betamethasone, betamethasone dipropionate, diflucortolone, fluticasone, cortisone, hydrocortisone, mometasone, methylprednisolone, beclomethasone dipropionate, clobetasol, prednisone, prednisolone, methylprednisolone, betamethasone, budesonide, and dexamethasone. The anti-infective agent can be a non-steroidal anti-inflammatory agent. Non-limiting examples of non-steroidal anti-inflammatory agents include celecoxib, nimesulide, rofecoxib, meclofenamic acid, meclofenamate sodium, flunixin, fluprofen, flurbiprofen, sulindac, meloxicam, piroxicam, etodolac, fenoprofen,

fenbuprofen, ketoprofen, suprofen, diclofenac, bromfenac sodium, phenylbutazone, thalidomide and indomethacin.

[0280] The anti-infective agent can be an anti-fungal. Non-limiting examples of anti-fungals include amphotericin B, caspofungin, fluconazole, flucytosine, itraconazole, ketoconazole, amroline, butenafine, naftifine, terbinafine, elubiol, econazole, econazole, itraconazole, isoconazole, imidazole, miconazole, sulconazole, clotrimazole, enilconazole, oxiconazole, tioconazole, terconazole, butoconazole, thiabendazole, voriconazole, saperconazole, sertaconazole, fenticonazole, posaconazole, bifonazole, flutrimazole, nystatin, pimarin, natamycin, tolnaftate, mafenide, dapson, actofunicone, griseofulvin, potassium iodide, Gentian Violet, ciclopirox, ciclopirox olamine, haloprogin, undecylenate, silver sulfadiazine, undecylenic acid, undecylenic alkanolamide, and Carbol-Fuchsin.

[0281] The therapeutic moiety can be an analgesic or pain-relieving agent. Non-limiting examples of analgesics or pain-relieving agents include aspirin, acetaminophen, ibuprofen, naproxen, procaine, lidocaine, tetracaine, dibucaine, benzocaine, p-buthylaminobenzoic acid 2-(diethylamino) ethyl ester HCl, mepivacaine, piperocaine, and dyclonine.

[0282] The therapeutic moiety can be an antibody or an antigen-binding fragment. Antibodies and antigen-binding fragments can be derived from any suitable source, including human, mouse, camelid (e.g., camel, alpaca, llama), rat, ungulates, or non-human primates (e.g., monkey, rhesus macaque).

[0283] It should furthermore be understood that the cargos including anti-infective agents and other therapeutic moieties described herein include possible salts thereof, of which pharmaceutically acceptable salts are of course especially relevant for the therapeutic applications. Salts include acid addition salts and basic salts. Examples of acid addition salts are hydrochloride salts, fumarate, oxalate, etc. Examples of basic salts are salts where the (remaining) counter ion can be selected from alkali metals, such as sodium and potassium, alkaline earth metals, such as calcium salts, potassium salts, and ammonium ions ($^+N(R')_4$, where the R's independently designate optionally substituted C₁₋₆-alkyl, optionally substituted C₂₋₆-alkenyl, optionally substituted aryl, or optionally substituted heteroaryl).

[0284] The therapeutic moiety can be an oligonucleotide. The oligonucleotide can be an antisense compound (AC). The oligonucleotide can include, for example, but is not limited to, antisense oligonucleotides, small interfering RNA (siRNA), microRNA (miRNA), ribozymes, immune

stimulating nucleic acids, antagomir, antimir, microRNA mimic, supermir, UI adaptors, CRISPR machinery and aptamers. The term "antisense oligonucleotide" or simply "antisense" is meant to include oligonucleotides that are complementary to a targeted polynucleotide sequence. Non-limiting examples of antisense oligonucleotides for treating Duchenne muscular dystrophy may be found in US Pub. No. 2019/0365918, US Pub. No. US2020/0040336, US Pat. No. 9,499,818, and US Pat. No. 9,447,417 each of which is incorporated by reference in its entirety for all purposes.

[0285] The therapeutic moiety can be used to treat any one of the following diseases: neuromuscular disorders, Pompe disease, β -thalassemia, dystrophin Kobe, Duchenne muscular dystrophy, Becker muscular dystrophy, diabetes, Alzheimer's disease, cancer, cystic fibrosis, Merosin-deficient congenital muscular dystrophy type 1A (MDC1A), proximal spinal muscular atrophy (SMA), Huntington's disease, Huntington disease-like 2 (HDL2), myotonic dystrophy, spinocerebellar ataxia, spinal and bulbar muscular atrophy (SBMA), dentatorubral-pallidoluyian atrophy (DRPLA), amyotrophic lateral sclerosis, frontotemporal dementia, Fragile X syndrome, fragile X mental retardation 1 (FMR1), fragile X mental retardation 2 (FMR2), Fragile XE mental retardation (FRAXE), Friedreich's ataxia (FRDA), fragile X-associated tremor/ataxia syndrome (FXTAS), myoclonic epilepsy, oculopharyngeal muscular dystrophy (OPMD), syndromic or non-syndromic X-linked mental retardation, myotonic dystrophy, myotonic dystrophy type 1, myotonic dystrophy type 2, epilepsy, Dravet syndrome, or Alzheimer's disease. The therapeutic moiety can be used to treat a cancer selected from glioma, acute myeloid leukemia, thyroid cancer, lung cancer, colorectal cancer, head and neck cancer, stomach cancer, liver cancer, pancreatic cancer, renal cancer, urothelial cancer, prostate cancer, testis cancer, breast cancer, cervical cancer, endometrial cancer, ovarian cancer, or melanoma. The therapeutic moiety can be used to treat an ocular disease. Non-limiting examples of ocular diseases include refractive errors, macular degeneration, cataracts, diabetic retinopathy, glaucoma, amblyopia, or strabismus.

[0286] The therapeutic moiety can comprise a targeting moiety. The targeting moiety can comprise, for example, a sequence of amino acids that can target one or more enzyme domains. The targeting moiety can comprise an inhibitor against an enzyme that can play a role in a disease, such as cancer, cystic fibrosis, diabetes, obesity, or combinations thereof. The targeting moiety targets one or more of the following genes: *FMR1*, *AFF2*, *FXN*, *DMPK*, *SCA8*, *PPP2R2B*, *ATN1*, *DRPLA*, *HTT*, *AR*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *TBP*, *ATP7B*, *HTT*, *SCN1A*, *BRCA1*, *LAMA2*, *CD33*, *VEGF*, *ABCA4*, *CEP290*, *RHO*, *USH2A*, *OPA1*, *CNGB3*, *PRPF31*, *GYS1*,

or *RPGR*. The therapeutic moiety can be an antisense compound (AC) described in U.S. Publication No. 2019/0365918, which is incorporated by reference herein in its entirety. For example, the targeting moiety can comprise any of the sequences listed in Table 7.

Table 7. Example targeting moieties

Abbreviation *	Sequence
PQGAYR	Pro-Pip-Gly-F ₂ Pmp-Tyr-Arg
SOIARR	Ser-Pip-Ile-F ₂ Pmp-F ₂ Pmp-Arg
IHIAIR	Ile-His-Ile-F ₂ Pmp-Ile-Arg
AalAOR	Ala-(D-Ala)-Ile-F ₂ Pmp-Pip-Arg
ΣSOAvR	Fpa-Ser-Pip-F ₂ Pmp-(D-Val)-Arg
OnPAAAR	Pip-(D-Asn)-Pro-F ₂ Pmp-Ala-Arg
TΨAAGR	Tyr-Phg-Ala-F ₂ Pmp-Gly-Arg
AHIAaR	Ala-His-Ile-F ₂ Pmp-(D-Ala)-Arg
GnGApR	Gly-(D-Asn)-Gly-F ₂ Pmp-(D-Pro)-Arg
fQΘAIR	(D-Phe)-Gln-Pip-F ₂ Pmp-Ile-Arg
SPGΔHR	Ser-Pro-Gly-F ₂ Pmp-His-Arg
ΘYIAHR	Pip-Tyr-Ile-F ₂ Pmp-His-Arg
SvPAHR	Ser-(D-Val)-Pro-F ₂ Pmp-His-Arg
AIPAnR	Ala-Ile-Pro-F ₂ Pmp-(D-Asn)-Arg
ΣSIAQF	Fpa-Ser-Ile-F ₂ Pmp-Gln-Arg
AaΨAIR	Ala-(D-Ala)-Phg-F ₂ Pmp-(D-Phe)-Arg
ntΨAΨR	(D-Asn)-(D-Thr)-Phg-F ₂ Pmp-Phg-Arg
IPΨAΩR	Ile-Pro-Phg-F ₂ Pmp-Nle-Arg
QΘΣAOR	Gln-Pip-Fpa-F ₂ Pmp-Pip-Arg
nAΣAGR	(D-Asn)-Ala-Fpa-F ₂ Pmp-Gly-Arg
ntYAAAR	(D-Asn)-(D-Thr)-Tyr-F ₂ Pmp-Ala-Arg
eAΨAvR	(D-Glu)-Ala-Phg-F ₂ Pmp-(D-Val)-Arg
IvΨAAR	Ile-(D-Val)-Phg-F ₂ Pmp-Ala-Arg
YtΨAAR	Tyr-(D-Thr)-Phg-F ₂ Pmp-Ala-Arg
nΘΨAIR	(D-Asn)-Pip-Phg-F ₂ Pmp-Ile-Arg
ΘnWΔHR	Pip-(D-Asn)-Trp-F ₂ Pmp-His-Arg
YΘvAIR	Tyr-Pip-(D-Val)-F ₂ Pmp-Ile-Arg
nSAAGR	(D-Asn)-Ser-(D-Ala)-F ₂ Pmp-Gly-Arg
tnvAaR	(D-Thr)-(D-Asn)-(D-Val)-F ₂ Pmp-(D-Ala)-Arg
ntvAtR	(D-Asn)-(D-Thr)-(D-Val)-F ₂ Pmp-(D-Thr)-Arg
StAYR	Ser-Ile-(D-Thr)-F ₂ Pmp-Tyr-Arg
nΣnAIR	(D-Asn)-Fpa-(D-Asn)-F ₂ Pmp-(D-Leu)-Arg
YnnAΩR	Tyr-(D-Asn)-(D-Asn)-F ₂ Pmp-Nle-Arg
nYnAGR	(D-Asn)-Tyr-(D-Asn)-F ₂ Pmp-Gly-Arg
AWnAAR	Ala-Trp-(D-Asn)-F ₂ Pmp-Ala-Arg
vtHAYR	(D-Val)-(D-Thr)-His-F ₂ Pmp-Tyr-Arg
PΨHAOR	Pro-Phg-His-F ₂ Pmp-Pip-Arg
nΨHAAGR	(D-Asn)-Phg-His-F ₂ Pmp-Gly-Arg
PAHAGR	Pro-Ala-His-F ₂ Pmp-Gly-Arg
AYHAIR	Ala-Tyr-His-F ₂ Pmp-Ile-Arg
nΘeAYR	(D-Asn)-Pip-(D-Glu)-F ₂ Pmp-Tyr-Arg
vSSAtR	(D-Val)-Ser-Ser-F ₂ Pmp-(D-Thr)-Arg
aΞt' θ Φ'YNNK	((D-Ala)-Sar-(D-pThr)-Pp-Nal-Tyr-Gln)-Lys
Tm(aΞt' θ Φ'RA)Dap	Tm((D-Ala)-Sar-(D-pThr)-Pp-Nal-Arg-Ala)-Dap

Abbreviation *	Sequence
Tm(a Ξ t' Θ 'RAa)Dap	Tm((D-Ala)-Sar-(D-pThr)-Pp-Nal-Arg-Ala-(D-Ala))-Dap
Tm(a Ξ t' Φ 'RAa)Dap	Tm((D-Ala)-Sar-(D-Thr)-Pp-Nal-Arg-Ala-(D-Ala))-Dap
Tm(a Ξ ta Φ 'RAa)Dap	Tm((D-Ala)-Sar-(D-Thr)-(D-Ala)-Nal-Arg-Ala-(D-Ala))-Dap

*Fpa, Σ : L-4-fluorophenylalanine; Pip, Θ : L-homoproline; Nle, Ω : L-norleucine; Phg, Ψ L-phenylglycine; F₂Pmp, Λ : L-4-(phosphonodifluoromethyl)phenylalanine; Dap, L-2,3-diaminopropionic acid; Nal, Φ ': L- β -naphthylalanine; Pp, ϑ : L-pipecolic acid; Sar, Ξ : sarcosine; Tm, trimethyl acid.

5 [0287] The targeting moiety and cell penetrating peptide can overlap. That is, the residues that form the cell penetrating peptide can also be part of a sequence that forms a targeting moiety, and vice versa.

[0288] The therapeutic moiety can be attached to the cell penetrating peptide at the amino group, the carboxylate group, or the side chain of any of the amino acids of the cell penetrating peptide
10 (e.g., at the amino group, the carboxylate group, or the side chain or any of amino acid of the cCPP). The therapeutic moiety can be attached to a detectable moiety.

[0289] The therapeutic moiety can comprise a targeting moiety that can act as an inhibitor against Ras (e.g., K-Ras), PTP1B, Pin1, Grb2 SH2, CAL PDZ, and the like, or combinations thereof.

[0290] Ras is a protein that in humans is encoded by the RAS gene. The normal Ras protein
15 performs an essential function in normal tissue signaling, and the mutation of a Ras gene is implicated in the development of many cancers. Ras can act as a molecular on/off switch, once it is turned on Ras recruits and activates proteins necessary for the propagation of growth factor and other receptors' signal. Mutated forms of Ras have been implicated in various cancers, including lung cancer, colon cancer, pancreatic cancer, and various leukemias.

20 [0291] Protein-tyrosine phosphatase 1B (PTP1B) is a prototypical member of the PTP superfamily and plays numerous roles during eukaryotic cell signaling. PTP1B is a negative regulator of the insulin signaling pathway, and is considered a promising potential therapeutic target, in particular for the treatment of type II diabetes. PIP1B has also been implicated in the development of breast cancer.

25 [0292] Pin1 is an enzyme that binds to a subset of proteins and plays a role as a post phosphorylation control in regulating protein function. Pin1 activity can regulate the outcome of proline-directed kinase signaling and consequently can regulate cell proliferation and cell survival. Deregulation of Pin1 can play a role in various diseases. The up-regulation of Pin1 may be

implicated in certain cancers, and the down-regulation of Pin1 may be implicated in Alzheimer's disease. Inhibitors of Pin1 can have therapeutic implications for cancer and immune disorders.

5 [0293] Grb2 is an adaptor protein involved in signal transduction and cell communication. The Grb2 protein contains one SH2 domain, which can bind tyrosine phosphorylated sequences. Grb2 is widely expressed and is essential for multiple cellular functions. Inhibition of Grb2 function can impair developmental processes and can block transformation and proliferation of various cell types.

[0294] It was recently reported that the activity of cystic fibrosis membrane conductance regulator (CFTR), a chloride ion channel protein mutated in cystic fibrosis (CF) patients, is negatively regulated by CFTR-associated ligand (CAL) through its PDZ domain (CAL-PDZ) (Wolde, M et al. *J. Biol. Chem.* 2007, 282, 8099). Inhibition of the CFTR/CAL-PDZ interaction was shown to improve the activity of Δ Phe508-CFTR, the most common form of CFTR mutation (Cheng, SH et al. *Cell* 1990, 63, 827; Kerem, BS et al. *Science* 1989, 245, 1073), by reducing its proteasome-mediated degradation (Cushing, PR et al. *Angew. Chem. Int. Ed.* 2010, 49, 9907). Thus, disclosed herein is a method for treating a subject having cystic fibrosis by administering an effective amount of a compound or composition disclosed herein. The compound or composition administered to the subject can comprise a therapeutic moiety that can comprise a targeting moiety that can act as an inhibitor against CAL PDZ. Also, the compositions or compositions disclosed herein can be administered with a molecule that corrects the CFTR function.

20 [0295] The therapeutic moiety can be attached to the cyclic peptide at an amino group or carboxylate group, or a side chain of any of the amino acids of the cyclic peptide (e.g., at an amino group or the carboxylate group on the side chain of an amino acid of the cyclic peptide). In some examples, the therapeutic moiety can be attached to a detectable moiety.

[0296] Also disclosed herein are compositions comprising the compounds described herein.

25 [0297] Also disclosed herein are pharmaceutically-acceptable salts and prodrugs of the disclosed compounds. Pharmaceutically-acceptable salts include salts of the disclosed compounds that are prepared with acids or bases, depending on the particular substituents found on the compounds. Under conditions where the compounds disclosed herein are sufficiently basic or acidic to form stable nontoxic acid or base salts, administration of the compounds as salts can be appropriate. Examples of pharmaceutically-acceptable base addition salts include sodium, potassium, calcium, ammonium, or magnesium salt. Examples of physiologically-acceptable acid addition salts include

hydrochloric, hydrobromic, nitric, phosphoric, carbonic, sulfuric, and organic acids like acetic, propionic, benzoic, succinic, fumaric, mandelic, oxalic, citric, tartaric, malonic, ascorbic, alpha-ketoglutaric, alpha-glycophosphoric, maleic, tosyl acid, methanesulfonic, and the like. Thus, disclosed herein are the hydrochloride, nitrate, phosphate, carbonate, bicarbonate, sulfate, acetate, 5 propionate, benzoate, succinate, fumarate, mandelate, oxalate, citrate, tartarate, malonate, ascorbate, alpha-ketoglutarate, alpha-glycophosphate, maleate, tosylate, and mesylate salts. Pharmaceutically acceptable salts of a compound can be obtained using standard procedures well known in the art, for example, by reacting a sufficiently basic compound such as an amine with a suitable acid affording a physiologically acceptable anion. Alkali metal (for example, sodium, 10 potassium or lithium) or alkaline earth metal (for example calcium) salts of carboxylic acids can also be made.

[0298] The therapeutic moiety can include a therapeutic polypeptide, an oligonucleotide or a small molecule. The therapeutic polypeptide can include a peptide inhibitor. The therapeutic polypeptide can include a binding reagent that specifically binds to a target of interest. The binding reagent can 15 include an antibody or antigen-binding fragment thereof that specifically binds to a target of interest. The antigen-binding fragments can include a Fab fragment, a F(ab') fragment, a F(ab')₂ fragment, a Fv fragment, a minibody, a diabody, a nanobody, a single domain antibody (dAb), a single-chain variable fragment (scFv), or a multispecific antibody.

[0299] The oligonucleotide can include an antisense compound (AC). The AC can include a 20 nucleotide sequence complementary to a target nucleotide sequence encoding a protein target of interest.

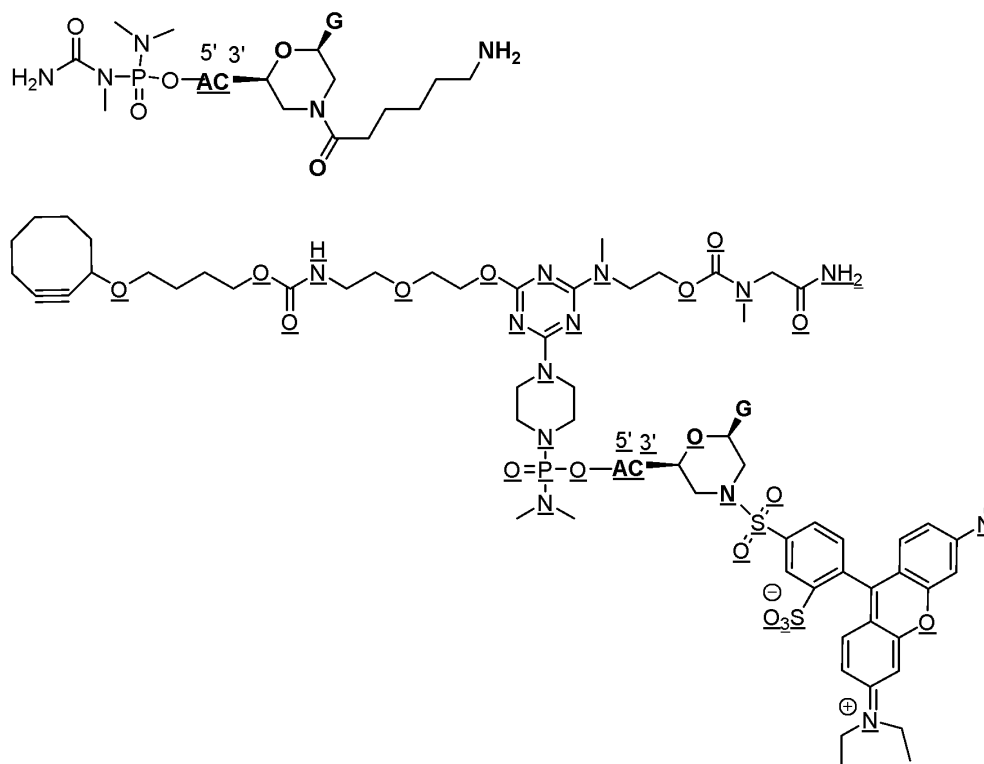
[0300] The therapeutic moiety (TM) can be conjugated to a chemically reactive side chain of an amino acid of the cCPP. Any amino acid side chain on the cCPP which is capable of forming a covalent bond, or which may be so modified, can be used to link the TM to the cCPP. The amino 25 acid on the cCPP can be a natural or non-natural amino acid. The chemically reactive side chain can include an amine group, a carboxylic acid, an amide, a hydroxyl group, a sulfhydryl group, a guanidiny group, a phenolic group, a thioether group, an imidazolyl group, or an indolyl group. The amino acid of the cCPP to which the TM is conjugated can include lysine, arginine, aspartic acid, glutamic acid, asparagine, glutamine, serine, threonine, tyrosine, cysteine, arginine, tyrosine, 30 methionine, histidine, tryptophan or analogs thereof. The amino acid on the cCPP used to conjugate the TM can be ornithine, 2,3-diaminopropionic acid, or analogs thereof. The amino acid

can be lysine, or an analog thereof. The amino acid can be glutamic acid, or an analog thereof. The amino acid can be aspartic acid, or an analog thereof. The side chain can be substituted with a bond to the TM or a linker.

[0301] The TM can include a therapeutic polypeptide and the cCPP can be conjugated to a chemically reactive side chain of an amino acid of the therapeutic polypeptide. Any amino acid side chain on the TM which is capable of forming a covalent bond, or which may be so modified, can be used to link the cCPP to the TM. The amino acid on the TM can be a natural or non-natural amino acid. The chemically reactive side chain can include an amine group, a carboxylic acid, an amide, a hydroxyl group, a sulfhydryl group, a guanidinyll group, a phenolic group, a thioether group, an imidazolyl group, or an indolyl group. The amino acid of the TM to which the cCPP is conjugated can include lysine, arginine, aspartic acid, glutamic acid, asparagine, glutamine, serine, threonine, tyrosine, cysteine, arginine, tyrosine, methionine, histidine, tryptophan or analogs thereof. The amino acid on the TM used to conjugate the cCPP can be ornithine, 2,3-diaminopropionic acid, or analogs thereof. The amino acid can be lysine, or an analog thereof. The amino acid can be glutamic acid, or an analog thereof. The amino acid can be aspartic acid, or an analog thereof. The side chain of the TM can be substituted with a bond to the cCPP or a linker.

[0302] The TM can be an antisense compound (AC) that includes an oligonucleotide where the 5' or 3' end of the oligonucleotide is conjugated to a chemically reactive side chain of an amino acid of the cCPP. The AC can be chemically conjugated to the cCPP through a moiety on the 5' or 3' end of the AC. The chemically reactive side chain of the cCPP can include an amine group, a carboxylic acid, an amide, a hydroxyl group, a sulfhydryl group, a guanidinyll group, a phenolic group, a thioether group, an imidazolyl group, or an indolyl group. The amino acid of the cCPP to which the AC is conjugated can include lysine, arginine, aspartic acid, glutamic acid, asparagine, glutamine, serine, threonine, tyrosine, cysteine, arginine, tyrosine, methionine, histidine or tryptophan. The amino acid of the cCPP to which the AC is conjugated can include lysine or cysteine.

[0303] Non-limiting examples of unconjugated AC structures (i.e. prior to conjugation to the CPP) are provided below. AC in the structures below refers to an antisense oligonucleotide.



[0304] Non-limiting examples of linear CPPs include Polyarginine (e.g., R₉ or R₁₁), Antennapedia sequences, HIV-TAT, Penetratin, Antp-3A (Antp mutant), Buforin II, Transportan, MAP (model amphipathic peptide), K-FGF, Ku70, Prion, pVEC, Pep-1, SynB1, Pep-7, HN-1, BGSC (Bis-Guanidinium-Spermidine-Cholesterol, and BGTC (Bis-Guanidinium-Tren-Cholesterol).

5 *Oligonucleotides*

[0305] The compounds can include a cyclic cell penetrating peptide (cCPP) conjugated to an antisense compound (AC) as the therapeutic moiety. The AC can include an antisense oligonucleotide, siRNA, microRNA, antagomir, aptamer, ribozyme, immunostimulatory oligonucleotide, decoy oligonucleotide, supermir, miRNA mimic, miRNA inhibitor, or
10 combinations thereof.

Antisense Oligonucleotides

[0306] The therapeutic moiety can include an antisense oligonucleotide. The term "antisense oligonucleotide" or simply "antisense" refers to oligonucleotides that are complementary to a targeted polynucleotide sequence. Antisense oligonucleotides can include single strands of DNA
15 or RNA that are complementary to a chosen sequence, e.g. a target gene mRNA.

[0307] The antisense oligonucleotides may modulate one or more aspects of protein transcription, translation, and expression and functions via hybridization of the antisense oligonucleotide with a target nucleic acid. Hybridization of the antisense oligonucleotide to its target sequence can suppress expression of the target protein. Hybridization of the antisense oligonucleotide to its
20 target sequence can suppress expression of one or more target protein isoforms. Hybridization of the antisense oligonucleotide to its target sequence can upregulate expression of the target protein. Hybridization of the antisense oligonucleotide to its target sequence can downregulate expression of the target protein.

[0308] The antisense compound can inhibit gene expression by binding to a complementary
25 mRNA. Binding to the target mRNA can lead to inhibition of gene expression either by preventing translation of complementary mRNA strands by binding to it or by leading to degradation of the target mRNA. Antisense DNA can be used to target a specific, complementary (coding or non-coding) RNA. If binding takes place this DNA/RNA hybrid can be degraded by the enzyme RNase H. The antisense oligonucleotide can include from about 10 to about 50 nucleotides, about
30 15 to about 30 nucleotides, or about 20 to about 25 nucleotides. The term also encompasses

antisense oligonucleotides that may not be fully complementary to the desired target gene. Thus, compounds disclosed herein can be utilized in instances where non-target specific-activities are found with antisense, or where an antisense sequence containing one or more mismatches with the target sequence is desired.

5 [0309] Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, can be used to specifically inhibit protein synthesis by a targeted gene. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established.

[0310] Methods of producing antisense oligonucleotides are known in the art and can be readily
10 adapted to produce an antisense oligonucleotide that targets any polynucleotide sequence of interest. Selection of antisense oligonucleotide sequences specific for a given target sequence is based upon analysis of the chosen target sequence and determination of secondary structure, T_m , binding energy, and relative stability. Antisense oligonucleotides may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or
15 prohibit specific binding to the target mRNA in a host cell. Target regions of the mRNA include those regions at or near the AUG translation initiation codon and those sequences that are substantially complementary' to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software (Molecular Biology Insights) and/or the BLASTN 2.0.5 algorithm software
20 (Altschul et al, Nucleic Acids Res. 1997, 25(17):3389-402).

RNA Interference Nucleic Acids

[0311] The therapeutic moiety can be a RNA interference (RNAi) molecule or a small interfering RNA molecule. RNA interference methods using RNAi or siRNA molecules may be used to disrupt the expression of a gene or polynucleotide of interest.

25 [0312] Small interfering RNAs (siRNAs) are RNA duplexes normally from about 16 to about 30 nucleotides long that can associate with a cytoplasmic multi-protein complex known as RNAi-induced silencing complex (RISC). RISC loaded with siRNA mediates the degradation of homologous mRNA transcripts, therefore siRNA can be designed to knock down protein expression with high specificity. Unlike other antisense technologies, siRNA function through a
30 natural mechanism evolved to control gene expression through non-coding RNA. A variety of RNAi reagents, including siRNAs targeting clinically relevant targets, are currently under

pharmaceutical development, as described, e.g., in de Fougères, A. et al., Nature Reviews 6:443-453 (2007).

[0313] While the first described RNAi molecules were RNA:RNA hybrids that include both an RNA sense and an RNA antisense strand, it has now been demonstrated that DNA sense:RNA antisense hybrids, RNA sense:DNA antisense hybrids, and DNA:DNA hybrids are capable of mediating RNAi (Lamberton, J.S. and Christian, A.T., (2003) Molecular Biotechnology 24:111-119). RNAi molecules can be used that include any of these different types of double-stranded molecules. In addition, it is understood that RNAi molecules may be used and introduced to cells in a variety of forms. RNAi molecules can encompass any and all molecules capable of inducing an RNAi response in cells, including, but not limited to, double-stranded oligonucleotides that include two separate strands, i.e. a sense strand and an antisense strand, e.g., small interfering RNA (siRNA); double-stranded oligonucleotide that includes two separate strands that are linked together by non-nucleotidyl linker; oligonucleotides that include a hairpin loop of complementary sequences, which forms a double-stranded region, e.g., shRNAi molecules, and expression vectors that express one or more polynucleotides capable of forming a double-stranded polynucleotide alone or in combination with another polynucleotide.

[0314] A "single strand siRNA compound" as used herein, is an siRNA compound which is made up of a single molecule. It may include a duplexed region, formed by intra-strand pairing, e.g., it may be, or include, a hairpin or pan-handle structure. Single strand siRNA compounds may be antisense with regard to the target molecule.

[0315] A single strand siRNA compound may be sufficiently long that it can enter the RISC and participate in RISC mediated cleavage of a target mRNA. A single strand siRNA compound is at least about 14, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, or up to about 50 nucleotides in length. The single strand siRNA is less than about 200, about 100, or about 60 nucleotides in length.

[0316] Hairpin siRNA compounds may have a duplex region equal to or at least about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, or about 25 nucleotide pairs. The duplex region may be equal to or less than about 200, about 100, or about 50 nucleotide pairs in length. Ranges for the duplex region are from about 15 to about 30, from about 17 to about 23, from about 19 to about 23, and from about 19 to about 21 nucleotides pairs in length. The hairpin may have a single strand overhang or terminal unpaired region. The overhangs may be from about

2 to about 3 nucleotides in length. The overhang can be at the sense side of the hairpin or on the antisense side of the hairpin.

5 [0317] A "double stranded siRNA compound" as used herein, is an siRNA compound which includes more than one, and in some cases two, strands in which interchain hybridization can form a region of duplex structure.

[0318] The antisense strand of a double stranded siRNA compound may be equal to or at least about 14, about 15, about 16 about 17, about 18, about 19, about 20, about 25, about 30, about 40, or about 60 nucleotides in length. It may be equal to or less than about 200, about 100, or about 50 nucleotides in length. Ranges may be from about 17 to about 25, from about 19 to about 23, and 10 from about 19 to about 21 nucleotides in length. As used herein, term "antisense strand" means the strand of an siRNA compound that is sufficiently complementary to a target molecule, e.g. a target RNA.

[0319] The sense strand of a double stranded siRNA compound may be equal to or at least about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 25, about 30, about 40, or 15 about 60 nucleotides in length. It may be equal to or less than about 200, about 100, or about 50, nucleotides in length. Ranges may be from about 17 to about 25, from about 19 to about 23, and from about 19 to about 21 nucleotides in length.

[0320] The double strand portion of a double stranded siRNA compound may be equal to or at least about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, 20 about 23, about 24, about 25, about 30, about 40, or about 60 nucleotide pairs in length. It may be equal to or less than about 200, about 100, or about 50, nucleotides pairs in length. Ranges may be from about 15 to about 30, from about 17 to about 23, from about 19 to about 23, and from about 19 to about 21 nucleotides pairs in length.

[0321] The siRNA compound can be sufficiently large that it can be cleaved by an endogenous 25 molecule, e.g., by Dicer, to produce smaller siRNA compounds, e.g., siRNAs agents.

[0322] The sense and antisense strands may be chosen such that the double-stranded siRNA compound includes a single strand or unpaired region at one or both ends of the molecule. Thus, a double- stranded siRNA compound may contain sense and antisense strands, paired to contain an overhang, e.g., one or two 5' or 3' overhangs, or a 3' overhang of 1 to 3 nucleotides. The overhangs 30 can be the result of one strand being longer than the other, or the result of two strands of the same

length being staggered. Some embodiments will have at least one 3' overhang. In embodiments, both ends of an siRNA molecule will have a 3' overhang. The overhang can be 2 nucleotides.

[0323] The length for the duplexed region can be from about 15 to about 30, or about 18, about 19, about 20, about 21, about 22, or about 23 nucleotides in length, e.g., in the ssiRNA (siRNA with sticky overhangs) compound range discussed above. ssiRNA compounds can resemble in length and structure the natural Dicer processed products from long dsRNAs. Embodiments in which the two strands of the ssiRNA compound are linked, e.g., covalently linked are also included. Hairpin, or other single strand structures which provide a double stranded region, and a 3' overhang are included.

[0324] The siRNA compounds described herein, including double-stranded siRNA compounds and single-stranded siRNA compounds can mediate silencing of a target RNA, e.g., mRNA, e.g., a transcript of a gene that encodes a protein. For convenience, such mRNA is also referred to herein as mRNA to be silenced. Such a gene is also referred to as a target gene. In general, the RNA to be silenced is an endogenous gene.

[0325] As used herein, the phrase "mediates RNAi" refers to the ability to silence, in a sequence specific manner, a target RNA. While not wishing to be bound by theory, it is believed that silencing uses the RNAi machinery or process and a guide RNA, e.g., an ssiRNA compound of from about 21 to about 23 nucleotides.

[0326] An siRNA compound that is "sufficiently complementary" to a target RNA, e.g., a target mRNA, can silence production of protein encoded by the target mRNA. A siRNA compound that is "sufficiently complementary" to the RNA encoding a protein of interest, can silence production of the protein of interest encoded by the mRNA. The siRNA compound can be "exactly complementary" to a target RNA, e.g., the target RNA and the siRNA compound anneal, for example to form a hybrid made exclusively of Watson-Crick base pairs in the region of exact complementarity. A "sufficiently complementary" target RNA can include an internal region (e.g., of at least about 10 nucleotides) that is exactly complementary to a target RNA. In embodiments, the siRNA compound specifically discriminates a single-nucleotide difference. In this case, the siRNA compound only mediates RNAi if exact complementarity is found in the region (e.g., within 7 nucleotides of) the single-nucleotide difference.

MicroRNAs

[0327] The therapeutic moiety can be a microRNA molecule. MicroRNAs (miRNAs) are a highly conserved class of small RNA molecules that are transcribed from DNA in the genomes of plants and animals, but are not translated into protein. Processed miRNAs are single stranded 17-25 nucleotide (nt) RNA molecules that become incorporated into the RNA-induced silencing complex (RISC) and have been identified as key regulators of development, cell proliferation, apoptosis and differentiation. They are believed to play a role in regulation of gene expression by binding to the 3'-untranslated region of specific mRNAs. RISC mediates down-regulation of gene expression through translational inhibition, transcript cleavage, or both. RISC is also implicated in transcriptional silencing in the nucleus of a wide range of eukaryotes.

Antagomirs

[0328] The therapeutic moiety can be an antagomir. Antagomirs are RNA-like oligonucleotides that harbor various modifications for RNase protection and pharmacologic properties, such as enhanced tissue and cellular uptake. They differ from normal RNA by, for example, complete 2'-O-methylation of sugar, phosphorothioate backbone and, for example, a cholesterol-moiety at 3'-end. Antagomirs may be used to efficiently silence endogenous miRNAs by forming duplexes that include the antagomir and endogenous miRNA, thereby preventing miRNA-induced gene silencing. An example of antagomir-mediated miRNA silencing is the silencing of miR-122, described in Krutzfeldt et al., Nature, 2005, 438: 685-689, which is expressly incorporated by reference herein in its entirety. Antagomir RNAs may be synthesized using standard solid phase oligonucleotide synthesis protocols. See U.S. Patent Application Ser. Nos. 11/502,158 and 11/657,341 (the disclosure of each of which are incorporated herein by reference).

[0329] An antagomir can include ligand-conjugated monomer subunits and monomers for oligonucleotide synthesis. Monomers are described in U.S. Application No. 10/916,185, filed on August 10, 2004. An antagomir can have a ZXY structure, such as is described in PCT Application No. PCT/US2004/07070 filed on March 8, 2004. An antagomir can be complexed with an amphipathic moiety. Amphipathic moieties for use with oligonucleotide agents are described in PCT Application No. PCT/US2004/07070, filed on March 8, 2004.

Aptamers

[0330] The therapeutic moiety can be an aptamer. Aptamers are nucleic acid or peptide molecules that bind to a particular molecule of interest with high affinity and specificity (Tuerk and Gold, Science 249:505 (1990); Ellington and Szostak, Nature 346:818 (1990)). DNA or RNA aptamers have been successfully produced which bind many different entities from large proteins to small organic molecules. See Eaton, Curr. Opin. Chem. Biol. 1: 10-16 (1997), Famulok, Curr. Opin. Struct. Biol. 9:324-9(1999), and Hermann and Patel, Science 287:820-5 (2000). Aptamers may be RNA or DNA based, and may include a riboswitch. A riboswitch is a part of an mRNA molecule that can directly bind a small target molecule, and whose binding of the target affects the gene's activity. Thus, an mRNA that contains a riboswitch is directly involved in regulating its own activity, depending on the presence or absence of its target molecule. Generally, aptamers are engineered through repeated rounds of in vitro selection or equivalently, SELEX (systematic evolution of ligands by exponential enrichment) to bind to various molecular targets such as small molecules, proteins, nucleic acids, and even cells, tissues and organisms. The aptamer may be prepared by any known method, including synthetic, recombinant, and purification methods, and may be used alone or in combination with other aptamers specific for the same target. Further, the term "aptamer" also includes "secondary aptamers" containing a consensus sequence derived from comparing two or more known aptamers to a given target. The aptamer can be an "intracellular aptamer", or "intramer", which specifically recognize intracellular targets. See Famulok et al., Chem Biol. 2001, Oct, 8(10):931-939; Yoon and Rossi, Adv Drug Deliv Rev. 2018, Sep, 134:22-35, each incorporated by reference herein.

Ribozymes

[0331] The therapeutic moiety can be a ribozyme. Ribozymes are RNA molecules complexes having specific catalytic domains that possess endonuclease activity (Kim and Cech, Proc Natl Acad Sci U S A. 1987 Dec;84(24):8788-92; Forster and Symons, Cell. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech et al, Cell. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, J Mol Biol. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, Nature. 1992 May 14;357(6374): 173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence("IGS") of the ribozyme prior to chemical reaction.

[0332] At least six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds in trans (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of an enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

[0333] The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif, for example. Specific examples of hammerhead motifs are described by Rossi et al. Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel et al. (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel et al, Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1 ;31(47): 11843-52; an example of the RNaseP motif is described by Guerrier-Takada et al , Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1 ;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in U. S. Patent 4,987,071. Enzymatic nucleic acid molecules can have a specific substrate binding site which is complementary to one or more of the target gene DNA or RNA regions, and that they have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

[0334] Methods of producing a ribozyme targeted to a polynucleotide sequence are known in the art. Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int.

Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference, and synthesized to be tested in vitro and in vivo, as described therein.

[0335] Ribozyme activity can be increased by altering the length of the ribozyme binding arms or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see e.g., Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711 ; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Immunostimulatory Oligonucleotides

[0336] The therapeutic moiety can be an immunostimulatory oligonucleotide. Immunostimulatory oligonucleotides (ISS; single-or double- stranded) are capable of inducing an immune response when administered to a patient, which may be a mammal or other patient. ISS include, e.g., certain palindromes leading to hairpin secondary structures (see Yamamoto S., et al. (1992) J. Immunol. 148: 4072-4076), or CpG motifs, as well as other known ISS features (such as multi-G domains, see WO 96/11266).

[0337] The immune response may be an innate or an adaptive immune response. The immune system is divided into a more innate immune system, and acquired adaptive immune system of vertebrates, the latter of which is further divided into humoral cellular components. The immune response may be mucosal.

[0338] Immunostimulatory nucleic acids are considered to be non-sequence specific when it is not required that they specifically bind to and reduce the expression of a target polynucleotide in order to provoke an immune response. Thus, certain immunostimulatory nucleic acids may include a sequence corresponding to a region of a naturally occurring gene or mRNA, but they may still be considered non-sequence specific immunostimulatory nucleic acids.

[0339] The immunostimulatory nucleic acid or oligonucleotide can include at least one CpG dinucleotide. The oligonucleotide or CpG dinucleotide may be unmethylated or methylated. The immunostimulatory nucleic acid can include at least one CpG dinucleotide having a methylated cytosine. The nucleic acid can include a single CpG dinucleotide, wherein the cytosine in said CpG dinucleotide is methylated. The nucleic acid can include the sequence 5' TAACGTTGAGGG'CAT

3'. The nucleic acid can include at least two CpG dinucleotides, wherein at least one cytosine in the CpG dinucleotides is methylated. Each cytosine in the CpG dinucleotides present in the sequence can be methylated. The nucleic acid can include a plurality of CpG dinucleotides, wherein at least one of said CpG dinucleotides includes a methylated cytosine.

- 5 [0340] Additional specific nucleic acid sequences of oligonucleotides (ODNs) suitable for use in the compositions and methods are described in Raney et al, Journal of Pharmacology and Experimental Therapeutics, 298:1185-1192 (2001). ODNs used in the compositions and methods can have a phosphodiester("PO") backbone or a phosphorothioate ("PS") backbone, and/or at least one methylated cytosine residue in a CpG motif.

10 *Decoy Oligonucleotides*

- [0341] The therapeutic moiety can be a decoy oligonucleotide. Because transcription factors recognize their relatively short binding sequences, even in the absence of surrounding genomic DNA, short oligonucleotides bearing the consensus binding sequence of a specific transcription factor can be used as tools for manipulating gene expression in living cells. This strategy involves the intracellular delivery of such "decoy oligonucleotides", which are then recognized and bound by the target factor. Occupation of the transcription factor's DNA-binding site by the decoy renders the transcription factor incapable of subsequently binding to the promoter regions of target genes. Decoys can be used as therapeutic agents, either to inhibit the expression of genes that are activated by a transcription factor, or to upregulate genes that are suppressed by the binding of a transcription factor. Examples of the utilization of decoy oligonucleotides may be found in Mann et al., J. Clin. Invest, 2000, 106: 1071-1075, which is expressly incorporated by reference herein, in its entirety.
- 15
20
Supermir

- [0342] The therapeutic moiety can be a supermir. A supermir refers to a single stranded, double stranded or partially double stranded oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or both or modifications thereof, which has a nucleotide sequence that is substantially identical to an miRNA and that is antisense with respect to its target, This term includes oligonucleotides composed of naturally-occurring nucleobases, sugars and covalent internucleoside (backbone) linkages and which contain at least one non-naturally- occurring portion which functions similarly. Such modified or substituted oligonucleotides have desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target and increased stability in the presence of nucleases. The supermir may not include a sense
- 25
30

strand. The supermir may not self-hybridize to a significant extent. A supermir can have secondary structure, but it is substantially single-stranded under physiological conditions. A supermir that is substantially single-stranded is single-stranded to the extent that less than about 50% {e.g., less than about 40%, about 30%, about 20%, about 10%, or about 5%) of the supermir is duplexed with itself. The supermir can include a hairpin segment, e.g., sequence, for example, at the 3' end can self hybridize and form a duplex region, e.g., a duplex region of at least about 1, about 2, about 3, or about 4 or less than about 8, about 7, about 6, or about 5 nucleotides, or about 5 nucleotides. The duplexed region can be connected by a linker, e.g., a nucleotide linker, e.g., about 3, about 4, about 5, or about 6 dTs, e.g., modified dTs. The supermir can be duplexed with a shorter oligo, e.g., of about 5, about 6, about 7, about 8, about 9, or about 10 nucleotides in length, e.g., at one or both of the 3' and 5' end or at one end and in the non-terminal or middle of the supermir.

miRNA mimics

[0343] The therapeutic moiety can be a miRNA mimic. miRNA mimics represent a class of molecules that can be used to imitate the gene silencing ability of one or more miRNAs. Thus, the term "microRNA mimic" refers to synthetic non-coding RNAs (i.e. the miRNA is not obtained by purification from a source of the endogenous miRNA) that are capable of entering the RNAi pathway and regulating gene expression. miRNA mimics can be designed as mature molecules (e.g. single stranded) or mimic precursors (e.g., pri- or pre-miRNAs). miRNA mimics can include nucleic acid (modified or modified nucleic acids) including oligonucleotides that include, without limitation, RNA, modified RNA, DNA, modified DNA, locked nucleic acids, or 2'-O,4'-C-ethylene-bridged nucleic acids (ENA), or any combination of the above (including DNA-RNA hybrids). In addition, miRNA mimics can include conjugates that can affect delivery, intracellular compartmentalization, stability, specificity, functionality, strand usage, and/or potency. In one design, miRNA mimics are double stranded molecules (e.g., with a duplex region of between about 16 and about 31 nucleotides in length) and contain one or more sequences that have identity with the mature strand of a given miRNA. Modifications can include 2' modifications (including 2'-O methyl modifications and 2' F modifications) on one or both strands of the molecule and internucleotide modifications (e.g. phosphorothioate modifications) that enhance nucleic acid stability and/or specificity. In addition, miRNA mimics can include overhangs. The overhangs can include from about 1 to about 6 nucleotides on either the 3' or 5' end of either strand and can be modified to enhance stability or functionality. The miRNA mimic can include a duplex region of

from about 16 to about 31 nucleotides and one or more of the following chemical modification patterns: the sense strand contains 2'-O-methyl modifications of nucleotides 1 and 2 (counting from the 5' end of the sense oligonucleotide), and all of the Cs and Us; the antisense strand modifications can include 2' F modification of all of the Cs and Us, phosphorylation of the 5' end of the oligonucleotide, and stabilized internucleotide linkages associated with a 2 nucleotide 3' overhang.
miRNA inhibitor

[0344] The therapeutic moiety can be a miRNA inhibitor. The terms "antimir" "microRNA inhibitor", "miR inhibitor", or "miRNA inhibitor" are synonymous and refer to oligonucleotides or modified oligonucleotides that interfere with the ability of specific miRNAs. In general, the inhibitors are nucleic acid or modified nucleic acids in nature including oligonucleotides that include RNA, modified RNA, DNA, modified DNA, locked nucleic acids (LNAs), or any combination of the above.

[0345] Modifications include 2' modifications (including 2'-O alkyl modifications and 2' F modifications) and internucleotide modifications (e.g. phosphorothioate modifications) that can affect delivery, stability, specificity, intracellular compartmentalization, or potency. In addition, miRNA inhibitors can include conjugates that can affect delivery, intracellular compartmentalization, stability, and/or potency. Inhibitors can adopt a variety of configurations including single stranded, double stranded (RNA/RNA or RNA/DNA duplexes), and hairpin designs, in general, microRNA inhibitors include contain one or more sequences or portions of sequences that are complementary or partially complementary with the mature strand (or strands) of the miRNA to be targeted, in addition, the miRNA inhibitor may also include additional sequences located 5' and 3' to the sequence that is the reverse complement of the mature miRNA. The additional sequences may be the reverse complements of the sequences that are adjacent to the mature miRNA in the pri-miRNA from which the mature miRNA is derived, or the additional sequences may be arbitrary sequences (having a mixture of A, G, C, or U). One or both of the additional sequences can be arbitrary sequences capable of forming hairpins. The sequence that is the reverse complement of the miRNA may be flanked on the 5' side and on the 3' side by hairpin structures. Micro-RNA inhibitors, when double stranded, may include mismatches between nucleotides on opposite strands. Furthermore, micro-RNA inhibitors may be linked to conjugate moieties in order to facilitate uptake of the inhibitor into a cell. For example, a micro-RNA inhibitor may be linked to cholesteryl 5-(bis(4-methoxyphenyl)(phenyl)methoxy)-3

hydroxypentylcarbamate) which allows passive uptake of a micro-RNA inhibitor into a cell. Micro-RNA inhibitors, including hairpin miRNA inhibitors, are described in detail in Vermeulen et al., "Double-Stranded Regions Are Essential Design Components Of Potent Inhibitors of RISC Function," RNA 13: 723- 730 (2007) and in WO2007/095387 and WO 2008/036825 each of which is incorporated herein by reference in its entirety. A person of ordinary skill in the art can select a sequence from the database for a desired miRNA and design an inhibitor useful for the methods disclosed herein.

Antisense compounds (ACs)

[0346] The therapeutic moiety includes an antisense compound (AC) that can alter one or more aspects of translation, or expression of a target gene. The principle behind antisense technology is that an antisense compound, which hybridizes to a target nucleic acid, modulates gene expression activities such as translation through one of a number of antisense mechanisms. Antisense technology is an effective means for changing the expression of one or more specific gene products and can therefore prove to be useful in a number of therapeutic, diagnostic, and research applications.

[0347] The compounds described herein may contain one or more asymmetric centers and thus give rise to enantiomers, diastereomers, and other stereoisomeric configurations that may be defined, in terms of absolute stereochemistry, as (R) or (S), α or β , or as (D) or (L). Included in the antisense compounds provided herein are all such possible isomers, as well as their racemic and optically pure forms.

Antisense compound hybridization site

[0348] Antisense mechanisms rely on hybridization of the antisense compound to the target nucleic acid.

[0349] The AC can hybridize with a sequence from about 5 to about 50 nucleic acids in length, which can also be referred to as the length of the AC. The AC can be from about 5 to about 10, from about 10 to about 15, from about 15 to about 20, from about 20 to about 25, from about 25 to about 30, from about 30 to about 35, from about 35 to about 40, from about 40 to about 45, or from about 45 to about 50 nucleic acids in length. The AC can be about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about

28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, or about 50 nucleic acids in length. The AC can be about 10 nucleic acids in length. The AC can be about 15 nucleic acids in length. The AC can be about 20 nucleic acids in length.

5 The AC can be about 25 nucleic acids in length. The AC can be about 30 nucleic acids in length.

[0350] The AC may be less than about 100 percent complementary to a target nucleic acid sequence. As used herein, the term "percent complementary" refers to the number of nucleobases of an AC that have nucleobase complementarity with a corresponding nucleobase of an oligomeric compound or nucleic acid divided by the total length (number of nucleobases) of the AC. One

10 skilled in the art recognizes that the inclusion of mismatches is possible without eliminating the activity of the antisense compound. The AC may contain up to about 20% nucleotides that disrupt base pairing of the AC to the target nucleic acid. The AC may contain no more than about 15%, no more than about 10%, no more than 5%, or no mismatches. The ACs can be at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%,

15 at least about 98%, at least about 99% or about 100% complementary to a target nucleic acid. Percent complementarity of an oligonucleotide is calculated by dividing the number of complementary nucleobases by the total number of nucleobases of the oligonucleotide. Percent complementarity of a region of an oligonucleotide is calculated by dividing the number of complementary nucleobases in the region by the total number of nucleobases region.

20 [0351] Incorporation of nucleotide affinity modifications can allow for a greater number of mismatches compared to an unmodified compound. Similarly, certain oligonucleotide sequences may be more tolerant to mismatches than other oligonucleotide sequences. One of ordinary skill in the art is capable of determining an appropriate number of mismatches between oligonucleotides, or between an oligonucleotide and a target nucleic acid, such as by determining

25 melting temperature (T_m). T_m or ΔT_m can be calculated by techniques that are familiar to one of ordinary skill in the art. For example, techniques described in Freier et al. (Nucleic Acids Research, 1997, 25, 22: 4429-4443) allow one of ordinary skill in the art to evaluate nucleotide modifications for their ability to increase the melting temperature of an RNA:DNA duplex.

Antisense mechanisms

30 [0352] The ACs according to the present disclosure may modulate one or more aspects of protein transcription, translation, and expression.

[0353] The AC can regulate transcription, translation, or protein expression through steric blocking. The following review article describes the mechanisms of steric blocking and applications thereof and is incorporated by reference herein in its entirety: Roberts et al. Nature Reviews Drug Discovery (2020) 19: 673-694.

5 [0354] The antisense mechanism functions via hybridization of an antisense compound with a target nucleic acid. The AC can hybridize to its target sequence and downregulate expression of the target protein. The AC can hybridize to its target sequence to downregulate expression of one or more target protein isomers. The AC can hybridize to its target sequence to upregulate expression of the target protein. The AC can hybridize to its target sequence to increase expression
10 of one or more target protein isomers.

[0355] The efficacy of the ACs may be assessed by evaluating the antisense activity effected by their administration. As used herein, the term "antisense activity" refers to any detectable and/or measurable activity attributable to the hybridization of an antisense compound to its target nucleic acid. Such detection and or measuring may be direct or indirect. In embodiments, antisense activity
15 is assessed by detecting and or measuring the amount of target protein. Antisense activity can be assessed by detecting and/or measuring the amount of target nucleic acids.

Antisense compound design

[0356] Design of ACs according to the present disclosure will depend upon the sequence being targeted. Targeting an AC to a particular target nucleic acid molecule can be a multistep process.
20 The process usually begins with the identification of a target nucleic acid whose expression is to be modulated. As used herein, the terms "target nucleic acid" and "nucleic acid encoding a target gene" encompass DNA encoding a selected target gene, RNA (including pre-mRNA and mRNA) transcribed from such DNA, and also cDNA derived from such RNA.

[0357] One of skill in the art will be able to design, synthesize, and screen antisense compounds
25 of different nucleobase sequences to identify a sequence that results in antisense activity. For example, one may design an antisense compound that inhibits expression of a target protein. Methods for designing, synthesizing and screening antisense compounds for antisense activity against a preselected target nucleic acid can be found, for example in "Antisense Drug Technology, Principles, Strategies, and Applications" Edited by Stanley T. Crooke, CRC Press, Boca Raton,
30 Florida, which is incorporated by reference in its entirety for any purpose.

[0358] Antisense compounds are provided that include from about 8 to about 30 linked nucleosides. The antisense compounds can include modified nucleosides, modified internucleoside linkages and/or conjugate groups.

5 [0359] The antisense compound can be a “tricyclo-DNA (tc-DNA)”, which refers to a class of constrained DNA analogs in which each nucleotide is modified by the introduction of a cyclopropane ring to restrict conformational flexibility of the backbone and to enhance the backbone geometry of the torsion angle γ . Homobasic adenine- and thymine-containing tc-DNAs form extraordinarily stable A-T base pairs with complementary RNAs.

Nucleosides

10 [0360] The antisense compounds can include linked nucleosides. Some or all of the nucleosides can be modified nucleosides. One or more nucleosides can include a modified nucleobase. One or more nucleosides can include a modified sugar. Chemically modified nucleosides are routinely used for incorporation into antisense compounds to enhance one or more properties, such as nuclease resistance, pharmacokinetics or affinity for a target RNA. Non-limiting examples of
15 nucleosides are provided in Khvorova et al. Nature Biotechnology (2017) 35: 238-248, which is incorporated by reference herein in its entirety.

[0361] In general, a nucleobase is any group that contains one or more atom or groups of atoms capable of hydrogen bonding to a base of another nucleic acid. In addition to “unmodified” or “natural” nucleobases such as the purine nucleobases adenine (A) and guanine (G), and the
20 pyrimidine nucleobases thymine (T), cytosine (C) and uracil (U), many modified nucleobases or nucleobase mimetics known to those skilled in the art are amenable with the compounds described herein. The terms modified nucleobase and nucleobase mimetic can overlap but generally a modified nucleobase refers to a nucleobase that is fairly similar in structure to the parent nucleobase, such as for example a 7-deaza purine, a 5-methyl cytosine, or a G-clamp, whereas a
25 nucleobase mimetic would include more complicated structures, such as for example a tricyclic phenoxazine nucleobase mimetic. Methods for preparation of the above noted modified nucleobases are well known to those skilled in the art.

[0362] ACs can include one or more nucleosides having a modified sugar moiety. The furanosyl sugar ring of a natural nucleoside can be modified in a number of ways including, but not limited
30 to, addition of a substituent group, bridging of two non-geminal ring atoms to form a bicyclic nucleic acid (BNA) and substitution of an atom or group such as -S-, -N(R)- or -C(R1)(R2) for the

ring oxygen at the 4'-position. Modified sugar moieties are well known and can be used to alter, typically increase, the affinity of the antisense compound for its target and/or increase nuclease resistance. A representative list of modified sugars includes but is not limited to non-bicyclic substituted sugars, especially non-bicyclic 2'-substituted sugars having a 2'-F, 2'-OCH₃ or a 2'-O(CH₂)₂-OCH₃ substituent group; and 4'-thio modified sugars. Sugars can also be replaced with sugar mimetic groups among others. Methods for the preparations of modified sugars are well known to those skilled in the art. Some representative patents and publications that teach the preparation of such modified sugars include, but are not limited to, U.S. Patents: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; 5,700,920; and 6,600,032; and WO 2005/121371 .

[0363] Nucleosides can include bicyclic modified sugars (BNA's), including LNA (4'-(CH₂)-O-2' bridge), 2'-thio-LNA (4'-(CH₂)-S-2' bridge), 2'-amino-LNA (4'-(CH₂)-NR-2' bridge), ENA (4'-(CH₂)₂-O-2' bridge), 4'-(CH₂)₃-2' bridged BNA, 4'-(CH₂CH(CH₃))-2' bridged BNA" cEt (4'-(CH(CH₃))-O-2' bridge), and cMOE BNAs (4'-(CH(CH₂OCH₃))-O-2' bridge). Certain such BNA's have been prepared and disclosed in the patent literature as well as in scientific literature (See, e.g., Srivastava, et al. J. Am. Chem. Soc. 2007, ACS Advanced online publication, 10.1021/ja071106y, Albaek et al. J. Org. Chem., 2006, 71, 7731 -7740, Fluiter, et al. Chembiochem 2005, 6, 1104-1109, Singh et al., Chem. Commun., 1998, 4, 455-456; Koshkin et al., Tetrahedron, 1998, 54, 3607-3630; Wahlestedt et al., Proc. Natl. Acad. Sci. U. S. A., 2000, 97, 5633-5638; Kumar et al., Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222; WO 94/14226; WO 2005/021570; Singh et al., J. Org. Chem., 1998, 63, 10035-10039, WO 2007/090071; Examples of issued US patents and published applications that disclose BNAs include, for example, U.S. Patent Nos. 7,053,207; 6,268,490; 6,770,748; 6,794,499; 7,034,133; and 6,525,191; and U.S. Pre-Grant Publication Nos. 2004-0171570; 2004-0219565; 2004-0014959; 2003-0207841; 2004-0143114; and 20030082807 .

[0364] Also provided herein are "Locked Nucleic Acids" (LNAs) in which the 2'-hydroxyl group of the ribosyl sugar ring is linked to the 4' carbon atom of the sugar ring thereby forming a 2'-C,4'-C-oxymethylene linkage to form the bicyclic sugar moiety (reviewed in Elayadi et al., Curr. Opin. Opin. Invers. Drugs, 2001, 2, 558-561; Braasch et al., Chem. Biol., 2001, 8 1-7; and Orum et al., Curr. Opin. Mol. Ther., 2001, 3, 239-243; see also U.S. Patents: 6,268,490 and 6,670,461).

The linkage can be a methylene (-CH₂-) group bridging the 2' oxygen atom and the 4' carbon atom, for which the term LNA is used for the bicyclic moiety; in the case of an ethylene group in this position, the term ENATM is used (Singh et al., Chem. Commun., 1998, 4, 455-456; ENATM: Morita et al., Bioorganic Medicinal Chemistry, 2003, 11, 2211-2226). LNA and other bicyclic sugar
5 analogs display very high duplex thermal stabilities with complementary DNA and RNA (T_m = +3 to +10° C), stability towards 3'-exonucleolytic degradation and good solubility properties. Potent and nontoxic antisense oligonucleotides containing LNAs have been described (Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A., 2000, 97, 5633-5638).

[0365] An isomer of LNA that has also been studied is alpha-L-LNA which has been shown to
10 have superior stability against a 3'-exonuclease. The alpha-L-LNA's were incorporated into antisense gapmers and chimeras that showed potent antisense activity (Frieden et al., Nucleic Acids Research, 2003, 21, 6365-6372).

[0366] The synthesis and preparation of the LNA monomers adenine, cytosine, guanine, 5-methyl-
cytosine, thymine and uracil, along with their oligomerization, and nucleic acid recognition
15 properties have been described (Koshkin et al., Tetrahedron, 1998, 54, 3607-3630). LNAs and preparation thereof are also described in WO 98/39352 and WO 99/14226.

[0367] Analogs of LNA, phosphorothioate-LNA and 2'-thio-LNAs, have also been prepared (Kumar et al.,
Bioorg. Med. Chem. Lett., 1998, 8, 2219-2222). Preparation of locked nucleoside analogs containing
oligodeoxyribonucleotide duplexes as substrates for nucleic acid polymerases has also been described (
20 Wengel et al., WO 99/14226). Furthermore, synthesis of 2'-amino-LNA, a novel conformationally restricted high-affinity oligonucleotide analog has been described in the art (Singh et al., J. Org. Chem., 1998, 63, 10035-10039). In addition, 2'-Amino- and 2'-methylamino-LNA's have been prepared and the thermal stability of their duplexes with complementary RNA and DNA strands has been previously reported.

Internucleoside Linkages

[0368] Described herein are internucleoside linking groups that link the nucleosides or otherwise
25 modified monomer units together thereby forming an antisense compound. The two main classes of internucleoside linking groups are defined by the presence or absence of a phosphorus atom. Representative phosphorus containing internucleoside linkages include, but are not limited to, phosphodiester, phosphotriester, methylphosphonates, phosphoramidate, and
30 phosphorothioates. Representative non-phosphorus containing internucleoside linking groups include, but are not limited to, methylenemethylimino (-CH₂-N(CH₃)-O-CH₂-), thiodiester (-O-C(O)-S-), thionocarbamate (-O-C(O)(NH)-S-); siloxane (-O-Si(H)₂-O-); and N,N'-

dimethylhydrazine (-CH₂-N(CH₃)-N(CH₃)-). Antisense compounds having non-phosphorus internucleoside linking groups are referred to as oligonucleosides. Modified internucleoside linkages, compared to natural phosphodiester linkages, can be used to alter, typically increase, nuclease resistance of the antisense compound. Internucleoside linkages having a chiral atom can be prepared racemic, chiral, or as a mixture. Representative chiral internucleoside linkages include, but are not limited to, alkylphosphonates and phosphorothioates. Methods of preparation of phosphorous-containing and non-phosphorous-containing linkages are well known to those skilled in the art.

[0369] A phosphate group can be linked to the 2', 3' or 5' hydroxyl moiety of the sugar. In forming oligonucleotides, the phosphate groups covalently link adjacent nucleosides to one another to form a linear polymeric compound. Within oligonucleotides, the phosphate groups are commonly referred to as forming the internucleoside backbone of the oligonucleotide. The normal linkage or backbone of RNA and DNA is a 3' to 5' phosphodiester linkage.

Conjugate Groups

[0370] Cargo can be modified by covalent attachment of one or more conjugate groups. In general, conjugate groups modify one or more properties of the cargo including but not limited to pharmacodynamic, pharmacokinetic, binding, absorption, cellular distribution, cellular uptake, charge and clearance. Conjugate groups are routinely used in the chemical arts and are linked directly or via an optional linking moiety or linking group to a parent compound. Conjugate groups include without limitation, intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, thioethers, polyethers, cholesterol, thiocholesterols, cholic acid moieties, folate, lipids, phospholipids, biotin, phenazine, phenanthridine, anthraquinone, adamantane, acridine, fluoresceins, rhodamines, coumarins and dyes. The conjugate group can include polyethylene glycol (PEG). PEG can be conjugated to either the cargo or the cCPP. The cargo can include a peptide, oligonucleotide or small molecule.

[0371] Conjugate groups can include lipid moieties such as a cholesterol moiety (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553); cholic acid (Manoharan et al., Bioorg. Med. Chem. Lett., 1994, 4, 1053); a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., Ann. N.Y. Acad. Sci., 1992, 660, 306; Manoharan et al., Bioorg. Med. Chem. Lett., 1993, 3, 2765); a thiocholesterol (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533); an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., EMBO J., 1991, 10, 111; Kabanov et al., FEBS Lett.,

1990, 259, 327; Svinarchuk et al., *Biochimie*, 1993, 75, 49); a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethylammonium-1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651; Shea et al., *Nucl. Acids Res.*, 1990, 18, 3777); a polyamine or a polyethylene glycol chain (Manoharan et al., *Nucleosides & Nucleotides*, 1995, 14, 969); adamantane acetic acid (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651); a palmityl moiety (Mishra et al., *Biochim. Biophys. Acta*, 1995, 1264, 229); or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke et al., *J. Pharmacol. Exp. Ther.*, 1996, 277, 923).

[0372] Linking groups or bifunctional linking moieties such as those known in the art are amenable to the compounds provided herein. Linking groups are useful for attachment of chemical functional groups, conjugate groups, reporter groups and other groups to selective sites in a parent compound such as for example an AC. In general a bifunctional linking moiety includes a hydrocarbyl moiety having two functional groups. One of the functional groups is selected to bind to a parent molecule or compound of interest and the other is selected to bind essentially any selected group such as chemical functional group or a conjugate group. Any of the linkers described here may be used. the linker can include a chain structure or an oligomer of repeating units such as ethylene glycol or amino acid units. Examples of functional groups that are routinely used in a bifunctional linking moiety include, but are not limited to, electrophiles for reacting with nucleophilic groups and nucleophiles for reacting with electrophilic groups. Bifunctional linking moieties can include amino, hydroxyl, carboxylic acid, thiol, unsaturations (e.g., double or triple bonds), and the like. Some nonlimiting examples of bifunctional linking moieties include 8-amino-3,6-dioxaoctanoic acid (ADO), succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC) and 6-aminohexanoic acid (AHX or AHA). Other linking groups include, but are not limited to, substituted C₁-C₁₀ alkyl, substituted or unsubstituted C₂-C₁₀ alkenyl or substituted or unsubstituted C₂-C₁₀ alkynyl, wherein a nonlimiting list of substituent groups includes hydroxyl, amino, alkoxy, carboxy, benzyl, phenyl, nitro, thiol, thioalkoxy, halogen, alkyl, aryl, alkenyl and alkynyl.

[0373] The AC may be from about 5 to about 50 nucleotides in length. The AC may be from about 5 to about 10 nucleotides in length. The AC may be from about 10 to about 15 nucleotides in length. The AC may be from about 15 to about 20 nucleotides in length. The AC may be from about 20 to about 25 nucleotides in length. The AC may be from about 25 to about 30 nucleotides

in length. The AC may be from about 30 to about 35 nucleotides in length. The AC may be from about 35 to about 40 nucleotides in length. The AC may be from about 40 to about 45 nucleotides in length. The AC may be from about 45 to about 50 nucleotides in length.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) Gene-Editing

5 Machinery

[0374] The compounds can include one or more cCPP (or cCPP) conjugated to CRISPR gene-editing machinery. As used herein, "CRISPR gene-editing machinery" refers to protein, nucleic acids, or combinations thereof, which may be used to edit a genome. Non-limiting examples of gene-editing machinery include gRNAs, nucleases, nuclease inhibitors, and combinations and
10 complexes thereof. The following patent documents describe CRISPR gene-editing machinery: U.S. Pat. No. 8,697,359, U.S. Pat. No. 8,771,945, U.S. Pat. No. 8,795,965, U.S. Pat. No. 8,865,406, U.S. Pat. No. 8,871,445, U.S. Pat. No. 8,889,356, U.S. Pat. No. 8,895,308, U.S. Pat. No. 8,906,616, U.S. Pat. No. 8,932,814, U.S. Pat. No. 8,945,839, U.S. Pat. No. 8,993,233, U.S. Pat. No. 8,999,641, U.S. Pat. App. No. 14/704,551, and U.S. Pat. App. No. 13/842,859. Each of the aforementioned
15 patent documents is incorporated by reference herein in its entirety.

[0375] A linker can conjugate the cCPP to the CRISPR gene-editing machinery. Any linker described in this disclosure or that is known to a person of skill in the art may be utilized.

gRNA

[0376] The compound can include the cCPP conjugated to a gRNA. A gRNA targets a genomic
20 loci in a prokaryotic or eukaryotic cell.

[0377] The gRNA can be a single-molecule guide RNA (sgRNA). A sgRNA includes a spacer sequence and a scaffold sequence. A spacer sequence is a short nucleic acid sequence used to target a nuclease (e.g., a Cas9 nuclease) to a specific nucleotide region of interest (e.g., a genomic DNA sequence to be cleaved). The spacer may be about 17-24 base pairs in length, such as about 20
25 base pairs in length. The spacer may be about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, or about 30 base pairs in length. The spacer may be at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 base pairs in length. The spacer may be about 15, about
30 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about

26, about 27, about 28, about 29, or about 30 base pairs in length. The spacer sequence can have between about 40% to about 80% GC content.

[0378] The spacer can target a site that immediately precedes a 5' protospacer adjacent motif (PAM). The PAM sequence may be selected based on the desired nuclease. For example, the PAM sequence may be any one of the PAM sequences shown in the table below, wherein N refers to any nucleic acid, R refers to A or G, Y refers to C or T, W refers to A or T, and V refers to A or C or G.

Table 8

PAM sequence (5' to 3')	Nuclease	Isolated from
NGG	SpCas9	Streptococcus pyogenes
NGRRT or NGRRN	SaCas9	Staphylococcus aureus
NNNNGATT	NmeCas9	Neisseria meningitidis
NNNNRYAC	CjCas9	Campylobacter jejuni
NNAGAAW	StCas9	Streptococcus thermophiles
TTTV	LbCpf1	Lachnospiraceae bacterium
TTTV	AsCpf1	Acidaminococcus sp.

[0379] A spacer may target a sequence of a mammalian gene, such as a human gene. The spacer may target a mutant gene. The spacer may target a coding sequence.

[0380] The scaffold sequence is the sequence within the sgRNA that is responsible for nuclease (e.g., Cas9) binding. The scaffold sequence does not include the spacer/targeting sequence. In embodiments, the scaffold may be about 1 to about 10, about 10 to about 20, about 20 to about 30, about 30 to about 40, about 40 to about 50, about 50 to about 60, about 60 to about 70, about 70 to about 80, about 80 to about 90, about 90 to about 100, about 100 to about 110, about 110 to about 120, or about 120 to about 130 nucleotides in length. The scaffold may be about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 51, about 52, about 53, about 54, about 55, about 56, about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70, about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about

92, about 93, about 94, about 95, about 96, about 97, about 98, about 99, about 100, about 101, about 102, about 103, about 104, about 105, about 106, about 107, about 108, about 109, about 110, about 111, about 112, about 113, about 114, about 115, about 116, about 117, about 118, about 119, about 120, about 121, about 122, about 123, about 124, or about 125 nucleotides in length. The scaffold may be at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, or at least 125 nucleotides in length.

[0381] The gRNA can be a dual-molecule guide RNA, e.g, crRNA and tracrRNA. The gRNA may further include a polyA tail.

10 [0382] A compound includes a cCPP conjugated to a nucleic acid that includes a gRNA. The nucleic acid can include about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, or about 20 gRNAs. The gRNA can recognize the same target. The gRNA can recognize different targets. The nucleic acid that includes a gRNA includes a sequence encoding a promoter, 15 wherein the promoter drives expression of the gRNA.

Nuclease

[0383] The compounds can include a cyclic cell penetrating peptide (cCPP) conjugated to a nuclease. The nuclease can be a Type II, Type V-A, Type V-B, Type VC, Type V-U, or Type VI-B nuclease. The nuclease can be a transcription, activator-like effector nuclease (TALEN), a 20 meganuclease, or a zinc-finger nuclease. The nuclease can be a Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B like, Cas13a (C2c2), Cas13b, or Cas14 nuclease. The nuclease can be a Cas9 nuclease or a Cpf1 nuclease.

[0384] The nuclease can be a modified form or variant of a Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B like, Cas13a (C2c2), Cas13b, or Cas14 nuclease. The nuclease can be a modified form or 25 variant of a TAL nuclease, a meganuclease, or a zinc-finger nuclease. A “modified” or “variant” nuclease is one that is, for example, truncated, fused to another protein (such as another nuclease), catalytically inactivated, etc. The nuclease may have at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, at least about 99%, or about 100% sequence identity to a naturally occurring Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B like, 30 Cas13a (C2c2), Cas13b, Cas14 nuclease, or a TALEN, meganuclease, or zinc-finger nuclease. The nuclease can be a Cas9 nuclease derived from *S. pyogenes* (SpCas9). The nuclease can have at

least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to a Cas9 nuclease derived from *S. pyogenes* (SpCas9). The nuclease can be a Cas9 derived from *S. aureus* (SaCas9). The nuclease can have at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to a Cas9 derived from *S. aureus* (SaCas9). Cpf1 can be a Cpf1 enzyme from *Acidaminococcus* (species BV3L6, UniProt Accession No. U2UMQ6). The nuclease can have at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to a Cpf1 enzyme from *Acidaminococcus* (species BV3L6, UniProt Accession No. U2UMQ6).

10 [0385] Cpf1 can be a Cpf1 enzyme from Lachnospiraceae (species ND2006, UniProt Accession No. A0A182DWE3). The nuclease can have at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to a Cpf1 enzyme from Lachnospiraceae. The sequence encoding the nuclease can be codon optimized for expression in mammalian cells. The sequence encoding the nuclease can be codon optimized for expression in human cells or mouse cells.

[0386] The compound can include a cCPP conjugated to a nuclease. The nuclease can be a soluble protein.

[0387] The compound can include a cCPP conjugated to a nucleic acid encoding a nuclease. The nucleic acid encoding a nuclease can include a sequence encoding a promoter, wherein the promoter drives expression of the nuclease.

gRNA and Nuclease Combinations

[0388] The compounds can include one or more cCPP conjugated to a gRNA and a nuclease. One or more cCPP can be conjugated to a nucleic acid encoding a gRNA and/or a nuclease. The nucleic acid encoding a nuclease and a gRNA can include a sequence encoding a promoter, wherein the promoter drives expression of the nuclease and the gRNA. The nucleic acid encoding a nuclease and a gRNA can include two promoters, wherein a first promoter controls expression of the nuclease and a second promoter controls expression of the gRNA. The nucleic acid encoding a gRNA and a nuclease can encode from about 1 to about 20 gRNAs, or from about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, or about 19, and up to about 20 gRNAs. The gRNAs can recognize different targets. The gRNAs can recognize the same target.

[0389] The compounds can include a cyclic cell penetrating peptide (or cCPP) conjugated to a ribonucleoprotein (RNP) that includes a gRNA and a nuclease.

[0390] A composition that includes: (a) a cCPP conjugated to a gRNA and (b) a nuclease can be delivered to a cell. A composition that includes: (a) a cCPP conjugated to a nuclease and (b) an gRNA can be delivered to a cell.

[0391] A composition that includes: (a) a first cCPP conjugated to a gRNA and (b) a second cCPP conjugated to a nuclease can be delivered to a cell. The first cCPP and second cCPP can be the same. The first cCPP and second cCPP can be different.

Genetic Element of Interest

[0392] The compounds can include a cyclic cell penetrating peptide (cCPP) conjugated to a genetic element of interest. A genetic element of interest can replace a genomic DNA sequence cleaved by a nuclease. Non-limiting examples of genetic elements of interest include genes, a single nucleotide polymorphism, promoter, or terminators.

Nuclease Inhibitors

[0393] The compound can include a cyclic cell penetrating peptide (cCPP) conjugated to an inhibitor of a nuclease (e.g. a Cas9 inhibitor). A limitation of gene editing is potential off-target editing. The delivery of a nuclease inhibitor may limit off-target editing. The nuclease inhibitor can be a polypeptide, polynucleotide, or small molecule. Nuclease inhibitors are described in U.S. Publication No. 2020/087354, International Publication No. 2018/085288, U.S. Publication No. 2018/0382741, International Publication No. 2019/089761, International Publication No. 2020/068304, International Publication No. 2020/041384, and International Publication No. 2019/076651, each of which is incorporated by reference herein in its entirety.

Therapeutic polypeptides

Antibodies

[0394] The therapeutic moiety can include an antibody or an antigen-binding fragment. Antibodies and antigen-binding fragments can be derived from any suitable source, including human, mouse, camelid (e.g., camel, alpaca, llama), rat, ungulates, or non-human primates (e.g., monkey, rhesus macaque).

[0395] The term “antibody” refers to an immunoglobulin (Ig) molecule capable of binding to a specific target, such as a carbohydrate, polynucleotide, lipid, or polypeptide, through at least one

epitope recognition site located in the variable region of the Ig molecule. As used herein, the term encompasses intact polyclonal or monoclonal antibodies and antigen-binding fragments thereof. A native immunoglobulin molecule generally includes two heavy chain polypeptides and two light chain polypeptides. Each of the heavy chain polypeptides associate with a light chain polypeptide by virtue of interchain disulfide bonds between the heavy and light chain polypeptides to form two heterodimeric proteins or polypeptides (i.e., a protein comprised of two heterologous polypeptide chains). The two heterodimeric proteins then associate by virtue of additional interchain disulfide bonds between the heavy chain polypeptides to form an immunoglobulin protein or polypeptide.

[0396] The term “antigen-binding fragment” as used herein refers to a polypeptide fragment that contains at least one complementarity-determining region (CDR) of an immunoglobulin heavy and/or light chain that binds to at least one epitope of the antigen of interest. An antigen-binding fragment may comprise 1, 2, or 3 CDRs of a variable heavy chain (VH) sequence from an antibody that specifically binds to a target molecule. An antigen-binding fragment may comprise 1, 2, or 3 CDRs of a variable light chain (VL) sequence from an antibody that specifically binds to a target molecule. An antigen-binding fragment may comprise 1, 2, 3, 4, 5, or all 6 CDRs of a variable heavy chain (VH) and variable light chain (VL) sequence from antibodies that specifically bind to a target molecule. Antigen-binding fragments include proteins that comprise a portion of a full length antibody, generally the antigen binding or variable region thereof, such as Fab, F(ab')₂, Fab', Fv fragments, minibodies, diabodies, single domain antibody (dAb), single-chain variable fragments (scFv), nanobodies, multispecific antibodies formed from antibody fragments, and any other modified configuration of the immunoglobulin molecule that can comprise an antigen-binding site or fragment of the required specificity.

[0397] The term “F(ab)” refers to two of the protein fragments resulting from proteolytic cleavage of IgG molecules by the enzyme papain. Each F(ab) can comprise a covalent heterodimer of the VH chain and VL chain and includes an intact antigen-binding site. Each F(ab) can be a monovalent antigen-binding fragment. The term “Fab” refers to a fragment derived from F(ab')₂ and may contain a small portion of Fc. Each Fab' fragment can be a monovalent antigen-binding fragment.

[0398] The term “F(ab')₂” refers to a protein fragment of IgG generated by proteolytic cleavage by the enzyme pepsin. Each F(ab')₂ fragment can comprise two F(ab') fragments and can be therefore a bivalent antigen-binding fragment.

[0399] An “Fv fragment” refers to a non-covalent VH::VL heterodimer which includes an antigen-binding site that retains much of the antigen recognition and binding capabilities of the native antibody molecule, but lacks the CH1 and CL domains contained within a Fab. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

[0400] Bispecific Antibodies (BsAbs) are antibodies that can simultaneously bind two separate and unique antigens (or different epitopes of the same antigen). The therapeutic moiety can include a bispecific antibody that can simultaneously bind to two different targets of interest. The BsAbs may redirect cytotoxic immune effector cells for enhanced killing of tumor cells by antibody-dependent cell-mediated cytotoxicity (ADCC) and other cytotoxic mechanisms mediated by the effector cells.

[0401] Recombinant antibody engineering has allowed for the creation of recombinant bispecific antibody fragments comprising the variable heavy (VH) and light (VL) domains of the parental monoclonal antibodies (mabs). Non-limiting examples include scFv (single-chain variable fragment), BsDb (bispecific diabody), scBsDb (single-chain bispecific diabody), scBsTaFv (single-chain bispecific tandem variable domain), DNL-(Fab)3 (dock-and-lock trivalent Fab), sdAb (single-domain antibody), and BssdAb (bispecific single-domain antibody).

[0402] BsAbs with an Fc region are useful for carrying out Fc mediated effector functions such as ADCC and CDC. They have the half-life of normal IgG. On the other hand, BsAbs without the Fc region (bispecific fragments) rely solely on their antigen-binding capacity for carrying out therapeutic activity. Due to their smaller size, these fragments have better solid-tumor penetration rates. BsAb fragments do not require glycosylation, and they may be produced in bacterial cells. The size, valency, flexibility and half-life of BsAbs to suit the application.

[0403] Using recombinant DNA technology, bispecific IgG antibodies can be assembled from two different heavy and light chains expressed in the same cell line. Random assembly of the different chains results in the formation of nonfunctional molecules and undesirable HC homodimers. To address this problem, a second binding moiety (e.g., single chain variable fragment) may be fused to the N or C terminus of the H or L chain resulting in tetravalent BsAbs containing two binding sites for each antigen. Additional methods to address the LC-HC mispairing and HC homodimerization follow.

[0404] *Knobs-into-holes BsAb IgG*. H chain heterodimerization is forced by introducing different mutations into the two CH3 domains resulting in asymmetric antibodies. Specifically a “knob” mutation is made into one HC and a “hole” mutation is created in the other HC to promote heterodimerization.

- 5 [0405] *Ig-scFv fusion*. The direct addition of a new antigen-binding moiety to full length IgG results in fusion proteins with tetravalency. Examples include IgG C-terminal scFv fusion and IgG N-terminal scFv fusion.

[0406] *Diabody-Fc fusion*. This involves replacing the Fab fragment of an IgG with a bispecific diabody (derivative of the scFv).

- 10 [0407] *Dual-Variable-Domain-IgG (DVD-IgG)*. VL and VH domains of IgG with one specificity were fused respectively to the N-terminal of VL and VH of an IgG of different specificity via a linker sequence to form a DVD-IgG.

[0408] The term “diabody” refers to a bispecific antibody in which VH and VL domains are expressed in a single polypeptide chain using a linker that is too short to allow for pairing between
15 the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen-binding sites (see, *e.g.*, Holliger et al., Proc. Natl. Acad. Sci. USA 90:6444-48 (1993) and Poljak et al., Structure 2:1121-23 (1994)). Diabodies may be designed to bind to two distinct antigens and are bi-specific antigen binding constructs.

[0409] The term “nanobody” or a “single domain antibody” refers to an antigen-binding fragment
20 of a single monomeric variable antibody domain comprising one variable domain (VH) of a heavy-chain antibody. They possess several advantages over traditional monoclonal antibodies (mAbs), including smaller size (15 kD), stability in the reducing intracellular environment, and ease of production in bacterial systems (Schumacher *et al.*, (2018) Nanobodies: Chemical Functionalization Strategies and Intracellular Applications. *Angew. Chem. Int. Ed.* 57, 2314;
25 Siontorou, (2013) Nanobodies as novel agents for disease diagnosis and therapy. *International Journal of Nanomedicine*, 8, 4215-27). These features render nanobodies amendable to genetic and chemical modifications (Schumacher *et al.*, (2018) Nanobodies: Chemical Functionalization Strategies and Intracellular Applications. *Angew. Chem. Int. Ed.* 57, 2314), facilitating their application as research tools and therapeutic agents (Bannas *et al.*, (2017) Nanobodies and
30 nanobody-based human heavy chain antibodies as antitumor therapeutics. *Frontiers in Immunology*, 8, 1603). Over the past decade, nanobodies have been used for protein

immobilization (Rothbauer *et al.*, (2008) A Versatile Nanotrap for Biochemical and Functional Studies with Fluorescent Fusion Proteins. *Mol. Cell. Proteomics*, 7, 282-289), imaging (Traenkle *et al.*, (2015) Monitoring Interactions and Dynamics of Endogenous Beta-catenin With Intracellular Nanobodies in Living Cells. *Mol. Cell. Proteomics*, 14, 707-723), detection of
5 protein-protein interactions (Herce *et al.*, (2013) Visualization and targeted disruption of protein interactions in living cells. *Nat. Commun*, 4, 2660; Massa *et al.*, (2014) Site-Specific Labeling of Cysteine-Tagged Camelid Single-Domain Antibody-Fragments for Use in Molecular Imaging. *Bioconjugate Chem*, 25, 979-988), and as macromolecular inhibitors (Truttmann *et al.*, (2015) HypE-specific Nanobodies as Tools to Modulate HypE-mediated Target AMPylation. *J. Biol.*
10 *Chem.* 290, 9087–9100).

[0410] The therapeutic moiety can be an antigen-binding fragment that binds to a target of interest. The antigen-binding fragment that binds to the target of interest may include 1, 2, or 3, CDRs of a variable heavy chain (VH) sequence from an antibody that specifically binds to the target of interest. The antigen-binding fragment that binds to the target of interest may include 1, 2, or 3
15 CDRs of a variable light chain (VL) sequence from an antibody that specifically binds to the target of interest. The antigen-binding fragment that binds to the target of interest may include 1, 2, 3, 4, 5, or all 6 CDRs of a variable heavy chain (VH) and/or a variable light chain (VL) sequence from an antibody that specifically binds to the target of interest. The antigen-binding fragment that binds to the target may be a portion of a full-length antibody, such as Fab, F(ab')₂, Fab', Fv
20 fragments, minibodies, diabodies, single domain antibody (dAb), single-chain variable fragments (scFv), nanobodies, multispecific antibodies formed from antibody fragments, or any other modified configuration of the immunoglobulin molecule that includes an antigen-binding site or fragment of the required specificity.

[0411] The therapeutic moiety can include a bispecific antibody. Bispecific Antibodies (BsAbs)
25 are antibodies that can simultaneously bind two separate and unique antigens (or different epitopes of the same antigen).

[0412] The therapeutic moiety can include a “diabody”.

[0413] The therapeutic moiety can include a nanobody or a single domain antibody (which can also be referred to herein as sdAbs or VHH).

30 [0414] The therapeutic moiety can include a “minibody.” Minibodies (Mb) include a CH3 domain fused or linked to an antigen-binding fragment (e.g., a CH3 domain fused or linked to an scFv, a

domain antibody, etc.). The term “Mb” can signify a CH3 single domain. A CH3 domain can signify a minibody. (S. Hu et al., *Cancer Res.*, 56, 3055-3061, 1996). See e.g., Ward, E. S. et al., *Nature* 341, 544-546 (1989); Bird et al., *Science*, 242, 423-426, 1988; Huston et al., *PNAS USA*, 85, 5879-5883, 1988); PCT/US92/09965; WO94/13804; P. Holliger et al., *Proc. Natl. Acad. Sci. USA* 90 6444-6448, 1993; Y. Reiter et al., *Nature Biotech*, 14, 1239-1245, 1996; S. Hu et al., *Cancer Res.*, 56, 3055-3061, 1996.

[0415] The therapeutic moiety can include a “monobody”. The term “monobody” refers to a synthetic binding protein constructed using a fibronectin type III domain (FN3) as a molecular scaffold.

10 [0416] The therapeutic moiety can be an antibody mimetic. Antibody mimetics are compounds that, like antibodies, can specifically bind antigens, but that are not structurally related to antibodies. They are usually artificial peptides or proteins with a molar mass of about 3 to 20 kD (compared to the molar mass of antibodies at ~150 kDa.). Examples of antibody mimetics include Affibody molecules (constructed on a scaffold of the domain of Protein A, See, Nygren (June
15 2008). *FEBS J.* 275 (11): 2668–76), Affilins (constructed on a scaffold of gamma-B crystalline or ubiquitin, See Ebersbach H et al. (September 2007). *J. Mol. Biol.* 372 (1): 172–85), Affimers (constructed on a Crystatin scaffold, See Johnson A et al., (Aug 7, 2012). *Anal. Chem.* 84 (15): 6553–60), Affitins (constructed on a Sac7d from *S. acidocaldarius* scaffold, See Krehenbrink M et al., (November 2008). *J. Mol. Biol.* 383 (5): 1058– 68), Alphabodies (constructed on a triple
20 helix coiled coil scaffold, See Desmet, J et al., (5 Feb 2014). *Nature Communications.* 5: 5237), Anticalins (constructs on scaffold of lipocalins, See Skerra A (June 2008). *FEBS J.* 275 (11): 2677–83), Avimers (constructed on scaffolds of various membrane receptors, See Silverman J et al. (December 2005). *Nat. Biotechnol.* 23 (12): 1556–61), DARPins (constructed on scaffolds of ankyrin repeat motifs, See Stumpp et al., (August 2008). *Drug Discov. Today.* 13 (15–16): 695–
25 701), Fynomers (constructed on a scaffold of the SH3 domain of Fyn, See Grabulovski et al., (2007). *J Biol Chem.* 282 (5): 3196–3204), Kunitz domain peptides (constructed on scaffolds of the Kunitz domains of various protease inhibitors, See Nixon et al (March 2006). *Curr Opin Drug Discov Dev.* 9 (2): 261–8), and Monobodies (constructed on scaffolds of type III domain of fibronectin, See Koide et al (2007). *Methods Mol. Biol.* 352: 95– 109).

[0417] The therapeutic moiety can include “designed ankryin repeats” or “DARPin”. DARPins are derived from natural ankyrin proteins comprised of at least three repeat motifs proteins, and usually comprise of four or five repeats.

[0418] The therapeutic moiety can include “dualvariable- domain-IgG” or “DVD-IgG”. DVD-IgGs are generated from two parental monoclonal antibodies by fusing VL and VH domains of IgG with one specificity to the N-terminal of VL and VH of an IgG of different specificity, respectively, via a linker sequence.

[0419] The therapeutic moiety can include a F(ab) fragment.

[0420] The therapeutic moiety can include a F(ab')₂ fragment.

[0421] The therapeutic moiety can include an Fv fragment.

[0422] The antigen-binding fragment can include a “single chain variable fragment” or “scFv”. An scFv refers to a fusion protein of the variabl regions of the heavy (VH) and light chains (VL) of immunoglobulins, connected with a short linker peptide of ten to about 25 amino acids. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. The linker can connect the N-terminus of the VH with the C-terminus of the VL, or vice versa. A number of methods have been described to discern chemical structures for converting the natur-ally aggregated - but chemi-ally separated - light and heavy polypeptide chains from an antibody V region into an scFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

[0423] The antigen binding constructcancomprise two or more antigen-binding moieties. The antigen binding constructs can bind to two separate and unique antigens or to different epitopes of the same antigen. Knobs-into-holes BsAb IgG. H chain heterodimerization is forced by introducing different mutations into the two CH3 domains resulting in asymmetric antibodies. Specifically, a “knob” mutation is made into one HC and a “hole” mutation is created in the other HC to promote heterodimerization.

Peptide inhibitors

[0424] The therapeutic moiety can include a peptide. the peptide can act as an agonist, increasing activity of a target protein. The peptide can act as an antagonist, decreasing activity of a target protein. The peptide can be configured to inhibit protein-protein interaction (PPI). Protein-protein interactions (PPIs) are important in many biochemical processes, including transcription of nucleic

acid and various post-translational modifications of translated proteins. PPIs can be experimentally determined by biophysical techniques such as X-ray crystallography, NMR spectroscopy, surface plasma resonance (SPR), bio-layer interferometry (BLI), isothermal titration calorimetry (ITC), radio-ligand binding, spectrophotometric assays and fluorescence spectroscopy. Peptides that inhibit protein-protein interaction can be referred to as peptide inhibitors.

[0425] The therapeutic moiety can include a peptide inhibitor. The peptide inhibitor can include from about 5 to about 100 amino acids, from about 5 to about 50 amino acids; from about 15 to about 30 amino acids; or from about 20 to about 40 amino acids. The peptide inhibitor can include one or more chemical modifications, for example, to reduce proteolytic degradation and/or to improve *in vivo* half-life. The peptide inhibitor can include one or more synthetic amino acids and/or a backbone modification. The peptide inhibitor can have an α -helical structure.

[0426] The peptide inhibitor can target the dimerization domain of a homodimeric or heterodimeric target protein of interest.

Small molecules

[0427] The therapeutic moiety can include a small molecule. The therapeutic moiety can include a small molecule kinase inhibitor. The therapeutic moiety can include a small molecule that inhibits a kinase that phosphorylates a target of interest. Inhibition of phosphorylation of target of interest can block nuclear translocation of the target of interest. The therapeutic moiety can include a small molecule inhibitor of MyD88.

Compositions

[0428] Compositions are provided that include the compounds described herein.

[0429] Pharmaceutically acceptable salts and/or prodrugs of the disclosed compounds are provided. Pharmaceutically acceptable salts include salts of the disclosed compounds that are prepared with acids or bases, depending on the particular substituents found on the compounds.

Under conditions where the compounds disclosed herein are sufficiently basic or acidic to form stable nontoxic acid or base salts, administration of the compounds as salts can be appropriate. Examples of pharmaceutically acceptable base addition salts include sodium, potassium, calcium, ammonium, or magnesium salt. Examples of physiologically acceptable acid addition salts include hydrochloric, hydrobromic, nitric, phosphoric, carbonic, sulfuric, and organic acids like acetic, propionic, benzoic, succinic, fumaric, mandelic, oxalic, citric, tartaric, malonic, ascorbic, alpha-

ketoglutaric, alpha-glycophosphoric, maleic, tosyl acid, methanesulfonic, and the like. Thus, disclosed herein are the hydrochloride, nitrate, phosphate, carbonate, bicarbonate, sulfate, acetate, propionate, benzoate, succinate, fumarate, mandelate, oxalate, citrate, tartarate, malonate, ascorbate, alpha-ketoglutarate, alpha-glycophosphate, maleate, tosylate, and mesylate salts.

5 Pharmaceutically acceptable salts of a compound can be obtained using standard procedures well known in the art, for example, by reacting a sufficiently basic compound such as an amine with a suitable acid affording a physiologically acceptable anion. Alkali metal (for example, sodium, potassium or lithium) or alkaline earth metal (for example calcium) salts of carboxylic acids can also be made.

10 **Mechanism of Oligonucleotide Therapeutics and Target Molecules**

[0430] Many types of oligonucleotides are capable of modulating gene transcription, translation and/or protein function in cells. Non-limiting examples of such oligonucleotides include, e.g. , small interfering RNA (siRNA), microRNA (miRNA), antisense oligonucleotides, ribozymes, plasmids, immune stimulating nucleic acids, antisense, antagomir, antimir, microRNA mimic, supermir, Ul adaptor, and aptamer. Additional examples include DNA-targeting, triplex-forming
15 oligonucleotide, strand-invading oligonucleotide, and synthetic guide strand for CRISPR/Cas, These nucleic acids act via a variety of mechanisms. See Smith and Zain, Annu Rev Pharmacol Toxicol. 2019, 59:605-630, incorporated by reference herein.

[0431] Splice-switching antisense oligonucleotides are short, synthetic, antisense, modified
20 nucleic acids that base-pair with a pre-mRNA and disrupt the normal splicing repertoire of the transcript by blocking the RNA–RNA base-pairing or protein–RNA binding interactions that occur between components of the splicing machinery and the pre-mRNA. Splicing of pre-mRNA is required for the proper expression of the vast majority of protein-coding genes, and thus, targeting the process offers a means to manipulate protein production from a gene. Splicing modulation is
25 particularly valuable in cases of disease caused by mutations that lead to disruption of normal splicing or when interfering with the normal splicing process of a gene transcript may be therapeutic. Such antisense oligonucleotides offer an effective and specific way to target and alter splicing in a therapeutic manner. See Havens and Hastings, Nucleic Acids Res. 2016 Aug 19;44(14):6549-6563, incorporated by reference herein.

30 [0432] In the case of siRNA or miRNA, these nucleic acids can down-regulate intracellular levels of specific proteins through a process termed RNA interference (RNAi). Following introduction

of siRNA or miRNA into the cell cytoplasm, these double-stranded RNA constructs can bind to a protein termed RISC. The sense strand of the siRNA or miRNA is displaced from the RISC complex providing a template within RISC that can recognize and bind mRNA with a complementary sequence to that of the bound siRNA or miRNA. Having bound the complementary mRNA the RISC complex cleaves the mRNA and releases the cleaved strands. RNAi can provide down-regulation of specific proteins by targeting specific destruction of the corresponding mRNA that encodes for protein synthesis.

[0433] The therapeutic applications of RNAi are extremely broad, since siRNA and miRNA constructs can be synthesized with any nucleotide sequence directed against a target protein. To date, siRNA constructs have shown the ability to specifically down-regulate target proteins in both in vitro and in vivo models, as well as in clinical studies.

[0434] Antisense oligonucleotides and ribozymes can also inhibit mRNA translation into protein. In the case of antisense constructs, these single stranded deoxynucleic acids have a complementary sequence to that of the target protein mRNA and can bind to the mRNA by Watson-Crick base pairing. This binding either prevents translation of the target mRNA and/or triggers RNase H degradation of the mRNA transcripts. Consequently, antisense oligonucleotides have tremendous potential for specificity of action (i.e., down-regulation of a specific disease-related protein). To date, these compounds have shown promise in several in vitro and in vivo models, including models of inflammatory disease, cancer, and HIV (reviewed in Agrawal, Trends in Biotech. 14:376-387 (1996)). Antisense can also affect cellular activity by hybridizing specifically with chromosomal DNA.

[0435] Immune-stimulating nucleic acids include deoxyribonucleic acids and ribonucleic acids. In the case of deoxyribonucleic acids, certain sequences or motifs have been shown to illicit immune stimulation in mammals. These sequences or motifs include the CpG motif, pyrimidine-rich sequences and palindromic sequences. It is believed that the CpG motif in deoxyribonucleic acids is specifically recognized by an endosomal receptor, tolllike receptor 9 (TLR-9), which then triggers both the innate and acquired immune stimulation pathway. Certain immune stimulating ribonucleic acid sequences have also been reported. It is believed that these RNA sequences trigger immune activation by binding to toll-like receptors 6 and 7 (TLR-6 and TLR-7). In addition, double-stranded RNA is also reported to be immune stimulating and is believed to activate via binding to TLR-3.

[0436] Non-limiting examples of mechanism and targets of antisense oligonucleotides (ASOs) to modulate gene transcription, translation and/or protein function are illustrated in Table 9A and 9B.

Table 9A. Mechanism of ASO Modulation and Target Molecules

Types	Location of target	Subcellular Location	Mechanism
mRNA	Intracellular	cytoplasm	inhibition of translation
Pre-mRNA	Intracellular	nucleus	alternative splicing
micro-RNA	Intracellular	cytoplasm and nucleus	miRNA inhibition or activation
long non-coding RNA	Intracellular	cytoplasm and nucleus	inhibition of lncRNA function
Telomerase RNA component	Intracellular	cytoplasm and nucleus	inhibition of telomerase
Protein	Extra- and intra-cellular	cytoplasm and nucleus	inhibition of protein target

5

Table 9B. Mechanism of ASO Modulation and Target Molecules

Mechanism	Target	Description	Examples of drugs
Regulation of pre-mRNA splicing	pre-mRNA	ASOs bind to pre-mRNA and alter the splicing by steric blocking, which result in disruption of the recognition by splicing factors	Nusinersen, Eteplirsen
Regulation of RNA translation by recruiting RNase H	pre-mRNA and mRNA	ASOs containing DNA bases bind to target RNA and induce the cleavage of RNA by RNase H	Mipomersen, Inotersen
Regulation of RNA translation by steric blocking	mRNA	ASOs and duplex RNA can both sterically block the translation machinery to inhibit RNA translation or enhance RNA translation by blocking aberrant sites that reduce RNA translation	
Regulation of RNA translation by RNAi	mRNA	siRNA and miRNA inhibit translation by RNA interference and induce the cleavage of target RNA	Patisiran, Inclisiran, Fitusiran, Givosiran

Regulation of protein activity by binding with target proteins	protein	Aptamers bind with target proteins as antagonists	Pegaptanib
--	---------	---	------------

[0437] Clustered regularly interspaced short palindromic repeats (CRISPR) and associated Cas proteins constitute the CRISPR-Cas system. CRISPR-Cas is a mechanism for gene-editing. The RNA-guided (e.g., gRNA) Cas9 endonuclease specifically targets and cleaves DNA in a sequence-dependent manner. The Cas9 endonuclease can be substituted with any nuclease of the disclosure. The gRNA targets a nuclease (e.g., a Cas9 nuclease) to a specific nucleotide region of interest (e.g., a genomic DNA sequence to be cleaved) and cleaves genomic DNA. Genomic DNA can then be replaced with a genetic element of interest.

Methods of Modulation of Tissue Distribution and/or Retention

[0438] Provided herein are compounds and methods for modulating tissue distribution and/or retention of a therapeutic agent in a subject. Tissue distribution relates to for example: increasing the concentration of a therapeutic agent in specific regions of a tissue, for example increasing the concentration in regions of the brain such as in the cerebellum, cortex, hippocampus, or olfactory bulb relative to the unconjugated therapeutic. Compounds that modulate tissue distribution of a therapeutic agent can include a cyclic cell penetrating peptide (cCCP) and an exocyclic peptide (EP). Methods for modulating tissue distribution can comprise administering to the subject a compound that includes a cyclic cell penetrating peptide (cCPP) and an exocyclic peptide (EP). Modulation of tissue distribution or retention of a compound can be assessed by measurement of the amount, expression, function or activity of the compound *in vivo* in different tissues. The tissues can be different tissues of the same biological system, such as different types of muscle tissues or different tissues within the central nervous system. The tissue can be muscle tissue and there is modulation of distribution or retention of the compound in cardiac muscle tissue as compared to at least one other type of muscle tissue (e.g., skeletal muscle, including but not limited to diaphragm, tibialis anterior and triceps, or smooth muscle). The tissue can be CNS tissue and there is modulation of distribution or retention of the compound in at least one CNS tissue as compared to at least one other type of CNS tissue.

[0439] Any of the EPs described herein are suitable for inclusion in the compound used in the method. The EP can be PKKKRKV. The EP can be KK, KR, RR, KKK, KGK, KBK, KBR, KRK,

KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKKB, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV and PKKKRKG. The EP can be selected from KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKKB, KKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV and PKKKRKG.

[0440] The EP can comprise PKKKRKV, RR, RRR, RHR, RBR, RBRBR, RBHBR, or HBRBH, wherein B is beta-alanine. The amino acids in the EP can have D or L stereochemistry. The EP can be PKKKRKV, RR, RRR, RHR, RBR, RBRBR, RBHBR, or HBRBH, wherein B is beta-alanine. The amino acids in the EP can have D or L stereochemistry.

[0441] The EP can comprise an amino acid sequence identified in the art as a nuclear localization sequence (NLS). The EP can consist of an amino acid sequence identified in the art as a nuclear localization sequence (NLS). The EP can comprise an NLS comprising the amino acid sequence PKKKRKV. The EP can consist of an NLS comprising the amino acid sequence PKKKRKV. The EP can comprise an NLS comprising an amino acid sequence selected from NLSKRPAAIKKAGQAKKKK, PAAKRVKLD, RQRRNELKRSF, RMRKFKNKGKDTAELRRRRVEVSVELR, KAKKDEQILKRRNV, VSRKRPRP, PPKKARED, PQPKKKPL, SALIKKKKKMAP, DRLRR, PKQKKRK, RKLKKKIKKL, REKKKFLKRR, KRKGDEV DGVDEVAKKKSKK and RKCLQAGMNLEARKTKK. The EP can consist of an NLS comprising an amino acid sequence selected from NLSKRPAAIKKAGQAKKKK, PAAKRVKLD, RQRRNELKRSF, RMRKFKNKGKDTAELRRRRVEVSVELR, KAKKDEQILKRRNV, VSRKRPRP, PPKKARED, PQPKKKPL, SALIKKKKKMAP, DRLRR, PKQKKRK, RKLKKKIKKL, REKKKFLKRR, KRKGDEV DGVDEVAKKKSKK and RKCLQAGMNLEARKTKK.

[0442] The amount, expression, function or activity of the compound may be increased at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450% or 500% in at least one tissue as compared to a second tissue.

[0443] The amount, expression, function or activity of the compound may be decreased at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%,

90%, 95%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450% or 500% in at least one tissue as compared to a second tissue.

[0444] Amount or expression of the compound can be assessed in different tissue types by methods known in the art, including but not limited to methodologies described in the Examples.

5 Tissue can be prepared by standard methods. The amount or expression of the compound in different tissues can be measured by techniques well-established in the art, for example by LC-MS/MS, Western blot analysis or ELISA. The function or activity of the compound in different tissues can be measured by techniques established for assessing the relevant function or activity, such as use of RT-PCR to evaluate the activity of oligonucleotide-based therapeutic moieties. For
10 example, for an antisense compound (AC) used as the therapeutic moiety (TM) to induces exon-skipping in a target mRNA of interest, RT-PCR can be used to quantify the level of exon-skipping in different tissues.

[0445] Tissue distribution and/or retention of the therapeutic agent in tissues of the central nervous system (CNS) can be modulated with a compound that includes a cyclic cell penetrating peptide
15 (cCPP) and an exocyclic peptide (EP). The compound may be administered to the subject intrathecally and the compound may modulate tissue distribution and/or retention of the therapeutic agent in tissues of the central nervous system (CNS). Non-limiting examples of tissues of the CNS include cerebellum, cortex, hippocampus, olfactory bulb, spinal cord, dorsal root ganglion (DRG) and cerebrospinal fluid (CSF). The compound comprising a cCPP and an EP can
20 be administered intrathecally and the level of expression, activity or function of the therapeutic agent may be higher in at least one CNS tissue as compared to another CNS tissue. The compound comprising a cCPP and an EP can be administered intrathecally and the level of expression, activity or function of the therapeutic agent may be lower in at least one CNS tissue as compared to another CNS tissue. The therapeutic agent can include a CD33-targeted therapeutic agent (e.g., a CD33-
25 targeted antisense compound), wherein the compound is administered intrathecally. The compound comprising a cCPP and an EP can be administered intrathecally at a dosage of at least 1 mg/kg, 5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg, 25 mg/kg or 50 mg/kg.

[0446] Method of modulating tissue distribution or retention of a therapeutic agent in the central nervous system (CNS) of a subject may comprise: administering intrathecally to the subject a
30 compound comprising:

(a) a cyclic cell penetrating peptide (cCPP);

(b) a therapeutic moiety (TM) comprising the therapeutic agent; and

(c) an exocyclic peptide (EP) comprising at least one positively-charged amino acid residue, wherein the amount, expression, function or activity of the therapeutic agent is modulated at least 10% in at least one tissue of the CNS of the subject as compared to a second tissue of the CNS of the subject.

[0447] The amount, expression, function or activity of the therapeutic agent can be modulated at least 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450% or 500% in at least one tissue of the CNS of the subject as compared to a second tissue of the CNS of the subject.

[0448] Any of the therapeutic agents described herein for CNS-related diseases or disorders are suitable for inclusion in the compound used in the method. The therapeutic agent can include a CD33-targeted therapeutic agent, such as any of the CD33-targeted antisense compounds described herein.

[0449] Any of the CPPs described herein are suitable for inclusion in the compound used in the method. The CPP can be a cyclic CPP (cCPP).

[0450] The compound can be used to treat a subject with a central nervous system disease or disorder or a neuroinflammatory disease or disorder. In embodiments, the subject has Alzheimer's disease or Parkinson's disease.

[0451] Tissue distribution and/or retention of the therapeutic agent in different types of muscle tissues can be modulated. Non-limiting examples of muscle tissues include the diaphragm, cardiac (heart) muscle, tibialis anterior muscle, triceps muscle, other skeletal muscles and smooth muscle. A compound comprising a cCPP, EP and therapeutic agent can be administered and the level of expression, activity or function of the therapeutic agent can be higher in at least one muscle tissue as compared to another muscle tissue. A compound comprising a cCPP, EP and therapeutic agent can be administered and the level of expression, activity or function of the therapeutic agent can be lower in at least one muscle tissue as compared to another muscle tissue. The therapeutic agent can be a dystrophin-targeted therapeutic agent (e.g., a DMD-targeted antisense compound). The compound can be administered at a dosage of at least 1 mg/kg, 5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg, 25 mg/kg or 50 mg/kg.

[0452] A method of modulating tissue distribution or retention of a therapeutic agent in the muscular system of a subject comprises: administering to the subject a compound comprising:

- (a) a cyclic cell penetrating peptide (cCPP);
- (b) a therapeutic moiety (TM) comprising the therapeutic agent; and
- (c) an exocyclic peptide (EP) comprising at least one positively-charged amino acid residue, wherein the amount, expression, function or activity of the therapeutic agent is modulated at least 10% in at least one tissue of the muscular system of the subject as compared to a second tissue of the muscular system of the subject.

[0453] The amount, expression, function or activity of the therapeutic agent can be modulated at least 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450% or 500% in at least one tissue of the muscular system of the subject as compared to a second tissue of the muscular system of the subject.

[0454] Any of the therapeutic agents described herein for muscular system-related diseases or disorders are suitable for inclusion in the compound used in the method. The therapeutic agent can be a DMD-targeted therapeutic agent, such as a DMD-targeted antisense compound.

[0455] Any of the CPPs described herein are suitable for inclusion in the compound used in the method. In embodiments, the CPP is a cyclic CPP (cCPP).

[0456] In embodiments, the subject has a neuromuscular disorder or a musculoskeletal disorder. In embodiments, the subject has Duchenne muscular dystrophy.

Diseases Associated with Aberrant Splicing and Exemplary Target Genes

[0457] The human genome comprises more than 40,000 genes, approximately half of which correspond to protein-coding genes. However, the number of human protein species is predicted to be orders of magnitude higher due to single amino acid polymorphisms, post translational modifications, and, importantly, alternative splicing. RNA splicing, generally taking place in the nucleus, is the process by which precursor messenger RNA (pre-mRNA) is transformed into mature messenger RNA (mRNA) by removing non-coding regions (introns) and joining together the remaining coding regions (exons). The resulting mRNA can then be exported from the nucleus and translated into protein. Alternative splicing, or differential splicing, is a regulated process during gene expression that results in a single gene coding for multiple proteins. In this process, particular exons of a gene may be included within or excluded from the final, processed mRNA produced from that gene. While alternative splicing is a normal phenomenon in eukaryotic

organisms, and contributes to the biodiversity of proteins encoded by a genome, abnormal variations in splicing are heavily implicated in disease. A large proportion of human genetic disorders result from splicing variants; abnormal splicing variants contribute to the development of cancer; and splicing factor genes are frequently mutated in different types of cancer.

5 [0458] About 10% of ~80,000 mutations reported in the human gene mutation database (HGMD) affect splice sites. In the HGMD, there are 3390 disease-causing mutations that occur at the +1 donor splice site. These mutations affect 2754 exons in 901 genes. The prevalence is even higher for neuromuscular disorders (NMDs) due to the unusually large size and multiexonic structure of genes encoding muscle structural proteins, further highlighting the importance of these mutations
10 in NMDs.

[0459] Previously, the correction of point mutations, e.g. splice site mutations, has been attempted via the homology-directed repair (HDR) pathway, which is extremely inefficient in post-mitotic tissues such as skeletal muscles, hampering its therapeutic utility in NMD. In addition, standard gene therapy approaches to reintroduce corrected coding regions into the genome are impeded by
15 the large size of genes encoding, e.g., muscular structural proteins. Furthermore, many existing therapies rely on inefficient introduction of the therapeutic compound into the disease cells, such that in vivo treatment is impractical and higher toxicities are experienced.

[0460] The target gene of the present disclosure may be any eukaryotic gene comprising one or more introns and one or more exons. The target gene can be a mammalian gene. The mammal can
20 be a human, mouse, bovine, rat, pig, horse, chicken, sheep, or the like. The target gene can be a human gene.

[0461] The target gene can be a gene comprising mutations leading to aberrant splicing. The target gene can be a gene that comprises one or more mutations. The target gene can be a gene that comprises one or more mutations, such that transcription and translation of the target gene does
25 not lead to a functional protein. The target gene can be a gene that comprises one or more mutations, such that transcription and translation of the target gene leads to a target protein that is less active or less functional than a wild type target protein.

[0462] The target gene can be a gene underlying a genetic disorder. The target gene may have abnormal gene expression in the central nervous system. The target gene can be a gene involved
30 in the pathogenesis of a neuromuscular disorder (NMD). The target gene can be a gene involved

in the pathogenesis of a musculoskeletal disorder (NMD). The neuromuscular disease can be Pompe disease, and the target gene can be GYS1.

5 [0463] Antisense compounds may be used to target genes comprising mutations that lead to aberrant splicing underlying a genetic disease in order to redirect splicing to give a desired splice product (Kole, *Acta Biochimica Polonica*, 1997, 44, 231-238).

[0464] CRISPR gene-editing machinery may be used to target aberrant genes for removal or to regulate gene transcription and translation.

10 [0465] The disease can include β -thalassemia (Dominski and Kole, *Proc. Natl. Acad. Sci. USA*, 1993, 90, 8673-8677; Sierakowska et al., *Nucleosides & Nucleotides*, 1997, 16, 1173-1182; Sierakowska et al., *Proc. Natl. Acad. Sci. USA*, 1996, 93, 12840-44; Lacerra et al., *Proc. Natl. Acad. Sci. USA*, 2000, 97, 9591-9596).

[0466] The disease can include dystrophin Kobe (Takeshima et al., *J. Clin. Invest.*, 1995, 95, 515-520).

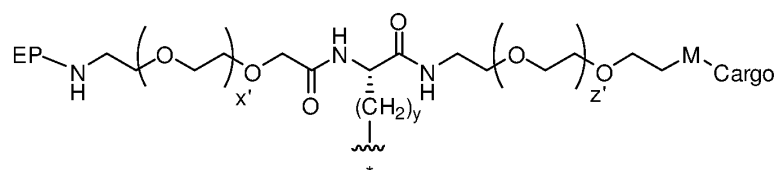
15 [0467] The disease can include Duchenne muscular dystrophy (Dunckley et al. *Nucleosides & Nucleotides*, 1997, 16, 1665-1668; Dunckley et al. *Human Mol. Genetics*, 1998, 5, 1083-90). The target gene can be the DMD gene, which codes for dystrophin. The protein consists of an N-terminal domain that binds to actin filaments, a central rod domain, and a C-terminal cysteine-rich domain that binds to the dystrophin-glycoprotein complex (Hoffman et al. 1987; Koenig et al. 1988; Yoshida and Ozawa 1990). Mutations in the DMD gene that interrupt the reading frame
20 result in a complete loss of dystrophin function, which causes severe Duchenne muscular dystrophy (DMD) [MIM 310200]. The milder Becker muscular dystrophy (BMD [MIM 300376]), on the other hand, is the result of mutations in the same gene that are not frameshifting and result in an internally deleted but partially functional dystrophin that has retained its N- and C-terminal ends (Koenig et al. 1989; Di Blasi et al. 1996). Over two-thirds of patients with DMD
25 and BMD have a deletion of >1 exon (den Dun-nen et al. 1989). Remarkably, patients have been described who exhibit very mild BMD and who lack up to 67% of the central rod domain (England et al. 1990; Winnard et al. 1993; Mirabella et al. 1998). This suggests that, despite large deletions, a partially functional dystrophin can be generated, provided that the deletions render the transcript in frame. These observations have led to the idea of using ACs to alter splicing so that the open
30 reading frame is restored and the severe DMD phenotype is converted into a milder BMD phenotype. Several studies have shown therapeutic AC-induced single-exon skipping in cells

derived from the mdx mouse model (Dunckley et al. 1998; Wilton et al. 1999; Mann et al. 2001, 2002; Lu et al. 2003) and various DMD patients (Takeshima et al. 2001; van Deutekom et al. 2001; Aartsma-Rus et al. 2002, 2003; De Angelis et al. 2002). The AC can be used to skip one or more exons selected from exons 2, 8, 11, 17, 19, 23, 29, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 51, 52, 53, 55, and 59 of DMD. See Aartsma-Rus et al. 2002, incorporated by reference herein. The AC can be used to skip one or more exons selected from exons 8, 11, 43, 44, 45, 50, 51, 53, and 55 of DMD. Of all patients with DMD, ~75% would benefit from the skipping of these exons. The skipping of exons flanking out-of-frame deletions or an in-frame exon containing a nonsense mutation can restore the reading frame and induce the synthesis of BMD-like dystrophins in treated cells. (van Deutekom et al. 2001; Aartsma-Rus et al. 2003). The AC hybridizing to its target sequence within a target DMD pre-mRNA can induce skipping of one or more exons. The AC can induce expression of a re-spliced target protein comprising an active fragment of dystrophin. Non-limiting examples of AC for exon 52 are described in US Pub. No. 2019/0365918, which is incorporated by reference in its entirety for all purposes. Compounds may comprise an EP, cCPP and a cargo that target the DMD gene.

Cyclic cell penetrating peptides (cCPPs) conjugated to a cargo moiety

[0468] The cyclic cell penetrating peptide (cCPP) can be conjugated to a cargo moiety.

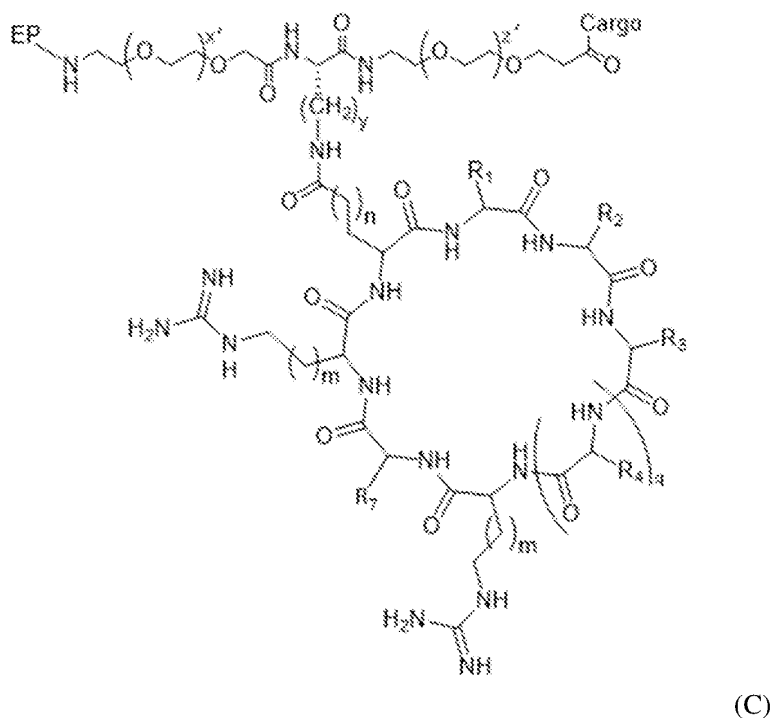
[0469] The cargo moiety can be conjugated to cCPP through a linker. The cargo moiety can comprise therapeutic moiety. The therapeutic moiety can comprise an oligonucleotide, a peptide or a small molecule. The oligonucleotide can comprise an antisense oligonucleotide. The cargo moiety can be conjugated to the linker at the terminal carbonyl group to provide the following structure:



, wherein:

EP is an exocyclic peptide and M, AASc, Cargo, x', y, and z' are as defined above, * is the point of attachment to the AASc. x' can be 1. y can be 4. z' can be 11. -(OCH₂CH₂)_{x'}- and/or -(OCH₂CH₂)_{z'}- can be independently replaced with one or more amino acids, including, for example, glycine, beta-alanine, 4-aminobutyric acid, 5-aminopentanoic acid, 6-aminohexanoic acid, or combinations thereof.

[0470] An endosomal escape vehicle (EEV) can comprise a cyclic cell penetrating peptide (cCPP), an exocyclic peptide (EP) and linker, and can be conjugated to a cargo to form an EEV-conjugate comprising the structure of Formula (C):



- 5 or a protonated form thereof,
wherein:

R_1 , R_2 , and R_3 can each independently be H or an amino acid residue having a side chain comprising an aromatic group;

R_4 is H or an amino acid side chain;

- 10 EP is an exocyclic peptide as defined herein;

Cargo is a moiety as defined herein;

each m is independently an integer from 0-3;

n is an integer from 0-2;

x' is an integer from 2-20;

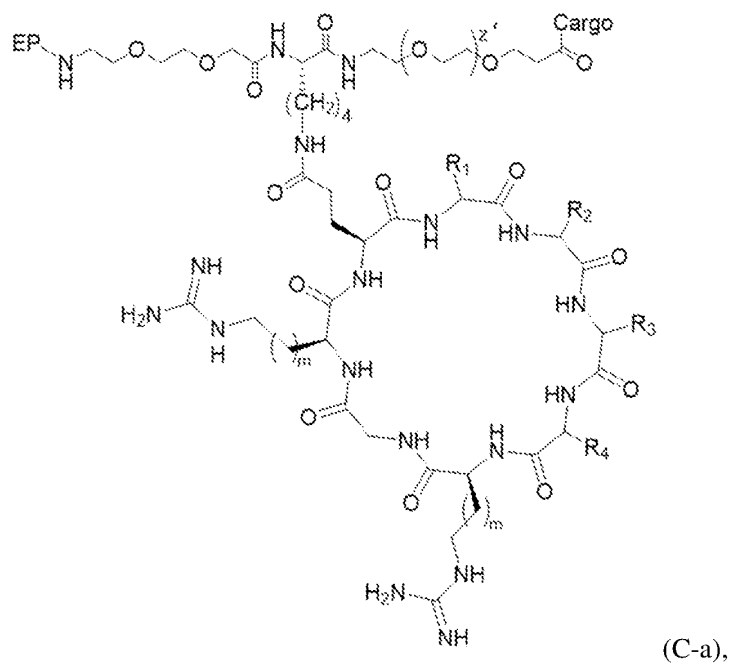
- 15 y is an integer from 1-5;

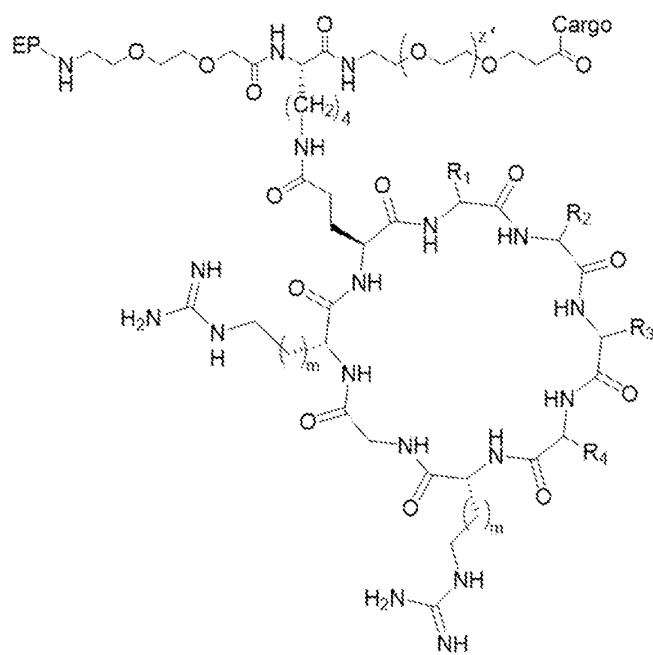
q is an integer from 1-4; and

z' is an integer from 2-20.

[0471] R_1 , R_2 , R_3 , R_4 , EP, cargo, m , n , x' , y , q , and z' are as defined herein.

[0472] The EEV can be conjugated to a cargo and the EEV-conjugate can comprise the structure of Formula (C-a) or (C-b):

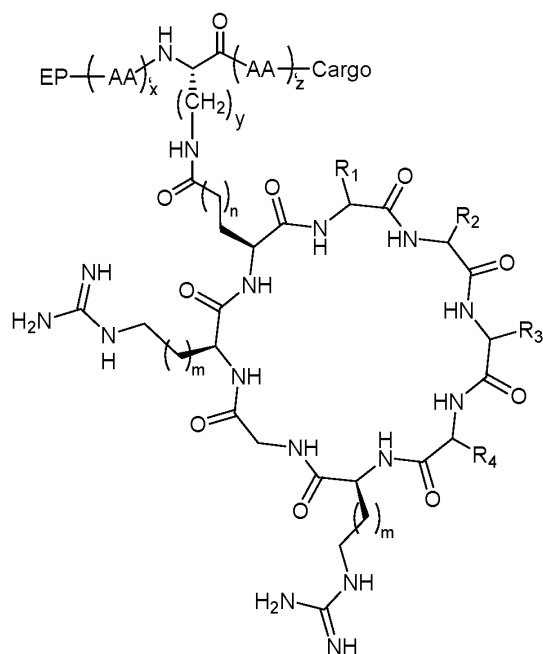




(C-b), or a protonated form thereof,

wherein EP , m and z are as defined above in Formula (C).

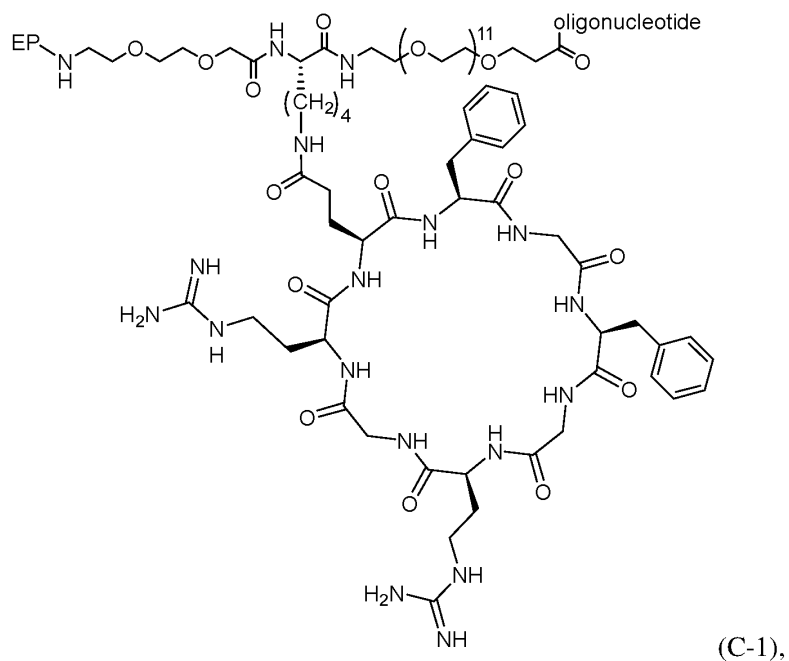
[0473] The EEV can be conjugated to a cargo and the EEV-conjugate can comprise the structure of Formula (C-c):



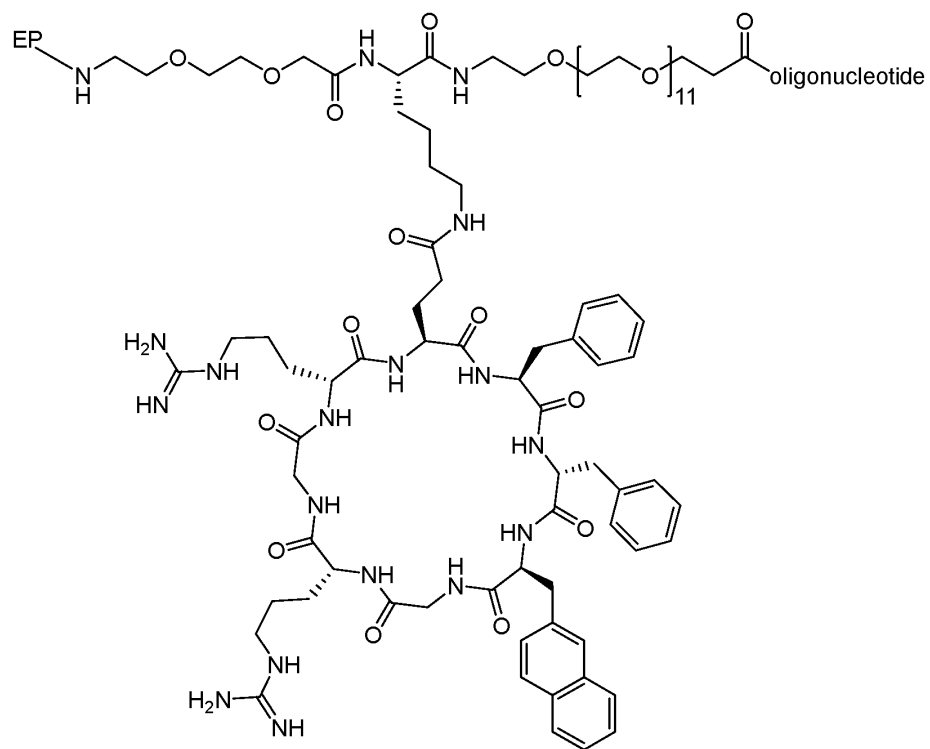
(C-c),

or a protonated form thereof, wherein EP, R¹, R², R³, R⁴, and m are as defined above in Formula (III); AA can be an amino acid as defined herein; n can be an integer from 0-2; x can be an integer from 1-10; y can be an integer from 1-5; and z can be an integer from 1-10.

- [0474] The EEV can be conjugated to an oligonucleotide cargo and the EEV-oligonucleotide conjugate can comprises a structure of Formula (C-1), (C-2), (C-3), or (C-4):

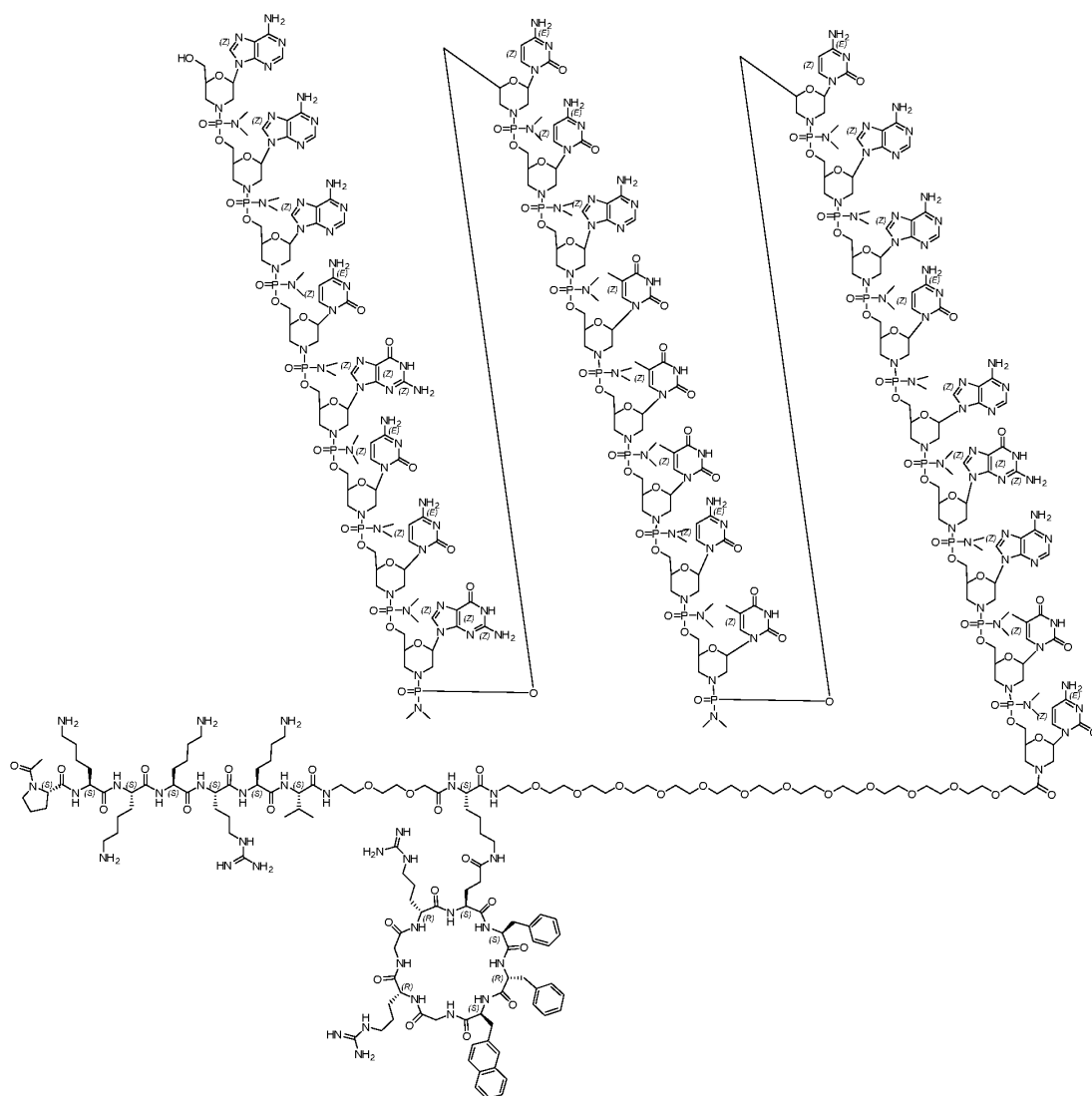






(C-4)

[0475] The EEV can be conjugated to an oligonucleotide cargo and the EEV-conjugate can comprise the structure:



5 *Cytosolic Delivery Efficiency*

[0476] Modifications to a cyclic cell penetrating peptide (cCPP) may improve cytosolic delivery efficiency. Improved cytosolic uptake efficiency can be measured by comparing the cytosolic delivery efficiency of a cCPP having a modified sequence to a control sequence. The control sequence does not include a particular replacement amino acid residue in the modified sequence
10 (including, but not limited to arginine, phenylalanine, and/or glycine), but is otherwise identical.

[0477] As used herein cytosolic delivery efficiency refers to the ability of a cCPP to traverse a cell membrane and enter the cytosol of a cell. Cytosolic delivery efficiency of the cCPP is not necessarily dependent on a receptor or a cell type. Cytosolic delivery efficiency can refer to absolute cytosolic delivery efficiency or relative cytosolic delivery efficiency.

15 [0478] Absolute cytosolic delivery efficiency is the ratio of cytosolic concentration of a cCPP (or a cCPP-cargo conjugate) over the concentration of the cCPP (or the cCPP-cargo conjugate) in the growth medium. Relative cytosolic delivery efficiency refers to the concentration of a cCPP in the cytosol compared to the concentration of a control cCPP in the cytosol. Quantification can be achieved by fluorescently labeling the cCPP (e.g., with a FITC dye) and measuring the
20 fluorescence intensity using techniques well-known in the art.

[0479] Relative cytosolic delivery efficiency is determined by comparing (i) the amount of a cCPP of the invention internalized by a cell type (e.g., HeLa cells) to (ii) the amount of a control cCPP internalized by the same cell type. To measure relative cytosolic delivery efficiency, the cell type may be incubated in the presence of a cCPP for a specified period of time (e.g., 30 minutes, 1 hour,
25 2 hours, etc.) after which the amount of the cCPP internalized by the cell is quantified using methods known in the art, e.g., fluorescence microscopy. Separately, the same concentration of the control cCPP is incubated in the presence of the cell type over the same period of time, and the amount of the control cCPP internalized by the cell is quantified.

[0480] Relative cytosolic delivery efficiency can be determined by measuring the IC_{50} of a cCPP
30 having a modified sequence for an intracellular target and comparing the IC_{50} of the cCPP having the modified sequence to a control sequence (as described herein).

[0481] The relative cytosolic delivery efficiency of the cCPPs can be in the range of from about 50% to about 450% compared to cyclo(FfΦRrRrQ), e.g., about 60%, about 70%, about 80%, about 90%, about 100%, about 110%, about 120%, about 130%, about 140%, about 150%, about 160%, about 170%, about 180%, about 190%, about 200%, about 210%, about 220%, about 230%, about 240%, about 250%, about 260%, about 270%, about 280%, about 290%, about 300%, about 310%, about 320%, about 330%, about 340%, about 350%, about 360%, about 370%, about 380%, about 390%, about 400%, about 410%, about 420%, about 430%, about 440%, about 450%, about 460%, about 470%, about 480%, about 490%, about 500%, about 510%, about 520%, about 530%, about 540%, about 550%, about 560%, about 570%, about 580%, or about 590%, inclusive of all values and subranges therebetween. The relative cytosolic delivery efficiency of the cCPPs can be improved by greater than about 600% compared to a cyclic peptide comprising cyclo(FfΦRrRrQ).

[0482] The absolute cytosolic delivery efficacy of from about 40% to about 100%, e.g., about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, inclusive of all values and subranges therebetween.

[0483] The cCPPs of the present disclosure can improve the cytosolic delivery efficiency by about 1.1 fold to about 30 fold, compared to an otherwise identical sequence, e.g., about 1.2, about 1.3, about 1.4, about 1.5, about 1.6, about 1.7, about 1.8, about 1.9, about 2.0, about 2.5, about 3.0, about 3.5, about 4.0, about 4.5, about 5.0, about 5.5, about 6.0, about 6.5, about 7.0, about 7.5, about 8.0, about 8.5, about 9.0, about 10, about 10.5, about 11.0, about 11.5, about 12.0, about 12.5, about 13.0, about 13.5, about 14.0, about 14.5, about 15.0, about 15.5, about 16.0, about 16.5, about 17.0, about 17.5, about 18.0, about 18.5, about 19.0, about 19.5, about 20, about 20.5, about 21.0, about 21.5, about 22.0, about 22.5, about 23.0, about 23.5, about 24.0, about 24.5, about 25.0, about 25.5, about 26.0, about 26.5, about 27.0, about 27.5, about 28.0, about 28.5, about 29.0, or about 29.5 fold, inclusive of all values and subranges therebetween.

Methods of Making

[0484] The compounds described herein can be prepared in a variety of ways known to one skilled in the art of organic synthesis or variations thereon as appreciated by those skilled in the art. The compounds described herein can be prepared from readily available starting materials. Optimum reaction conditions can vary with the particular reactants or solvents used, but such conditions can be determined by one skilled in the art.

[0485] Variations on the compounds described herein include the addition, subtraction, or movement of the various constituents as described for each compound. Similarly, when one or more chiral centers are present in a molecule, the chirality of the molecule can be changed. Additionally, compound synthesis can involve the protection and deprotection of various chemical groups. The use of protection and deprotection, and the selection of appropriate protecting groups can be determined by one skilled in the art. The chemistry of protecting groups can be found, for example, in Wuts and Greene, *Protective Groups in Organic Synthesis*, 4th Ed., Wiley & Sons, 2006, which is incorporated herein by reference in its entirety.

[0486] The starting materials and reagents used in preparing the disclosed compounds and compositions are either available from commercial suppliers such as Aldrich Chemical Co., (Milwaukee, WI), Acros Organics (Morris Plains, NJ), Fisher Scientific (Pittsburgh, PA), Sigma (St. Louis, MO), Pfizer (New York, NY), GlaxoSmithKline (Raleigh, NC), Merck (Whitehouse Station, NJ), Johnson & Johnson (New Brunswick, NJ), Aventis (Bridgewater, NJ), AstraZeneca (Wilmington, DE), Novartis (Basel, Switzerland), Wyeth (Madison, NJ), Bristol-Myers-Squibb (New York, NY), Roche (Basel, Switzerland), Lilly (Indianapolis, IN), Abbott (Abbott Park, IL), Schering Plough (Kenilworth, NJ), or Boehringer Ingelheim (Ingelheim, Germany), or are prepared by methods known to those skilled in the art following procedures set forth in references such as Fieser and Fieser's *Reagents for Organic Synthesis*, Volumes 1-17 (John Wiley and Sons, 1991); Rodd's *Chemistry of Carbon Compounds*, Volumes 1-5 and Supplementals (Elsevier Science Publishers, 1989); *Organic Reactions*, Volumes 1-40 (John Wiley and Sons, 1991); March's *Advanced Organic Chemistry*, (John Wiley and Sons, 4th Edition); and Larock's *Comprehensive Organic Transformations* (VCH Publishers Inc., 1989). Other materials, such as the pharmaceutical carriers disclosed herein can be obtained from commercial sources.

[0487] Reactions to produce the compounds described herein can be carried out in solvents, which can be selected by one of skill in the art of organic synthesis. Solvents can be substantially nonreactive with the starting materials (reactants), the intermediates, or products under the conditions at which the reactions are carried out, *i.e.*, temperature and pressure. Reactions can be carried out in one solvent or a mixture of more than one solvent. Product or intermediate formation can be monitored according to any suitable method known in the art. For example, product formation can be monitored by spectroscopic means, such as nuclear magnetic resonance spectroscopy (*e.g.*, ^1H or ^{13}C) infrared spectroscopy, spectrophotometry (*e.g.*, UV-visible), or mass

spectrometry, or by chromatography such as high performance liquid chromatography (HPLC) or thin layer chromatography.

[0488] The disclosed compounds can be prepared by solid phase peptide synthesis wherein the amino acid α -N-terminus is protected by an acid or base protecting group. Such protecting groups should have the properties of being stable to the conditions of peptide linkage formation while being readily removable without destruction of the growing peptide chain or racemization of any of the chiral centers contained therein. Suitable protecting groups are 9-fluorenylmethyloxycarbonyl (Fmoc), t-butyloxycarbonyl (Boc), benzyloxycarbonyl (Cbz), biphenylisopropyloxycarbonyl, t-amylloxycarbonyl, isobornyloxycarbonyl, α,α -dimethyl-3,5-dimethoxybenzyloxycarbonyl, o-nitrophenylsulfonyl, 2-cyano-t-butyloxycarbonyl, and the like. The 9-fluorenylmethyloxycarbonyl (Fmoc) protecting group is particularly preferred for the synthesis of the disclosed compounds. Other preferred side chain protecting groups are, for side chain amino groups like lysine and arginine, 2,2,5,7,8-pentamethylchroman-6-sulfonyl (pmc), nitro, p-toluenesulfonyl, 4-methoxybenzene-sulfonyl, Cbz, Boc, and adamantyloxycarbonyl; for tyrosine, benzyl, o-bromobenzyloxy-carbonyl, 2,6-dichlorobenzyl, isopropyl, t-butyl (t-Bu), cyclohexyl, cyclopentyl and acetyl (Ac); for serine, t-butyl, benzyl and tetrahydropyranyl; for histidine, trityl, benzyl, Cbz, p-toluenesulfonyl and 2,4-dinitrophenyl; for tryptophan, formyl; for aspartic acid and glutamic acid, benzyl and t-butyl and for cysteine, triphenylmethyl (trityl).

[0489] In the solid phase peptide synthesis method, the α -C-terminal amino acid is attached to a suitable solid support or resin. Suitable solid supports useful for the above synthesis are those materials which are inert to the reagents and reaction conditions of the stepwise condensation-deprotection reactions, as well as being insoluble in the media used. Solid supports for synthesis of α -C-terminal carboxy peptides is 4-hydroxymethylphenoxymethyl-copoly(styrene-1% divinylbenzene) or 4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)phenoxyacetamidoethyl resin available from Applied Biosystems (Foster City, Calif.). The α -C-terminal amino acid is coupled to the resin by means of N,N'-dicyclohexylcarbodiimide (DCC), N,N'-diisopropylcarbodiimide (DIC) or O-benzotriazol-1-yl-N,N,N',N'-tetramethyluroniumhexafluorophosphate (HBTU), with or without 4-dimethylaminopyridine (DMAP), 1-hydroxybenzotriazole (HOBT), benzotriazol-1-yloxy-tris(dimethylamino)phosphoniumhexafluorophosphate (BOP) or bis(2-oxo-3-oxazolidinyl)phosphine chloride (BOPCl), mediated coupling for from about 1 to about 24 hours at a temperature of between 10°C and 50°C in a solvent such as dichloromethane or DMF. When

the solid support is 4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)phenoxy-acetamidoethyl resin, the Fmoc group is cleaved with a secondary amine, preferably piperidine, prior to coupling with the α -C-terminal amino acid as described above. One method for coupling to the deprotected 4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)phenoxy-acetamidoethyl resin is O-benzotriazol-1-yl-N,N,N',N'-tetramethyluroniumhexafluorophosphate (HBTU, 1 equiv.) and 1-hydroxybenzotriazole (HOBT, 1 equiv.) in DMF. The coupling of successive protected amino acids can be carried out in an automatic polypeptide synthesizer. In one example, the α -N-terminus in the amino acids of the growing peptide chain are protected with Fmoc. The removal of the Fmoc protecting group from the α -N-terminal side of the growing peptide is accomplished by treatment with a secondary amine, preferably piperidine. Each protected amino acid is then introduced in about 3-fold molar excess, and the coupling is preferably carried out in DMF. The coupling agent can be O-benzotriazol-1-yl-N,N,N',N'-tetramethyluroniumhexafluorophosphate (HBTU, 1 equiv.) and 1-hydroxybenzotriazole (HOBT, 1 equiv.). At the end of the solid phase synthesis, the polypeptide is removed from the resin and deprotected, either successively or in a single operation. Removal of the polypeptide and deprotection can be accomplished in a single operation by treating the resin-bound polypeptide with a cleavage reagent comprising thianisole, water, ethanedithiol and trifluoroacetic acid. In cases wherein the α -C-terminal of the polypeptide is an alkylamide, the resin is cleaved by aminolysis with an alkylamine. Alternatively, the peptide can be removed by transesterification, *e.g.* with methanol, followed by aminolysis or by direct transamidation. The protected peptide can be purified at this point or taken to the next step directly. The removal of the side chain protecting groups can be accomplished using the cleavage cocktail described above. The fully deprotected peptide can be purified by a sequence of chromatographic steps employing any or all of the following types: ion exchange on a weakly basic resin (acetate form); hydrophobic adsorption chromatography on underivitized polystyrene-divinylbenzene (for example, Amberlite XAD); silica gel adsorption chromatography; ion exchange chromatography on carboxymethylcellulose; partition chromatography, *e.g.* on Sephadex G-25, LH-20 or countercurrent distribution; high performance liquid chromatography (HPLC), especially reverse-phase HPLC on octyl- or octadecylsilyl-silica bonded phase column packing.

[0490] Methods of synthesizing oligomeric antisense compounds are known in the art. The present disclosure is not limited by the method of synthesizing the AC. In embodiments, provided herein are compounds having reactive phosphorus groups useful for forming internucleoside linkages

including for example phosphodiester and phosphorothioate internucleoside linkages. Methods of preparation and/or purification of precursors or antisense compounds are not a limitation of the compositions or methods provided herein. Methods for synthesis and purification of DNA, RNA, and the antisense compounds are well known to those skilled in the art.

5 [0491] Oligomerization of modified and unmodified nucleosides can be routinely performed according to literature procedures for DNA (Protocols for Oligonucleotides and Analogs, Ed. Agrawal (1993), Humana Press) and/or RNA (Scaringe, Methods (2001), 23, 206-217. Gait et al., Applications of Chemically synthesized RNA in RNA: Protein Interactions, Ed. Smith (1998), 1-36. Gallo et al., Tetrahedron (2001), 57, 5707-5713).

10 [0492] Antisense compounds provided herein can be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, CA). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated
15 derivatives. The invention is not limited by the method of antisense compound synthesis.

[0493] Methods of oligonucleotide purification and analysis are known to those skilled in the art. Analysis methods include capillary electrophoresis (CE) and electrospray-mass spectroscopy. Such synthesis and analysis methods can be performed in multi-well plates. The method of the invention is not limited by the method of oligomer purification.

20

Methods of Use

[0494] Also provided herein are methods of use of the compounds or compositions described herein. Also provided herein are methods for treating a disease or pathology in a subject in need thereof comprising administering to the subject an effective amount of any of the compounds or
25 compositions described herein. The compounds or compositions can be used to treat any disease or condition that is amenable to treatment with the therapeutic moieties disclosed herein.

[0495] Also provided herein are methods of treating cancer in a subject. The methods include administering to a subject an effective amount of one or more of the compounds or compositions described herein, or a pharmaceutically acceptable salt thereof. The compounds and compositions
30 described herein or pharmaceutically acceptable salts thereof are useful for treating cancer in humans, *e.g.*, pediatric and geriatric populations, and in animals, *e.g.*, veterinary applications. The

disclosed methods can optionally include identifying a patient who is or can be in need of treatment of a cancer. Examples of cancer types treatable by the compounds and compositions described herein include bladder cancer, brain cancer, breast cancer, colorectal cancer, cervical cancer, gastrointestinal cancer, genitourinary cancer, head and neck cancer, lung cancer, ovarian cancer, 5 pancreatic cancer, prostate cancer, renal cancer, skin cancer, and testicular cancer. Further examples include cancer and/or tumors of the anus, bile duct, bone, bone marrow, bowel (including colon and rectum), eye, gall bladder, kidney, mouth, larynx, esophagus, stomach, testis, cervix, mesothelioma, neuroendocrine, penis, skin, spinal cord, thyroid, vagina, vulva, uterus, liver, muscle, blood cells (including lymphocytes and other immune system cells). Further examples of 10 cancers treatable by the compounds and compositions described herein include carcinomas, Karposi's sarcoma, melanoma, mesothelioma, soft tissue sarcoma, pancreatic cancer, lung cancer, leukemia (acute lymphoblastic, acute myeloid, chronic lymphocytic, chronic myeloid, and other), and lymphoma (Hodgkin's and non-Hodgkin's), and multiple myeloma.

[0496] The methods of treatment or prevention of cancer described herein can further include 15 treatment with one or more additional agents (*e.g.*, an anti-cancer agent or ionizing radiation). The one or more additional agents and the compounds and compositions or pharmaceutically acceptable salts thereof as described herein can be administered in any order, including simultaneous administration, as well as temporally spaced order of up to several days apart. The methods can also include more than a single administration of the one or more additional agents 20 and/or the compounds and compositions or pharmaceutically acceptable salts thereof as described herein. The administration of the one or more additional agents and the compounds and compositions or pharmaceutically acceptable salts thereof as described herein can be by the same or different routes. When treating with one or more additional agents, the compounds and compositions or pharmaceutically acceptable salts thereof as described herein can be combined 25 into a pharmaceutical composition that includes the one or more additional agents.

[0497] For example, the compounds or compositions or pharmaceutically acceptable salts thereof as described herein can be combined into a pharmaceutical composition with an additional anti-cancer agent, such as 13-cis-Retinoic Acid, 2-Amino-6-Mercaptopurine, 2-CdA, 2-Chlorodeoxyadenosine, 5-fluorouracil, 6-Thioguanine, 6-Mercaptopurine, Accutane, 30 Actinomycin-D, Adriamycin, Atracuril, Agrylin, Ala-Cort, Aldesleukin, Alectinib, Alisertib, Alitretinoin, Alkaban-AQ, Alkeran, All-transretinoic acid, Alpha interferon, Altretamine,

Amethopterin, Amifostine, Aminoglutethimide, Anagrelide, Anandron, Anastrozole, Arabinosylcytosine, Aranesp, Aredia, Arimidex, Aromasin, Arsenic trioxide, Asparaginase, ATRA, Avastin, BCG, BCNU, Bevacizumab, Bexarotene, Bicalutamide, BiCNU, Blenoxane, Bleomycin, Bortezomib, Busulfan, Busulfex, C225, Calcium Leucovorin, Campath, Camptosar, 5 Camptothecin-11, Capecitabine, Carac, Carboplatin, Carmustine, Carmustine wafer, Casodex, CCNU, CDDP, CeeNU, Cerubidine, cetuximab, Chlorambucil, Cisplatin, Citrovorium Factor, Cladribine, Cortisone, Cosmegen, CPT-11, Cyclophosphamide, Cytadren, Cytarabine, Cytarabine liposomal, Cytosar-U, Cytosan, Dacarbazine, Dactinomycin, Darbepoetin alfa, Daunomycin, Daunorubicin, Daunorubicin hydrochloride, Daunorubicin liposomal, DaunoXome, Decadron, 10 Delta-Cortef, Deltasone, Denileukin diftitox, DepoCyt, Dexamethasone, Dexamethasone acetate, Dexamethasone sodium phosphate, Dexasone, Dexrazoxane, DHAD, DIC, Diodex, Docetaxel, Doxil, Doxorubicin, Doxorubicin liposomal, Droxia, DTIC, DTIC-Dome, Duralone, Efudex, Eligard, Ellence, Eloxatin, Elspar, Emcyt, Epirubicin, Epoetin alfa, Erbitux, Erwinia L-asparaginase, Estramustine, Ethyol, Etopophos, Etoposide, Etoposide phosphate, Eulexin, Evista, 15 Exemestane, Fareston, Faslodex, Femara, Filgrastim, Floxuridine, Fludara, Fludarabine, Fluoroplex, Fluorouracil, Fluorouracil (cream), Fluoxymesterone, Flutamide, Folinic Acid, FUDR, Fulvestrant, G-CSF, Gefitinib, Gemcitabine, Gemtuzumab ozogamicin, Gemzar, Gleevec, Lupron, Lupron Depot, Matulane, Maxidex, Mechlorethamine, -Mechlorethamine Hydrochlorine, Medralone, Medrol, Megace, Megestrol, Megestrol Acetate, Melphalan, Mercaptopurine, Mesna, 20 Mesnex, Methotrexate, Methotrexate Sodium, Methylprednisolone, Mylocel, Letrozole, Neosar, Neulasta, Neumega, Neupogen, Nilandron, Nilutamide, Nitrogen Mustard, Novaldex, Novantrone, Octreotide, Octreotide acetate, Oncospar, Oncovin, Ontak, Onxal, Oprevelkin, Orapred, Orasone, Oxaliplatin, Paclitaxel, Pamidronate, Panretin, Paraplatin, Pediapred, PEG Interferon, Pegaspargase, Pegfilgrastim, PEG-INTRON, PEG-L-asparaginase, Phenylalanine Mustard, 25 Platinol, Platinol-AQ, Prednisolone, Prednisone, Prelone, Procarbazine, PROCRIT, Proleukin, Prolifeprospan 20 with Carmustine implant, Purinethol, Raloxifene, Rheumatrex, Rituxan, Rituximab, Roveron-A (interferon alfa-2a), Rubex, Rubidomycin hydrochloride, Sandostatin, Sandostatin LAR, Sargramostim, Solu-Cortef, Solu-Medrol, STI-571, Streptozocin, Tamoxifen, Targretin, Taxol, Taxotere, Temodar, Temozolomide, Teniposide, TESPA, Thalidomide, 30 Thalomid, TheraCys, Thioguanine, Thioguanine Tabloid, Thiophosphoamide, Thioplex, Thiotepa, TICE, Toposar, Topotecan, Toremifene, Trastuzumab, Tretinoin, Trexall, Trisenox, TSPA, VCR,

Velban, Velcade, VePesid, Vesanoid, Viadur, Vinblastine, Vinblastine Sulfate, Vincasar Pfs, Vincristine, Vinorelbine, Vinorelbine tartrate, VLB, VP-16, Vumon, Xeloda, Zanosar, Zevalin, Zinecard, Zoladex, Zoledronic acid, Zometa, Gliadel wafer, Glivec, GM-CSF, Goserelin, granulocyte colony stimulating factor, Halotestin, Herceptin, Hexadrol, Hexalen, 5 Hexamethylmelamine, HMM, Hycamtin, Hydrea, Hydrocort Acetate, Hydrocortisone, Hydrocortisone sodium phosphate, Hydrocortisone sodium succinate, Hydrocortone phosphate, Hydroxyurea, Ibritumomab, Ibritumomab Tiuxetan, Idamycin, Idarubicin, Ifex, IFN-alpha, Ifosfamide, IL 2, IL-11, Imatinib mesylate, Imidazole Carboxamide, Interferon alfa, Interferon Alfa-2b (PEG conjugate), Interleukin 2, Interleukin-11, Intron A (interferon alfa-2b), Leucovorin, 10 Leukeran, Leukine, Leuprolide, Leurocristine, Leustatin, Liposomal Ara-C, Liquid Pred, Lomustine, L-PAM, L-Sarcosine, Meticorten, Mitomycin, Mitomycin-C, Mitoxantrone, M-Prednisol, MTC, MTX, Mustargen, Mustine, Mutamycin, Myleran, Iressa, Irinotecan, Isotretinoin, Kidrolase, Lanacort, L-asparaginase, and LCR. The additional anti-cancer agent can also include biopharmaceuticals such as, for example, antibodies.

15 [0498] Many tumors and cancers have viral genome present in the tumor or cancer cells. For example, Epstein-Barr Virus (EBV) is associated with a number of mammalian malignancies. The compounds disclosed herein can also be used alone or in combination with anticancer or antiviral agents, such as ganciclovir, azidothymidine (AZT), lamivudine (3TC), *etc.*, to treat patients infected with a virus that can cause cellular transformation and/or to treat patients having a tumor 20 or cancer that is associated with the presence of viral genome in the cells. The compounds disclosed herein can also be used in combination with viral based treatments of oncologic disease.

[0499] Also described herein are methods of killing a tumor cell in a subject. The method includes contacting the tumor cell with an effective amount of a compound or composition as described herein, and optionally includes the step of irradiating the tumor cell with an effective amount of 25 ionizing radiation. Additionally, methods of radiotherapy of tumors are provided herein. The methods include contacting the tumor cell with an effective amount of a compound or composition as described herein, and irradiating the tumor with an effective amount of ionizing radiation. As used herein, the term ionizing radiation refers to radiation comprising particles or photons that have sufficient energy or can produce sufficient energy via nuclear interactions to produce 30 ionization. An example of ionizing radiation is x-radiation. An effective amount of ionizing radiation refers to a dose of ionizing radiation that produces an increase in cell damage or death

when administered in combination with the compounds described herein. The ionizing radiation can be delivered according to methods as known in the art, including administering radiolabeled antibodies and radioisotopes.

[0500] The methods and compounds as described herein are useful for both prophylactic and therapeutic treatment. As used herein the term treating or treatment includes prevention; delay in onset; diminution, eradication, or delay in exacerbation of signs or symptoms after onset; and prevention of relapse. For prophylactic use, a therapeutically effective amount of the compounds and compositions or pharmaceutically acceptable salts thereof as described herein are administered to a subject prior to onset (*e.g.*, before obvious signs of cancer), during early onset (*e.g.*, upon initial signs and symptoms of cancer), or after an established development of cancer. Prophylactic administration can occur for several days to years prior to the manifestation of symptoms of an infection. Prophylactic administration can be used, for example, in the chemopreventative treatment of subjects presenting precancerous lesions, those diagnosed with early stage malignancies, and for subgroups with susceptibilities (*e.g.*, family, racial, and/or occupational) to particular cancers. Therapeutic treatment involves administering to a subject a therapeutically effective amount of the compounds and compositions or pharmaceutically acceptable salts thereof as described herein after cancer is diagnosed.

[0501] In some examples of the methods of treating of treating cancer or a tumor in a subject, the compound or composition administered to the subject can comprise a therapeutic moiety that can comprise a targeting moiety that can act as an inhibitor against Ras (*e.g.*, K-Ras), PTP1B, Pin1, Grb2 SH2, or combinations thereof.

[0502] The disclosed subject matter also concerns methods for treating a subject having a metabolic disorder or condition. An effective amount of one or more compounds or compositions disclosed herein can be administered to a subject having a metabolic disorder and who is in need of treatment thereof. In some examples, the metabolic disorder can comprise type II diabetes. In some examples of the methods of treating of treating the metabolic disorder in a subject, the compound or composition administered to the subject can comprise a therapeutic moiety that can comprise a targeting moiety that can act as an inhibitor against PTP1B. In one particular example of this method the subject is obese, and the method can comprise treating the subject for obesity by administering a composition as disclosed herein.

[0503] The disclosed subject matter also concerns methods for treating a subject having an immune disorder or condition. An effective amount of one or more compounds or compositions disclosed herein is administered to a subject having an immune disorder and who is in need of treatment thereof. In some examples of the methods of treating of treating the immune disorder in a subject, the compound or composition administered to the subject can comprise a therapeutic moiety that can comprise a targeting moiety that can act as an inhibitor against Pin1.

[0504] The disclosed subject matter also concerns methods for treating a subject having an inflammatory disorder or condition. An effective amount of one or more compounds or compositions disclosed herein can be administered to a subject having an inflammatory disorder and who is in need of treatment thereof.

[0505] The disclosed subject matter also concerns methods for treating a subject having cystic fibrosis. An effective amount of one or more compounds or compositions disclosed herein can be administered to a subject having cystic fibrosis and who is in need of treatment thereof. In some examples of the methods of treating the cystic fibrosis in a subject, the compound or composition administered to the subject can comprise a therapeutic moiety that can comprise a targeting moiety that can act as an inhibitor against CAL PDZ.

[0506] The compounds disclosed herein can be used for detecting or diagnosing a disease or condition in a subject. For example, a cCPP can comprise a targeting moiety and/or a detectable moiety that can interact with a target, e.g., a tumor.

[0507] In some embodiments, the disease is associated with insulin resistance. In some embodiments, the disease is diabetes. In some embodiments, the target gene is PTP.

[0508] In some embodiments, the disease is a CNS disorder. In some embodiments, the disease is Alzheimer's Disease (AD) (Zhao et al. Gerontology 2019;65:323–331). In some embodiments, the target gene is the CD33 gene. The CD33 gene maps on chromosome 19q13.33 in humans encoding a 67-kDa transmembrane glycoprotein. Human CD33 binds preferentially to alpha-2,6-linked sialic acid. CD33 is expressed exclusively on immune cells. CD33 is an inhibitory receptor that recruits inhibitory proteins such as SHP phosphatases via its immunoreceptor tyrosine-based inhibition motif (ITIM). CD33 is also involved in adhesion processes in immune or malignant cells, inhibition of cytokine release by monocytes, immune cell growth and survival through the inhibition of proliferation, and induction of apoptosis. Polymorphisms of CD33 have been implicated in modulating AD susceptibility. rs3865444C is an allele associated with an increased

risk of AD in European, Chinese, and North American populations due to increased expression of CD33. The skipping of exon 2 of CD33 leads to a decreased expression of CD33 and an increased expression of D2-CD33, which is a CD33 isoform that lacks a ligand binding domain. Expression of D2-CD33 is associated with a decreased risk of developing AD. In some embodiments, the AC of the present disclosure is used to skip an exon of CD33 selected from the group consisting of exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7a, and exon 7b. In some embodiments, the exon is exon 2. In some embodiments, the AC hybridizing to its target sequence within a target CD33 pre-mRNA induces the skipping of one or more exons. In some embodiments, the AC induces expression of a re-spliced target protein comprising an inactive fragment of CD33.

[0509] In some embodiments, the disease is cancer (Laszlo et al. Oncotarget. 2016 Jul 12; 7(28): 43281–43294.). In some embodiments, the cancer is acute myeloid leukemia (AML). In some embodiments, the cancer is glioma, thyroid cancer, lung cancer, colorectal cancer, head and neck cancer, stomach cancer, liver cancer, pancreatic cancer, renal cancer, urothelial cancer, prostate cancer, testis cancer, breast cancer, cervical cancer, endometrial cancer, ovarian cancer, or melanoma. In some embodiments, the target gene is the CD33 gene. Each of the aforementioned cancers express CD33. In some embodiments, the AC of the present disclosure is used to skip an exon of CD33. In some embodiments, the exon is selected from exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7a, and exon 7b. In some embodiments, the target gene is Myc, STAT3, MDM4, ERBB4, BCL2L1, GLDC, PKM2, MCL1, MDM2, BRCA2, IL5R, FGFR1, MSTR1, USP5, or CD33.

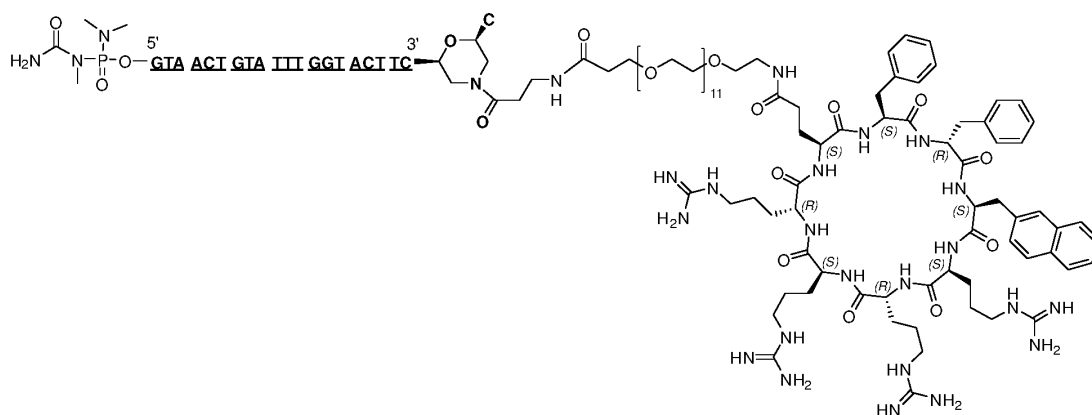
[0510] In some embodiments, provided herein are compounds comprising an AC and CPP that target the CD33 gene. Non-limiting examples of the aforementioned compounds are shown below, which can be further modified to conjugate an exocyclic peptide (EP) as described herein. The antisense oligonucleotides are underlined.

[0511] ENTR-0036:



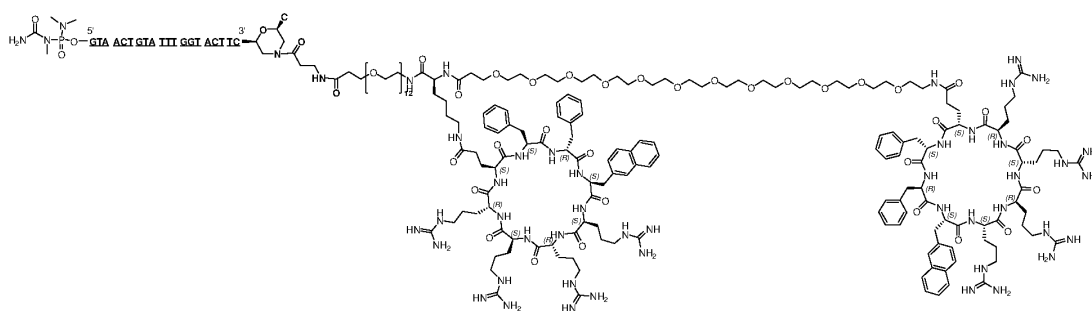
[0512]

[0513] ENTR-0081:



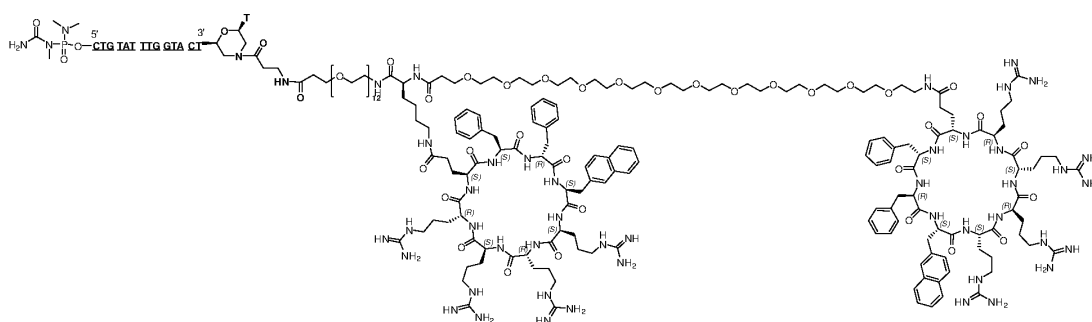
[0514]

[0515] ENTR-0087:



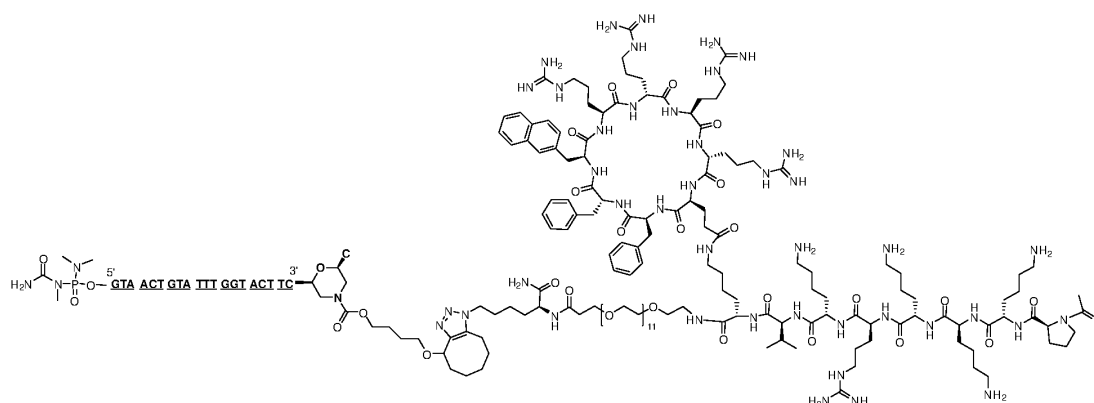
[0516]

[0517] ENTR-0085:



5 [0518]

[0519] ENTR-0179:



[0520]

[0521] In some embodiments, the disease is an inflammatory or autoimmune disease. In some embodiments, the target gene is NLRP3 or CD6.

[0522] In some embodiments, the disease is osteogenesis imperfecta (Wang and Marini, J. Clin Invest., 1996, 97, 448-454).

[0523] In some embodiments, the disease is cystic fibrosis (Friedman et al., J. Biol. Chem., 1999, 274, 36193-36199).

[0524] In some embodiments, the disease is Merosin-deficient congenital muscular dystrophy type 1A (MDC1A). MDC1A is an autosomal recessive neuromuscular disease characterized by neonatal onset of muscle weakness, hypotonia, dysmyelinating neuropathy, and minor brain abnormalities. Splice site mutations are estimated to affect ~40% of the MDC1A patient population. Causative mutations are located in the LAMA2 gene, which encodes the $\alpha 2$ chain (Lama2) of laminin-211 (or merosin) heterotrimeric protein complex expressed in the basement membrane of muscle and Schwann cells. In MDC1A, laminin-211 loses its proper interactions with receptors such as integrin $\alpha 7 \beta 1$ and dystroglycan, resulting in muscle and Schwann cells apoptosis and degeneration, which leads to fibrosis and loss of muscle function. In some embodiments, the AC hybridizes with a LAMA2 target pre-mRNA. So far, development of therapeutic strategies for MDC1A have been mainly focused on preventing fibrosis and apoptosis. The degree of LAMA2 deficiency highly correlates with the clinical severity in patients and mouse models. The lack of a functional Lama2 leads to the development of severe muscle atrophy and hind limb paralysis in mice. Therefore, restoration of LAMA2 expression holds a tremendous potential for the treatment of MDC1A. It has previously been demonstrated that muscle-specific

overexpression of Laminin-211 in merosin-deficient mice improved muscle pathology, but not the associated paralysis, indicating that correction of the peripheral neuropathy requires restoration of Lama2 beyond skeletal muscles. In some embodiments, the AC restores proper splicing to the gene.

5 **[0525]** In some embodiments, antisense compounds may be used to alter the ratio of the long and short forms of bcl-x pre-mRNA. See U.S. Pat. Nos. 6,172,216; 6,214,986; Taylor et al., Nat. Biotechnol. 1999, 17, 1097-1100, each incorporated herein by reference. An increasing number of genes and gene products have been implicated in apoptosis. One of these is bcl-2, which is an intracellular membrane protein shown to block or delay apoptosis. Overexpression of bcl-2 has
10 been shown to be related to hyperplasia, autoimmunity and resistance to apoptosis, including that induced by chemotherapy (Fang et al., J. Immunol. 1994, 153, 4388-4398). A family of bcl-2-related genes has been described. All bcl-2 family members share two highly conserved domains, BH1 and BH2. These family members include, but are not limited to, A-1, mcl-1, bax and bcl-x. Bcl-x was isolated using a bcl-2 cDNA probe at low stringency due to its sequence homology with
15 bcl-2. Bcl-x was found to function as a bcl-2-independent regulator of apoptosis (Boise et al., Cell, 1993, 74, 597-608). Two isoforms of bcl-x were reported in humans. Bcl-xl (long) contains the highly conserved BH1 and BH2 domains. When transfected into an IL-3 dependent cell line, bcl-xl inhibited apoptosis during growth factor withdrawal in a manner similar to bcl-2. In contrast, the bcl-x short isoform, bcl-xs, which is produced by alternative splicing and lacks a 63-amino
20 acid region of exon 1 containing the BH1 and BH2 domains, antagonizes the anti-apoptotic effect of either bcl-2 or bcl-xl. As numbered in Boise et al., Cell, 1993 74:, 597-608, the bcl-x transcript can be categorized into regions described by those of skill in the art as follows: nucleotides 1-134, 5' untranslated region (5'-UTR); nucleotides 135-137, translation initiation codon (AUG); nucleotides 135-836, coding region, of which 135-509 are the shorter exon 1 of the bcl-xs
25 transcript and 135-698 are the longer exon 1 of the bcl-xl transcript; nucleotides 699-836, exon 2; nucleotides 834-836, stop codon; and nucleotides 837-926, 3' untranslated region (3'-UTR). Between exons 1 and 2 (between nucleotide 698 and 699) an intron is spliced out of the pre-mRNA when the mature bcl-xl (long) mRNA transcript is produced. An alternative splice from position 509 to position 699 produces the bcl-xs (short) mRNA transcript which is 189 nucleotides shorter
30 than the long transcript, encoding a protein product (bcl-xs) which is 63 amino acids shorter than bcl-xl. Thus nucleotide position 698 is sometimes referred to in the art as the "5' splice site" and

position 509 as the “cryptic 5’ splice site,” with nucleotide 699 sometimes referred to as the “3’ splice site.” In some embodiments, the AC hybridizes with a sequence comprising the cryptic 5’ splice site of the bcl-x pre-mRNA, thereby inhibiting production of the short isoform and increasing the ratio of bcl-xl to bcl-xs isoforms.

5 [0526] In some embodiments, the AC promotes skipping of specific exons containing premature termination codons. See Wilton et al., *Neuromuscul. Disord.*, 1999, 9, 330-338, incorporated by reference herein.

[0527] In some embodiments, the AC counteracts or corrects aberrant splicing in a target pre-mRNA. See U.S. Pat. No. 5,627,274 and WO 94/26887, each of which is incorporated by reference
10 herein, and which disclose compositions and methods for combating aberrant splicing in a pre-mRNA molecule containing a mutation using antisense oligonucleotides which do not activate RNase H.

[0528] In some embodiments, the disease is proximal spinal muscular atrophy (SMA). SMA is a genetic, neurodegenerative disorder characterized by the loss of spinal motor neurons. SMA is an
15 autosomal recessive disease of early onset and is currently the leading cause of death among infants. SMA is caused by the loss of both copies of survival of motor neuron 1 (SMN1), a protein that is part of a multi-protein complex thought to be involved in snRNP biogenesis and recycling. A nearly identical gene, SMN2, exists in a duplicated region on chromosome 5q13. Although SMN1 and SMN2 have the potential to code for the same protein, SMN2 contains a translationally
20 silent mutation at position +6 of exon 7, which results in inefficient inclusion of exon 7 in SMN2 transcripts. Thus, the predominant form of SMN2 is a truncated version, lacking exon 7, which is unstable and inactive (Cartegni and Drainer, *Nat. Genet.*, 2002, 30, 377-384). In some embodiments, the AC is targeted to intron 6, exon 7 or intron 7 of SMN2. In some embodiments, the AC modulates splicing of SMN2 pre-mRNA. In some embodiments, modulation of splicing
25 results in an increase in exon 7 inclusion.

[0529] In some embodiments, the target gene is the beta globin gene. See Sierakowska et al. 1996, incorporated by reference herein. In some embodiments, the target gene is the cystic fibrosis transmembrane conductance regulator gene. See Friedman et al. 1999, incorporated by reference
30 herein. In some embodiments, the target gene is the BRCA1 gene. In some embodiments, the target gene is the eIF4E gene. In some embodiments, the target gene is a gene involved in the

pathogenesis of Duchenne muscular dystrophy, spinal muscular atrophy, or Steinert myotonic dystrophy. In some embodiments, the target gene is a DMD gene. In some embodiments, the target gene is BRCA1. In some embodiments, the target gene is a gene encoding a muscular structural protein. In some embodiments, the target gene is a gene implicated in a neuromuscular disorder (NMD). In some embodiments, the target gene is a gene implicated in cancer.

[0530] In some embodiments, the target gene is a gene that is subject to alternative splicing. In some embodiments, the present compounds and methods may be used to preferentially increase the ratio of a protein isoform by preferentially increasing the splicing of the target pre-mRNA to produce the mRNA encoding that isoform.

[0531] In some embodiments, the disease is a disease that is caused by repeat expansions of nucleotide repeat (e.g., trinucleotide repeat expansions, tetranucleotide repeat expansions, pentanucleotide repeat expansions, or hexanucleotide repeat expansions). In some embodiments, the disease is Huntington's disease, Huntington disease-like 2 (HDL2), myotonic dystrophy, spinocerebellar ataxia, spinal and bulbar muscular atrophy (SBMA), dentatorubral-pallidoluysian atrophy (DRPLA), amyotrophic lateral sclerosis, frontotemporal dementia, Fragile X syndrome, fragile X mental retardation 1 (FMR1), fragile X mental retardation 2 (FMR2), Fragile XE mental retardation (FRAXE), Friedreich's ataxia (FRDA), fragile X-associated tremor/ataxia syndrome (FXTAS), myoclonic epilepsy, oculopharyngeal muscular dystrophy (OPMD), or syndromic or non-syndromic X-linked mental retardation. In some embodiments, the disease is Huntington's disease. In some embodiments, the disease is amyotrophic lateral sclerosis. In some embodiments, the disease is a form of spinocerebellar ataxia (e.g., SCA1, SCA2, SAC3/MJD, SCA6, SCA7, SCA8, SCA10, SCA12, or SCA17).

[0532] In some embodiments, the disease is Friedreich's ataxia. In some embodiments, the target gene is *FXN*, which encodes for frataxin. In some embodiments, the compounds provided herein comprise an antisense oligonucleotide that targets *FXN*. Exemplary oligonucleotides that target *FXN* are provided in Table 9.

[0533] Table 9. Exemplary Oligonucleotides targeting FXN

Oligo chemistry	Name (oligo chemistry)	Design	Target
MOE blocker	M-4	5'-TT ^m C TT ^m C TT ^m C TT ^m C TT ^m C TT ^m C-3' (all 2'-O-MOE, all PS bonds, m=5-methyl C)	FXN (GAA) _n
MOE gapmer	Gap-0039	5'- ^m CTT ^m CTT ^m CTT ^m CTT ^m CTT ^m CTT ^m CT-3' (all PS bonds, not bold=DNA, bold=2-MOE, m=5-methyl C)	FXN (GAA) _n
MOE gapmer	Gap-0040	5'-T ^m CT T ^m CT T ^m CT T ^m CT T ^m CT T ^m CT T ^m C-3' (all PS bonds, not-bold=DNA, bold=2-MOE, m=5-methyl C)	FXN (GAA) _n
2'-F, siRNA	ENTR-siRNA-0027	ss 5'-GSASAS GAA GAA GAA GAA GASASG-3' as 5'-CSUSUC UUC UUC UUC UUC SUSUSC-3' (all 2'-F, S =PS bond)	FXN (GAA) _n
2'-F, siRNA EEV	ENTR-siRNA-0027A	ss CPP ₁₂ -NH-5'-GSASAS GAA GAA GAA GAA GASASG-3' as 5'-CSUSUC UUC UUC UUC UUC SUSUSC-3' (all 2'-F, S =PS bond)	FXN (GAA) _n
2'-F, siRNA EEV	ENTR-siRNA-0027B	ss 5'-GSASAS GAA GAA GAA GAA GASASG-3'-NH-CPP ₁₂ as 5'-CSUSUC UUC UUC UUC UUC SUSUSC-3' (all 2'-F, S =PS bond)	FXN (GAA) _n
siRNA	siGAA	ss 5'-GAAGAAGAAGAAGAAGT _d T _d -3' as 5'-CUUCUUCUUCUUCUUCUUCT _d T _d -3' (d=DNA)	FXN (GAA) _n
siRNA	Control (Ctrl)	as 5'-GCUAUACCAGCGUCGUAUT _d T _d -3' ss 5'-ATGACGACGCTGGTATAGCT _d T _d -3' (d=DNA)	Mismatched negative control
siRNA	siExon2	ss 5'-GAGUGUCUAUUUGAUGAAUT _d T _d -3' as 5'-AUUCAUCAAAUAGACACUCT _d T _d -3' (d=DNA)	FXN exon2, positive control for transfection
MOE blocker	ET4	5'- ^m C ^m CT ^m CAA AAG ^m CAG GAA UA-3' (all 2'-O-MOE, all PS bonds, m=5-methyl C)	FXN 3'- UTR

Oligo chemistry	Name (oligo chemistry)	Design	Target
MOE blocker	ET14	5'- ^m C ^m CG GGT ^m CTG ^m C ^m CG ^m C ^m C ^m C-3' (all 2'-O-MOE, all PS bonds, m=5-methyl C)	FXN 5'-UTR
PMO	ENTR-Oligo-0180	5'-CCT CAA AAG CAG GAA TA-3' (all PMO monomers)	FXN 3'-UTR
PMO	ENTR-Oligo-0181	5'-CCG GGT CTG CCG CCC-3' (all PMO monomers)	FXN 5'-UTR
PMO	ENTR-Oligo-0182	5'-CCA ACT GTC CTC AAA AGC AGG AAT A-3' (all PMO monomers)	FXN 3'-UTR
PMO	ENTR-Oligo-0183	5'-CCG GGT CTG CCG CCC GCT CCG CCC T-3' (all PMO monomers)	FXN 5'-UTR
PMO	ENTR-Oligo-0000	5'-CTT CTT CTT CTT CTT CTT CTT CTT C-3' (all PMO monomers)	FXN (GAA) _n
O-MOE	ENTR-Oligo-0002	5'-CTT CTT CTT CTT CTT CTT-3' (all 2'-O-MOE RNA monomers, all PS bonds)	FXN (GAA) _n
LNA	ENTR-Oligo-0004	5'-LnT-dC-LnT-LnT-dC-LnT-LnT-dC-LnT-LnT-dC-LnT-LnT-dC-LnT-LnT-3' (LnT=LNA T; dC=DNA C; all PS bonds)	FXN (GAA) _n
Gapmer (all PS bonds)	ENTR-Oligo-0039	5'- CTTCTTCTTCTTCTTCTTCTTCT -3' (all PS bonds, non-bold=DNA, bold=2-MOE)	FXN (GAA) _n
Gapmer (all PS bonds)	ENTR-Oligo-0040	5'- TCTTCTTCTTCTTCTTCTTCTTCT -3' (all PS bonds, non-bold=DNA, bold=2-MOE)	FXN (GAA) _n
MOE (all PS bonds)	ENTR-Oligo-0041	5'-CTT CTT CTT CTT CTT CTT-3' (all 2'-O-MOE RNA monomers, all PS bonds)	FXN (GAA) _n

[0534]

[0535] In some embodiments, the disease is a form of myotonic dystrophy (e.g., myotonic dystrophy type 1 or myotonic dystrophy type 2). In some embodiments, the target gene is the *DMPK* gene, which encodes myotonic-protein kinase. In some embodiments, the compounds

provided herein comprise an antisense oligonucleotide that targets *DMPK*. Exemplary oligonucleotides that target *DMPK* are provided in Table 10.

[0536] Table 10. Exemplary Oligonucleotides targeting DMPK

Oligo ID	Oligo name	Target	Sequence (5'-3')
ENTR-Oligo-0022	DMPK-2MOE-M+N+H (quote as DM+N+H)	DMPK	alkyne-5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, black=DNA, black=2-MOE)
ENTR-Oligo-0023	DMPK-2MOE-M+N+H-dual modification-1	DMPK	cyclooctyne-5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, black=DNA, black=2-MOE)
ENTR-Oligo-0023A	DMPK-2MOE-M+N+H-NH2-1	DMPK	5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, black=DNA, black=2-MOE)
ENTR-Oligo-0028	DMPK-cEt-M+N+H	DMPK	cyclooctyne-5'- ACAATAAATACCGAGG -3'- primary amine (all PS bonds, black=DNA, bold=(s)-cEt)
ENTR-Oligo-0029	DMPK-cEt-M+N+H-CP12	DMPK	cyclooctyne-5'- ACAATAAATACCGAGG -3'- CP12 (all PS bonds, black=DNA, bold=(s)-cEt)
ENTR-Oligo-0031	DMPK-2MOE-H-2 (quote as DH-1)	DMPK	5'- CGGAGCGGTTGTGA ACTGGC-3'- primary amine (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-Oligo-0032	DMPK-2MOE-N-2 (wrongly labeled as DM-1 in quote)	DMPK	alkyne-5'- CGGAGCGGTTGTGA ACTGGCA-3'- primary amine (20 bases, all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-Oligo-0053	DMPK-2MOE-M+N+H	DMPK	5'- ACAGACAATAAATACCGAGG -3' (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-oligo-0077	ENTR-oligo-0022-PEG12-CPP12-Amide	DMPK	alkyne-5'- ACAGACAATAAATACCGAGG -3'- PEG12-CPP12 (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-oligo-0078	CPP12-PEG12-click-ENTR-oligo-0022	DMPK	CPP12-PEG12-Lys-click-5'- ACAGACAATAAATACCGAGG -3' (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-oligo-0189	ETRDWUXI-617+ENTR-oligo-0023 (click)	DMPK	CPP12-PEG12-click-5'- ACAGACAATAAATACCGAGG -3'-

Oligo ID	Oligo name	Target	Sequence (5'-3')
			primary amine (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-oligo-0190	ETRDWUXI-642+ ENTR-oligo-0023 (click)	DMPK	CPP12-K(CPP12) PEG12-K-click-5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-oligo-0191	ETRDWUXI-684+ ENTR-oligo-0023 (click)	DMPK	Ac-NLS-Lys(CPP12)-PEG12-K-click-5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, non-bold=DNA, bold=2-MOE) (
ENTR-oligo-0192	ENTR-oligo-0023a+SMCC	DMPK	5'- ACAGACAATAAATACCGAGG -3'- primary amine+SMCC (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-Oligo-0071	PMO ^{CAG}	DMPK	5'-CAG CAG CAG CAG CAG CAG CAG-3'-NH ₂ (all PMO monomers)
ENTR-Oligo-0034	PMO ^{CAG} -CP12	DMPK	5'-CAG CAG CAG CAG CAG CAG CAG-3'-CP12 (all PMO monomers)
ENTR-Oligo-0022	DMPK-2MOE-M+N+H (quote as DM+N+H)	DMPK	alkyne-5'- ACAGACAATAAATACCGAGG -3'- primary amine (all PS bonds, non-bold=DNA, bold=2-MOE)
ENTR-Oligo-0031	DMPK-2MOE-H-2 (quote as DH-1)	DMPK	5'- CGGAGCGGTTGTGA ACTGGC-3'- primary amine (all PS bonds, non-bold=DNA, bold=2-MOE)

[0537]

[0538] In some embodiments, the disease is Dravet syndrome. Dravet syndrome is a severe and progressive genetic epilepsy. Dravet syndrome is an autosomal dominant condition caused by more than 1250 *de novo* mutations in *SCN1A*, resulting in 50 % NaV1.1 protein expression. Dravet syndrome is caused by pathogenic mutation or deletion of the *SCN1A* gene in 85 % of patients. Existing antiepileptic drug only address the occurrence of seizures, and more than 90 % of Dravet syndrome patients still report suffering from incomplete seizure control. In some embodiments, the antisense oligonucleotide targets *SCN1A*. In some embodiments, the antisense oligonucleotide targeting *SCN1A* has a sequence of 5'-CCATAATAAAGGGCTCAG-3'. In some embodiments, the efficacy of antisense compounds targeting *SCN1A* is evaluated in a mouse model. Non-limiting examples of mouse models include mouse models with a targeted deletion of *SCN1A* exon 1 (Scn1a^{tm1Kex}) and exon 26 (Scn1a^{tm1Wac}), mouse models with specific point mutation knock-ins,

such as *Scn1a* R1407X, *Scn1a* R1648H, and *Scn1a* E1099X, and a transgenic mouse model expressing a bacterial artificial chromosome (BAC) with a human *SCN1A* R1648H mutation. In some embodiments, the efficacy of antisense compounds targeting *SCN1A* is evaluated in an in vitro model, for example, in wild-type fibroblasts.

5 [0539] In some embodiments, the disease is Fragile X Syndrome (FXS). FXS is the most common form of inherited intellectual and developmental disease. FXS is caused by silenced expression of fragile X mental retardation protein (FMRP) due to the presence of > 200 CGG trinucleotide repeats in *FMRI* which encodes for FMRP. FMRP is encoded by *FMRI*. In some embodiments, an antisense compound of the disclosure targets *FMRI*. In some embodiments, the efficacy of
10 antisense compounds targeting *FMRI* is evaluated in a mouse model (e.g., those described in Dahlhaus et al.), which is incorporated by reference herein in its entirety: Dahlhaus, R. (2018). Of men and mice: modeling the fragile X syndrome. *Frontiers in molecular neuroscience*, 11, 41.

[0540] In some embodiments, the disease is Fragile X tremor ataxia syndrome (FXTAS). FXTAS is a late-onset, progressive neurodegenerative disorder characterized by cerebellar ataxia and
15 intention tremor. FXTAS is caused by an *FMRI* premutation, which is defined as having 55 to 200 CGG repeats in the 5' untranslated region of *FMRI*. In some embodiments, an antisense compound of the disclosure targets *FMRI*.

[0541] In some embodiments, the disease is Huntington's Disease (HD). HD is an autosomal dominant disease, characterized by cognitive decline, psychiatric illness, and chorea. HD is often
20 fatal. HD is caused by an expanded CAG triplet repeat in the *HTT* gene, which results in the production of mutant huntingtin protein (mHTT). Accumulation of mHTT causes progressive loss of neurons in the brain. In some embodiments, the target gene is *HTT*. In some embodiments, an antisense compound of the disclosure targets *HTT*. In some embodiments, the efficacy of antisense compounds and/or oligonucleotides are evaluated in in vivo models. Exemplary models are
25 described in Pouladi et al. which is incorporated by reference herein in its entirety: Pouladi, Mahmoud A., et al. "Choosing an animal model for the study of Huntington's disease." *Nature Reviews Neuroscience* 14.10 (2013): 708-721. In some embodiments, the antisense oligonucleotide is non-allele selective. In some embodiments, the non-allele selective antisense oligonucleotide is an HTTRx gapmer (Ionis) or a divalent siRNA (UMass). In some embodiments,
30 the antisense oligonucleotide is allele selective. In some embodiments, the allele selective

antisense oligonucleotide is a stereopure gapmer targeting a single nucleotide polymorphism in *HTT*. In some embodiments, the antisense oligonucleotide targets exon 1, exon 30, exon 36, exon 50, or exon 67 of *HTT*. The following references describe exemplary antisense oligonucleotides and are incorporated herein by reference in their entirety: Yu, Dongbo, et al. *Cell* 150.5 (2012): 895-908; Alterman, Julia F., et al. *Nature biotechnology* 37.8 (2019): 884-894. Tabrizi, Sarah J., et al. *New England Journal of Medicine* 380.24 (2019): 2307-2316.; Kordasiewicz, Holly B., et al. *Neuron* 74.6 (2012): 1031-1044.

[0542] In some embodiments, the disease is Wilson's Disease (WD). WD is a recessive fatal copper homeostasis disorder, typically diagnosed in patients between the ages of 5 and 35, leading to hepatic and neurologic symptoms due to free copper accumulation. WD is caused by loss-of-function mutations in the *ATP7B* gene. *ATP7B* encodes copper-transporting ATPase 2, which is a transmembrane copper transporter and responsible in the transport of copper from the liver to other parts of the body. In some embodiments, provided herein is an antisense oligonucleotide or compound thereof that targets *ATP7B*. In some embodiments, the antisense oligonucleotide or compound thereof targets a T1934G (or Met-645-Arg) mutation in *ATP7B*. The aforementioned *ATP7B* variant is described in Merico et al. which is incorporated by reference herein in its entirety: Merico, Daniele, et al. *NPJ Genomic Medicine* 5.1 (2020): 1-7. In some embodiments, the antisense oligonucleotide has a sequence of 5'-CAGCTGGAGTTTATCTTTTG-3'.

[0543] In some embodiments of this aspect, the sequence of the corresponding gene underlying such diseases is prone to forming clusters of RNA comprises tandem nucleotide repeats (e.g., multiple nucleotide repeats comprising at least 10, 15, 20, 25, 30, 40, 50, 60, 70 or more adjacent repeated nucleotide sequences). In some embodiments, the tandem nucleotide repeats are trinucleotide repeats. The trinucleotide repeat sequences may be CAG repeats, CGG repeats, GCC repeats, GAA repeats, or CUG repeats. In some embodiments, the trinucleotide repeat is a CAG repeat. In some embodiments, the RNA sequence comprises at least 10 trinucleotide repeats (e.g., CAG, CGG, GCC, GAA, or CUG repeats), e.g., at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, or at least 70 trinucleotide repeats. In some embodiments, the target gene is selected from the group consisting of FMR1, AFF2, FXN, DMPK, SCA8, PPP2R2B, ATN1, DRPLA, HTT, AR, ATXN1, ATXN2, ATXN3, CACNA1A, ATXN7, TBP. See U.S. Pat. Appl. Publ. No. 2016/0355796 and U.S. Pat. Appl. Publ. No.

2018/0344817, each of which is incorporated by reference herein, and which discloses diseases and corresponding genes prone to forming and/or expanding tandem nucleotide repeats.

[0544] In some embodiments, an AC of the disclosure is administered to treat any disease described by the disclosure, for example, Huntington's disease, Huntington disease-like 2 (HDL2),
5 myotonic dystrophy, spinocerebellar ataxia, spinal and bulbar muscular atrophy (SBMA), dentatorubral-pallidoluysian atrophy (DRPLA), amyotrophic lateral sclerosis, frontotemporal dementia, Fragile X syndrome, fragile X mental retardation 1 (FMR1), fragile X mental retardation 2 (FMR2), Fragile XE mental retardation (FRAXE), Friedreich's ataxia (FRDA), fragile X-associated tremor/ataxia syndrome (FXTAS), myoclonic epilepsy, oculopharyngeal muscular
10 dystrophy (OPMD), syndromic or non-syndromic X-linked mental retardation, Cystic fibrosis, proximal spinal muscular atrophy, of Duchenne muscular dystrophy, spinal muscular atrophy, Steinert myotonic dystrophy, Merosin-deficient congenital muscular dystrophy type 1A, osteogenesis imperfect, cancer, glioma, thyroid cancer, lung cancer, colorectal cancer, head and neck cancer, stomach cancer, liver cancer, pancreatic cancer, renal cancer, urothelial cancer,
15 prostate cancer, testis cancer, breast cancer, cervical cancer, endometrial cancer, ovarian cancer, melanoma, or Alzheimer's Disease.

[0545] In some embodiments, an AC of the disclosure is a gapmer oligonucleotide as disclose in U.S. Patent No. 9,550,988, the disclosure of which is incorporated by reference herein.

[0546] In some embodiments, an AC of the disclosure comprises the sequence and/or structure of
20 any one of the ACs targeting SMN2 disclosed in U.S. Patent No. 8,361,977, the disclosure of which is incorporated by reference herein.

[0547] In some embodiments, an AC of the disclosure comprises the sequence and/or structure of any one of the ACs targeting DMD, SMN2, or DMPK disclosed in U.S. Patent Publication No. 2017/0260524, the disclosure of which is incorporated by reference herein.

[0548] In some embodiments, an AC of the disclosure comprises the sequence and/or structure of
25 any one of the ACs or oligonucleotides disclosed in U.S. Patent Publications US20030235845A1, US20060099616A1, US 2013/0072671 A1, US 2014/0275212 A1, US 2009/0312532 A1, US20100125099A1, US 2010/0125099 A1, US 2009/0269755 A1, US 2011/0294753 A1, US 2012/0022134 A1, US 2011/0263682 A1, US 2014/0128592 A1, US 2015/0073037 A1, and

US20120059042A1, the contents of each of which are incorporated herein in their entirety for all purposes.

Compositions, Formulations and Methods of Administration

5 [0549] *In vivo* application of the disclosed compounds, and compositions containing them, can be accomplished by any suitable method and technique presently or prospectively known to those skilled in the art. For example, the disclosed compounds can be formulated in a physiologically- or pharmaceutically-acceptable form and administered by any suitable route known in the art including, for example, oral, nasal, rectal, topical, and parenteral routes of administration. As used
10 herein, the term parenteral includes subcutaneous, intradermal, intravenous, intramuscular, intraperitoneal, and intrasternal administration, such as by injection. Administration of the disclosed compounds or compositions can be a single administration, or at continuous or distinct intervals as can be readily determined by a person skilled in the art.

[0550] The compounds disclosed herein, and compositions comprising them, can also be
15 administered utilizing liposome technology, slow release capsules, implantable pumps, and biodegradable containers. These delivery methods can, advantageously, provide a uniform dosage over an extended period of time. The compounds can also be administered in their salt derivative forms or crystalline forms.

[0551] The compounds disclosed herein can be formulated according to known methods for
20 preparing pharmaceutically acceptable compositions. Formulations are described in detail in a number of sources which are well known and readily available to those skilled in the art. For example, *Remington's Pharmaceutical Science* by E.W. Martin (1995) describes formulations that can be used in connection with the disclosed methods. In general, the compounds disclosed herein can be formulated such that an effective amount of the compound is combined with a suitable
25 carrier in order to facilitate effective administration of the compound. The compositions used can also be in a variety of forms. These include, for example, solid, semi-solid, and liquid dosage forms, such as tablets, pills, powders, liquid solutions or suspension, suppositories, injectable and infusible solutions, and sprays. The preferred form depends on the intended mode of administration and therapeutic application. The compositions also preferably include conventional
30 pharmaceutically-acceptable carriers and diluents which are known to those skilled in the art. Examples of carriers or diluents for use with the compounds include ethanol, dimethyl sulfoxide,

glycerol, alumina, starch, saline, and equivalent carriers and diluents. To provide for the administration of such dosages for the desired therapeutic treatment, compositions disclosed herein can advantageously comprise between about 0.1% and 100% by weight of the total of one or more of the subject compounds based on the weight of the total composition including carrier or diluent.

5 [0552] Formulations suitable for administration include, for example, aqueous sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient; and aqueous and nonaqueous sterile suspensions, which can include suspending agents and thickening agents. The formulations can be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and can
10 be stored in a freeze dried (lyophilized) condition requiring only the condition of the sterile liquid carrier, for example, water for injections, prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powder, granules, tablets, *etc.* It should be understood that in addition to the ingredients particularly mentioned above, the compositions disclosed herein can include other agents conventional in the art having regard to the type of formulation in
15 question.

[0553] Compounds disclosed herein, and compositions comprising them, can be delivered to a cell either through direct contact with the cell or via a carrier means. Carrier means for delivering compounds and compositions to cells are known in the art and include, for example, encapsulating the composition in a liposome moiety. Another means for delivery of compounds and
20 compositions disclosed herein to a cell can comprise attaching the compounds to a protein or nucleic acid that is targeted for delivery to the target cell. U.S. Patent No. 6,960,648 and U.S. Application Publication Nos. 20030032594 and 20020120100 disclose amino acid sequences that can be coupled to another composition and that allows the composition to be translocated across biological membranes. U.S. Application Publication No. 20020035243 also describes
25 compositions for transporting biological moieties across cell membranes for intracellular delivery. Compounds can also be incorporated into polymers, examples of which include poly (D-L lactide-co-glycolide) polymer for intracranial tumors; poly[bis(p-carboxyphenoxy) propane:sebacic acid] in a 20:80 molar ratio (as used in GLIADEL); chondroitin; chitin; and chitosan.

[0554] For the treatment of oncological disorders, the compounds disclosed herein can be
30 administered to a patient in need of treatment in combination with other antitumor or anticancer substances and/or with radiation and/or photodynamic therapy and/or with surgical treatment to

remove a tumor. These other substances or treatments can be given at the same as or at different times from the compounds disclosed herein. For example, the compounds disclosed herein can be used in combination with mitotic inhibitors such as taxol or vinblastine, alkylating agents such as cyclophosphamide or ifosfamide, antimetabolites such as 5-fluorouracil or hydroxyurea, DNA intercalators such as adriamycin or bleomycin, topoisomerase inhibitors such as etoposide or camptothecin, antiangiogenic agents such as angiostatin, antiestrogens such as tamoxifen, and/or other anti-cancer drugs or antibodies, such as, for example, GLEEVEC (Novartis Pharmaceuticals Corporation) and HERCEPTIN (Genentech, Inc.), respectively, or an immunotherapeutic such as ipilimumab and bortezomib.

10 [0555] In certain examples, compounds and compositions disclosed herein can be locally administered at one or more anatomical sites, such as sites of unwanted cell growth (such as a tumor site or benign skin growth, *e.g.*, injected or topically applied to the tumor or skin growth), optionally in combination with a pharmaceutically acceptable carrier such as an inert diluent. Compounds and compositions disclosed herein can be systemically administered, such as
15 intravenously or orally, optionally in combination with a pharmaceutically acceptable carrier such as an inert diluent, or an assimilable edible carrier for oral delivery. They can be enclosed in hard or soft shell gelatin capsules, can be compressed into tablets, or can be incorporated directly with the food of the patient's diet. For oral therapeutic administration, the active compound can be combined with one or more excipients and used in the form of ingestible tablets, buccal tablets,
20 troches, capsules, elixirs, suspensions, syrups, wafers, aerosol sprays, and the like.

[0556] The disclosed compositions are bioavailable and can be delivered orally. Oral compositions can be tablets, troches, pills, capsules, and the like, and can also contain the following: binders such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as
25 magnesium stearate; and a sweetening agent such as sucrose, fructose, lactose or aspartame or a flavoring agent such as peppermint, oil of wintergreen, or cherry flavoring can be added. When the unit dosage form is a capsule, it can contain, in addition to materials of the above type, a liquid carrier, such as a vegetable oil or a polyethylene glycol. Various other materials can be present as coatings or to otherwise modify the physical form of the solid unit dosage form. For instance,
30 tablets, pills, or capsules can be coated with gelatin, wax, shellac, or sugar and the like. A syrup or elixir can contain the active compound, sucrose or fructose as a sweetening agent, methyl and

propylparabens as preservatives, a dye and flavoring such as cherry or orange flavor. Of course, any material used in preparing any unit dosage form should be pharmaceutically acceptable and substantially non-toxic in the amounts employed. In addition, the active compound can be incorporated into sustained-release preparations and devices.

5 [0557] Compounds and compositions disclosed herein, including pharmaceutically acceptable salts or prodrugs thereof, can be administered intravenously, intramuscularly, or intraperitoneally by infusion or injection. Solutions of the active agent or its salts can be prepared in water, optionally mixed with a nontoxic surfactant. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, triacetin, and mixtures thereof and in oils. Under ordinary conditions of
10 storage and use, these preparations can contain a preservative to prevent the growth of microorganisms.

[0558] The pharmaceutical dosage forms suitable for injection or infusion can include sterile aqueous solutions or dispersions or sterile powders comprising the active ingredient, which are adapted for the extemporaneous preparation of sterile injectable or infusible solutions or
15 dispersions, optionally encapsulated in liposomes. The ultimate dosage form should be sterile, fluid and stable under the conditions of manufacture and storage. The liquid carrier or vehicle can be a solvent or liquid dispersion medium comprising, for example, water, ethanol, a polyol (for example, glycerol, propylene glycol, liquid polyethylene glycols, and the like), vegetable oils, nontoxic glyceryl esters, and suitable mixtures thereof. The proper fluidity can be maintained, for
20 example, by the formation of liposomes, by the maintenance of the required particle size in the case of dispersions or by the use of surfactants. Optionally, the prevention of the action of microorganisms can be brought about by various other antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, buffers or sodium chloride.
25 Prolonged absorption of the injectable compositions can be brought about by the inclusion of agents that delay absorption, for example, aluminum monostearate and gelatin.

[0559] Sterile injectable solutions are prepared by incorporating a compound and/or agent disclosed herein in the required amount in the appropriate solvent with various other ingredients enumerated above, as required, followed by filter sterilization. In the case of sterile powders for
30 the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum

drying and the freeze drying techniques, which yield a powder of the active ingredient plus any additional desired ingredient present in the previously sterile-filtered solutions.

[0560] For topical administration, compounds and agents disclosed herein can be applied in as a liquid or solid. However, it will generally be desirable to administer them topically to the skin as compositions, in combination with a dermatologically acceptable carrier, which can be a solid or a liquid. Compounds and agents and compositions disclosed herein can be applied topically to a subject's skin to reduce the size (and can include complete removal) of malignant or benign growths, or to treat an infection site. Compounds and agents disclosed herein can be applied directly to the growth or infection site. Preferably, the compounds and agents are applied to the growth or infection site in a formulation such as an ointment, cream, lotion, solution, tincture, or the like.

[0561] Useful solid carriers include finely divided solids such as talc, clay, microcrystalline cellulose, silica, alumina and the like. Useful liquid carriers include water, alcohols or glycols or water-alcohol/glycol blends, in which the compounds can be dissolved or dispersed at effective levels, optionally with the aid of non-toxic surfactants. Adjuvants such as fragrances and additional antimicrobial agents can be added to optimize the properties for a given use. The resultant liquid compositions can be applied from absorbent pads, used to impregnate bandages and other dressings, or sprayed onto the affected area using pump-type or aerosol sprayers, for example.

[0562] Thickeners such as synthetic polymers, fatty acids, fatty acid salts and esters, fatty alcohols, modified celluloses or modified mineral materials can also be employed with liquid carriers to form spreadable pastes, gels, ointments, soaps, and the like, for application directly to the skin of the user.

[0563] Useful dosages of the compounds and agents and pharmaceutical compositions disclosed herein can be determined by comparing their *in vitro* activity, and *in vivo* activity in animal models. Methods for the extrapolation of effective dosages in mice, and other animals, to humans are known to the art.

[0564] The dosage ranges for the administration of the compositions are those large enough to produce the desired effect in which the symptoms or disorder are affected. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of the disease in the patient and can be determined by one of skill in the art. The dosage can be

adjusted by the individual physician in the event of any counterindications. Dosage can vary, and can be administered in one or more dose administrations daily, for one or several days.

[0565] Also disclosed are pharmaceutical compositions that comprise a compound disclosed herein in combination with a pharmaceutically acceptable carrier. Pharmaceutical compositions adapted for oral, topical or parenteral administration, comprising an amount of a compound constitute a preferred aspect. The dose administered to a patient, particularly a human, should be sufficient to achieve a therapeutic response in the patient over a reasonable time frame, without lethal toxicity, and preferably causing no more than an acceptable level of side effects or morbidity. One skilled in the art will recognize that dosage will depend upon a variety of factors including the condition (health) of the subject, the body weight of the subject, kind of concurrent treatment, if any, frequency of treatment, therapeutic ratio, as well as the severity and stage of the pathological condition.

[0566] Also disclosed are kits that comprise a compound disclosed herein in one or more containers. The disclosed kits can optionally include pharmaceutically acceptable carriers and/or diluents. A kit can include one or more other components, adjuncts, or adjuvants as described herein. kit includes one or more anti-cancer agents, such as those agents described herein. A kit can include instructions or packaging materials that describe how to administer a compound or composition of the kit. Containers of the kit can be of any suitable material, *e.g.*, glass, plastic, metal, *etc.*, and of any suitable size, shape, or configuration. A compound and/or agent disclosed herein can be provided in the kit as a solid, such as a tablet, pill, or powder form. A compound and/or agent disclosed herein can be provided in the kit as a liquid or solution. A kit can comprise an ampoule or syringe containing a compound and/or agent disclosed herein in liquid or solution form.

Certain Definitions

[0567] As used in the description and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “a composition” includes mixtures of two or more such compositions, reference to “an agent” includes mixtures of two or more such agents, reference to “the component” includes mixtures of two or more such components, and the like.

[0568] The term “about” when immediately preceding a numerical value means a range (e.g., plus or minus 10% of that value). For example, “about 50” can mean 45 to 55, “about 25,000” can mean 22,500 to 27,500, etc., unless the context of the disclosure indicates otherwise, or is inconsistent with such an interpretation. For example, in a list of numerical values such as “about 49, about 50, about 55, ...”, “about 50” means a range extending to less than half the interval(s) between the preceding and subsequent values, e.g., more than 49.5 to less than 52.5. Furthermore, the phrases “less than about” a value or “greater than about” a value should be understood in view of the definition of the term “about” provided herein. Similarly, the term “about” when preceding a series of numerical values or a range of values (e.g., “about 10, 20, 30” or “about 10-30”) refers, respectively to all values in the series, or the endpoints of the range.

[0569] As used herein, the term “cyclic cell penetrating peptide” or “cCPP” refers to a peptide that facilitates the delivery of a cargo, e.g., a therapeutic moiety, into a cell.

[0570] As used herein, the term “endosomal escape vehicle” (EEV) refers to a cCPP that is conjugated by a chemical linkage (i.e., a covalent bond or non-covalent interaction) to a linker and/or an exocyclic peptide (EP). The EEV can be an EEV of Formula (B).

[0571] As used herein, the term “EEV-conjugate” refers to an endosomal escape vehicle defined herein conjugated by a chemical linkage (i.e., a covalent bond or non-covalent interaction) to a cargo. The cargo can be a therapeutic moiety (e.g., an oligonucleotide, peptide or small molecule) that can be delivered into a cell by the EEV. The EEV-conjugate can be an EEV-conjugate of Formula (C).

[0572] As used herein, the term “exocyclic peptide” (EP) and “modulatory peptide” (MP) may be used interchangeably to refer to two or more amino acid residues linked by a peptide bond that can be conjugated to a cyclic cell penetrating peptide (cCPP) disclosed herein. The EP, when conjugated to a cyclic peptide disclosed herein, may alter the tissue distribution and/or retention of the compound. Typically, the EP comprises at least one positively charged amino acid residue,

e.g., at least one lysine residue and/or at least one arginine residue. Non-limiting examples of EP are described herein. The EP can be a peptide that has been identified in the art as a “nuclear localization sequence” (NLS). Non-limiting examples of nuclear localization sequences include the nuclear localization sequence of the SV40 virus large T-antigen, the minimal functional unit of which is the seven amino acid sequence PKKKRKV, the nucleoplasmin bipartite NLS with the sequence NLSKRPAAIKKAGQAKKKK, the c-myc nuclear localization sequence having the amino acid sequence PAAKRVKLD or RQRRNELKRSF, the sequence RMRKFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV of the IBB domain from importin-alpha, the sequences VSRKRPRP and PPKKARED of the myoma T protein, the sequence PQPKKKPL of human p53, the sequence SALIKKKKKMAP of mouse c-abl IV, the sequences DRLRR and PKQKKRK of the influenza virus NS1, the sequence RKLKKKIKKL of the Hepatitis virus delta antigen and the sequence REKKKFLKRR of the mouse Mx1 protein, the sequence KRKGDEV DGVDEVAKKKSKK of the human poly(ADP-ribose) polymerase and the sequence RKCLQAGMNLEARKTKK of the steroid hormone receptors (human) glucocorticoid.

International Publication No. 2001/038547 describes additional examples of NLSs and is incorporated by reference herein in its entirety.

[0573] As used herein, “linker” or “L” refers to a moiety that covalently bonds one or more moieties (e.g., an exocyclic peptide (EP) and a cargo, e.g., an oligonucleotide, peptide or small molecule) to the cyclic cell penetrating peptide (cCPP). The linker can comprise a natural or non-natural amino acid or polypeptide. The linker can be a synthetic compound containing two or more appropriate functional groups suitable to bind the cCPP to a cargo moiety, to thereby form the compounds disclosed herein. The linker can comprise a polyethylene glycol (PEG) moiety. The linker can comprise one or more amino acids. The cCPP may be covalently bound to a cargo via a linker.

[0574] As used herein, the term “oligonucleotide” refers to an oligomeric compound comprising a plurality of linked nucleotides or nucleosides. One or more nucleotides of an oligonucleotide can be modified. An oligonucleotide can comprise ribonucleic acid (RNA) or deoxyribonucleic acid (DNA). Oligonucleotides can be composed of natural and/or modified nucleobases, sugars and covalent internucleoside linkages, and can further include non-nucleic acid conjugates.

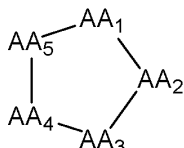
[0575] The terms “peptide,” “protein,” and “polypeptide” are used interchangeably to refer to a natural or synthetic molecule comprising two or more amino acids linked by the carboxyl group

of one amino acid to the alpha amino group of another. Two or more amino acid residues can be linked by the carboxyl group of one amino acid to the alpha amino group. Two or more amino acids of the polypeptide can be joined by a peptide bond. The polypeptide can include a peptide backbone modification in which two or more amino acids are covalently attached by a bond other than a peptide bond. The polypeptide can include one or more non-natural amino acids, amino acid analogs, or other synthetic molecules that are capable of integrating into a polypeptide. The term polypeptide includes naturally occurring and artificially occurring amino acids. The term polypeptide includes peptides, for example, that include from about 2 to about 100 amino acid residues as well as proteins, that include more than about 100 amino acid residues, or more than about 1000 amino acid residues, including, but not limited to therapeutic proteins such as antibodies, enzymes, receptors, soluble proteins and the like.

[0576] The term “therapeutic polypeptide” refers to a polypeptide that has therapeutic, prophylactic or other biological activity. The therapeutic polypeptide can be produced in any suitable manner. For example, the therapeutic polypeptide may isolated or purified from a naturally occurring environment, may be chemically synthesized, may be recombinantly produced, or a combination thereof.

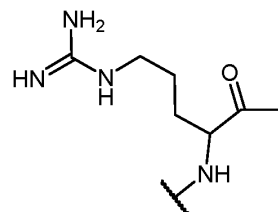
[0577] The term “small molecule” refers to an organic compound with pharmacological activity and a molecular weight of less than about 2000 Daltons, or less than about 1000 Daltons, or less than about 500 Daltons. Small molecule therapeutics are typically manufactured by chemical synthesis.

[0578] As used herein, the term “contiguous” refers to two amino acids, which are connected by a covalent bond. For example, in the context of a representative cyclic cell penetrating peptide

(cCPP) such as , AA₁/AA₂, AA₂/AA₃, AA₃/AA₄, and AA₅/AA₁ exemplify pairs of contiguous amino acids.

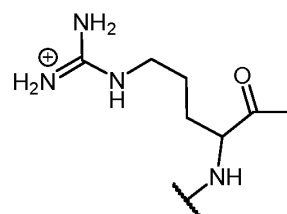
[0579] A residue of a chemical species, as used herein, refers to a derivative of the chemical species that is present in a particular product. To form the product, at least one atom of the species is replaced by a bond to another moiety, such that the product contains a derivative, or residue, of the chemical species. For example, the cyclic cell penetrating peptides (cCPP) described herein

have amino acids (e.g., arginine) incorporated therein through formation of one or more peptide bonds. The amino acids incorporated into the cCPP may be referred to residues, or simply as an



amino acid. Thus, arginine or an arginine residue refers to

[0580] The term “protonated form thereof” refers to a protonated form of an amino acid. For example, the guanidine group on the side chain of arginine may be protonated to form a



guanidinium group. The structure of a protonated form of arginine is

[0581] As used herein, the term “chirality” refers to a molecule that has more than one stereoisomer that differs in the three-dimensional spatial arrangement of atoms, in which one stereoisomer is a non-superimposable mirror image of the other. Amino acids, except for glycine, have a chiral carbon atom adjacent to the carboxyl group. The term “enantiomer” refers to stereoisomers that are chiral. The chiral molecule can be an amino acid residue having a “D” and “L” enantiomer. Molecules without a chiral center, such as glycine, can be referred to as “achiral.”

[0582] As used herein, the term “hydrophobic” refers to a moiety that is not soluble in water or has minimal solubility in water. Generally, neutral moieties and/or non-polar moieties, or moieties that are predominately neutral and/or non-polar are hydrophobic. Hydrophobicity can be measured by one of the methods disclosed herein below.

[0583] As used herein “aromatic” refers to an unsaturated cyclic molecule having $4n + 2 \pi$ electrons, wherein n is any integer. The term “non-aromatic” refers to any unsaturated cyclic molecule which does not fall within the definition of aromatic.

[0584] “Alkyl”, “alkyl chain” or “alkyl group” refer to a fully saturated, straight or branched hydrocarbon chain radical having from one to forty carbon atoms, and which is attached to the rest of the molecule by a single bond. Alkyls comprising any number of carbon atoms from 1 to 40 are

included. An alkyl comprising up to 40 carbon atoms is a C₁-C₄₀ alkyl, an alkyl comprising up to 10 carbon atoms is a C₁-C₁₀ alkyl, an alkyl comprising up to 6 carbon atoms is a C₁-C₆ alkyl and an alkyl comprising up to 5 carbon atoms is a C₁-C₅ alkyl. A C₁-C₅ alkyl includes C₅ alkyls, C₄ alkyls, C₃ alkyls, C₂ alkyls and C₁ alkyl (*i.e.*, methyl). A C₁-C₆ alkyl includes all moieties described above for C₁-C₅ alkyls but also includes C₆ alkyls. A C₁-C₁₀ alkyl includes all moieties described above for C₁-C₅ alkyls and C₁-C₆ alkyls, but also includes C₇, C₈, C₉ and C₁₀ alkyls. Similarly, a C₁-C₁₂ alkyl includes all the foregoing moieties, but also includes C₁₁ and C₁₂ alkyls. Non-limiting examples of C₁-C₁₂ alkyl include methyl, ethyl, *n*-propyl, *i*-propyl, *sec*-propyl, *n*-butyl, *i*-butyl, *sec*-butyl, *t*-butyl, *n*-pentyl, *t*-amyl, *n*-hexyl, *n*-heptyl, *n*-octyl, *n*-nonyl, *n*-decyl, *n*-undecyl, and *n*-dodecyl. Unless stated otherwise specifically in the specification, an alkyl group can be optionally substituted.

[0585] “Alkylene”, “alkylene chain” or “alkylene group” refers to a fully saturated, straight or branched divalent hydrocarbon chain radical, having from one to forty carbon atoms. Non-limiting examples of C₂-C₄₀ alkylene include ethylene, propylene, *n*-butylene, ethenylene, propenylene, *n*-butenylene, propynylene, *n*-butynylene, and the like. Unless stated otherwise specifically in the specification, an alkylene chain can be optionally substituted.

[0586] “Alkenyl”, “alkenyl chain” or “alkenyl group” refers to a straight or branched hydrocarbon chain radical having from two to forty carbon atoms and having one or more carbon-carbon double bonds. Each alkenyl group is attached to the rest of the molecule by a single bond. Alkenyl groups comprising any number of carbon atoms from 2 to 40 are included. An alkenyl group comprising up to 40 carbon atoms is a C₂-C₄₀ alkenyl, an alkenyl comprising up to 10 carbon atoms is a C₂-C₁₀ alkenyl, an alkenyl group comprising up to 6 carbon atoms is a C₂-C₆ alkenyl and an alkenyl comprising up to 5 carbon atoms is a C₂-C₅ alkenyl. A C₂-C₅ alkenyl includes C₅ alkenyls, C₄ alkenyls, C₃ alkenyls, and C₂ alkenyls. A C₂-C₆ alkenyl includes all moieties described above for C₂-C₅ alkenyls but also includes C₆ alkenyls. A C₂-C₁₀ alkenyl includes all moieties described above for C₂-C₅ alkenyls and C₂-C₆ alkenyls, but also includes C₇, C₈, C₉ and C₁₀ alkenyls. Similarly, a C₂-C₁₂ alkenyl includes all the foregoing moieties, but also includes C₁₁ and C₁₂ alkenyls. Non-limiting examples of C₂-C₁₂ alkenyl include ethenyl (vinyl), 1-propenyl, 2-propenyl (allyl), iso-propenyl, 2-methyl-1-propenyl, 1-butenyl, 2-butenyl, 3-butenyl, 1-pentenyl, 2-pentenyl, 3-pentenyl, 4-pentenyl, 1-hexenyl, 2-hexenyl, 3-hexenyl, 4-hexenyl, 5-hexenyl, 1-heptenyl, 2-heptenyl, 3-heptenyl, 4-heptenyl, 5-heptenyl, 6-heptenyl, 1-octenyl, 2-octenyl, 3-

octenyl, 4-octenyl, 5-octenyl, 6-octenyl, 7-octenyl, 1-nonenyl, 2-nonenyl, 3-nonenyl, 4-nonenyl, 5-nonenyl, 6-nonenyl, 7-nonenyl, 8-nonenyl, 1-decenyl, 2-decenyl, 3-decenyl, 4-decenyl, 5-decenyl, 6-decenyl, 7-decenyl, 8-decenyl, 9-decenyl, 1-undecenyl, 2-undecenyl, 3-undecenyl, 4-undecenyl, 5-undecenyl, 6-undecenyl, 7-undecenyl, 8-undecenyl, 9-undecenyl, 10-undecenyl, 1-dodecenyl, 2-dodecenyl, 3-dodecenyl, 4-dodecenyl, 5-dodecenyl, 6-dodecenyl, 7-dodecenyl, 8-dodecenyl, 9-dodecenyl, 10-dodecenyl, and 11-dodecenyl. Unless stated otherwise specifically in the specification, an alkyl group can be optionally substituted.

[0587] “Alkenylene”, “alkenylene chain” or “alkenylene group” refers to a straight or branched divalent hydrocarbon chain radical, having from two to forty carbon atoms, and having one or more carbon-carbon double bonds. Non-limiting examples of C₂-C₄₀ alkenylene include ethene, propene, butene, and the like. Unless stated otherwise specifically in the specification, an alkenylene chain can be optionally.

[0588] “Alkoxy” or “alkoxy group” refers to the group -OR, where R is alkyl, alkenyl, alkynyl, cycloalkyl, or heterocyclyl as defined herein. Unless stated otherwise specifically in the specification, an alkoxy group can be optionally substituted.

[0589] “Acyl” or “acyl group” refers to groups -C(O)R, where R is hydrogen, alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, as defined herein. Unless stated otherwise specifically in the specification, acyl can be optionally substituted.

[0590] “Alkylcarbamoyl” or “alkylcarbamoyl group” refers to the group -O-C(O)-NR_aR_b, where R_a and R_b are the same or different and are independently an alkyl, alkenyl, alkynyl, aryl, heteroaryl, as defined herein, or R_aR_b can be taken together to form a cycloalkyl group or heterocyclyl group, as defined herein. Unless stated otherwise specifically in the specification, an alkylcarbamoyl group can be optionally substituted.

[0591] “Alkylcarboxamidyl” or “alkylcarboxamidyl group” refers to the group -C(O)-NR_aR_b, where R_a and R_b are the same or different and are independently an alkyl, alkenyl, alkynyl, aryl, heteroaryl, cycloalkyl, cycloalkenyl, cycloalkynyl, or heterocyclyl group, as defined herein, or R_aR_b can be taken together to form a cycloalkyl group, as defined herein. Unless stated otherwise specifically in the specification, an alkylcarboxamidyl group can be optionally substituted.

[0592] “Aryl” refers to a hydrocarbon ring system radical comprising hydrogen, 6 to 18 carbon atoms and at least one aromatic ring. For purposes of this invention, the aryl radical can be a monocyclic, bicyclic, tricyclic or tetracyclic ring system, which can include fused or bridged ring

systems. Aryl radicals include, but are not limited to, aryl radicals derived from aceanthrylene, acenaphthylene, acephenanthrylene, anthracene, azulene, benzene, chrysene, fluoranthene, fluorene, *as*-indacene, *s*-indacene, indane, indene, naphthalene, phenalene, phenanthrene, pleiadene, pyrene, and triphenylene. Unless stated otherwise specifically in the specification, the term “aryl” is meant to include aryl radicals that are optionally substituted.

[0593] “Heteroaryl” refers to a 5- to 20-membered ring system radical comprising hydrogen atoms, one to thirteen carbon atoms, one to six heteroatoms selected from nitrogen, oxygen and sulfur, and at least one aromatic ring. For purposes of this invention, the heteroaryl radical can be a monocyclic, bicyclic, tricyclic or tetracyclic ring system, which can include fused or bridged ring systems; and the nitrogen, carbon or sulfur atoms in the heteroaryl radical can be optionally oxidized; the nitrogen atom can be optionally quaternized. Examples include, but are not limited to, azepinyl, acridinyl, benzimidazolyl, benzothiazolyl, benzindolyl, benzodioxolyl, benzofuranyl, benzooxazolyl, benzothiazolyl, benzothiadiazolyl, benzo[*b*][1,4]dioxepinyl, 1,4-benzodioxanyl, benzonaphthofuranyl, benzoxazolyl, benzodioxolyl, benzodioxinyl, benzopyranyl, benzopyranonyl, benzofuranyl, benzofuranonyl, benzothienyl (benzothiophenyl), benzotriazolyl, benzo[4,6]imidazo[1,2-*a*]pyridinyl, carbazolyl, cinnolinyl, dibenzofuranyl, dibenzothiophenyl, furanyl, furanonyl, isothiazolyl, imidazolyl, indazolyl, indolyl, indazolyl, isoindolyl, indolinyl, isoindolinyl, isoquinolyl, indoliziny, isoxazolyl, naphthyridinyl, oxadiazolyl, 2-oxoazepinyl, oxazolyl, oxiranyl, 1-oxidopyridinyl, 1-oxidopyrimidinyl, 1-oxidopyrazinyl, 1-oxidopyridazinyl, 1-phenyl-1*H*-pyrrolyl, phenazinyl, phenothiazinyl, phenoxazinyl, phthalazinyl, pteridinyl, purinyl, pyrrolyl, pyrazolyl, pyridinyl, pyrazinyl, pyrimidinyl, pyridazinyl, quinazolinyl, quinoxalinyl, quinolinyl, quinuclidinyl, isoquinolinyl, tetrahydroquinolinyl, thiazolyl, thiadiazolyl, triazolyl, tetrazolyl, triazinyl, and thiophenyl (*i.e.* thienyl). Unless stated otherwise specifically in the specification, a heteroaryl group can be optionally substituted.

[0594] The term “substituted” used herein means any of the above groups (*i.e.*, alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkenyl, cycloalkynyl, heterocyclyl, aryl, heteroaryl, alkoxy, aryloxy, acyl, alkylcarbamoyl, alkylcarboxamidyl, alkoxycarbonyl, alkylthio, or arylthio) wherein at least one atom is replaced by a non-hydrogen atoms such as, but not limited to: a halogen atom such as F, Cl, Br, and I; an oxygen atom in groups such as hydroxyl groups, alkoxy groups, and ester groups; a sulfur atom in groups such as thiol groups, thioalkyl groups, sulfone groups, sulfonyl groups, and sulfoxide groups; a nitrogen atom in groups such as amines, amides, alkylamines,

dialkylamines, arylamines, alkylarylamines, diarylamines, N-oxides, imides, and enamines; a silicon atom in groups such as trialkylsilyl groups, dialkylarylsilyl groups, alkylarylsilyl groups, and triarylsilyl groups; and other heteroatoms in various other groups. “Substituted” also means any of the above groups in which one or more atoms are replaced by a higher-order bond (e.g., a double- or triple-bond) to a heteroatom such as oxygen in oxo, carbonyl, carboxyl, and ester groups; and nitrogen in groups such as imines, oximes, hydrazones, and nitriles. For example, “substituted” includes any of the above groups in which one or more atoms are replaced with $-NR_gR_h$, $-NR_gC(=O)R_h$, $-NR_gC(=O)NR_gR_h$, $-NR_gC(=O)OR_h$, $-NR_gSO_2R_h$, $-OC(=O)NR_gR_h$, $-OR_g$, $-SR_g$, $-SOR_g$, $-SO_2R_g$, $-OSO_2R_g$, $-SO_2OR_g$, $=NSO_2R_g$, and $-SO_2NR_gR_h$. “Substituted also means any of the above groups in which one or more hydrogen atoms are replaced with $-C(=O)R_g$, $-C(=O)OR_g$, $-C(=O)NR_gR_h$, $-CH_2SO_2R_g$, $-CH_2SO_2NR_gR_h$. In the foregoing, R_g and R_h are the same or different and independently hydrogen, alkyl, alkenyl, alkynyl, alkoxy, alkylamino, thioalkyl, aryl, aralkyl, cycloalkyl, cycloalkenyl, cycloalkynyl, cycloalkylalkyl, haloalkyl, haloalkenyl, haloalkynyl, heterocyclyl, *N*-heterocyclyl, heterocyclylalkyl, heteroaryl, *N*-heteroaryl and/or heteroarylalkyl. “Substituted” further means any of the above groups in which one or more atoms are replaced by an amino, cyano, hydroxyl, imino, nitro, oxo, thioxo, halo, alkyl, alkenyl, alkynyl, alkoxy, alkylamino, thioalkyl, aryl, aralkyl, cycloalkyl, cycloalkenyl, cycloalkynyl, cycloalkylalkyl, haloalkyl, haloalkenyl, haloalkynyl, heterocyclyl, *N*-heterocyclyl, heterocyclylalkyl, heteroaryl, *N*-heteroaryl and/or heteroarylalkyl group. “Substituted” can also mean an amino acid in which one or more atoms on the side chain are replaced by alkyl, alkenyl, alkynyl, acyl, alkylcarboxamidyl, alkoxycarbonyl, carbocyclyl, heterocyclyl, aryl, or heteroaryl. In addition, each of the foregoing substituents can also be optionally substituted with one or more of the above substituents.

[0595] As used herein, by a “subject” is meant an individual. Thus, the “subject” can include domesticated animals (*e.g.*, cats, dogs, etc.), livestock (*e.g.*, cattle, horses, pigs, sheep, goats, etc.), laboratory animals (*e.g.*, mouse, rabbit, rat, guinea pig, etc.), and birds. “Subject” can also include a mammal, such as a primate or a human. Thus, the subject can be a human or veterinary patient. The term “patient” refers to a subject under the treatment of a clinician, *e.g.*, physician.

[0596] The term “inhibit” refers to a decrease in an activity, response, condition, disease, or other biological parameter. This can include but is not limited to the complete ablation of the activity, response, condition, or disease. This can also include, for example, a 10% reduction in the activity,

response, condition, or disease as compared to the native or control level. Thus, the reduction can be a 10, 20, 30, 40, 50, 60, 70, 80, 90, 100%, or any amount of reduction in between as compared to native or control levels.

[0597] By “reduce” or other forms of the word, such as “reducing” or “reduction,” is meant lowering of an event or characteristic (*e.g.*, tumor growth). It is understood that this is typically in relation to some standard or expected value, in other words it is relative, but that it is not always necessary for the standard or relative value to be referred to. For example, “reduces tumor growth” means reducing the rate of growth of a tumor relative to a standard or a control (*e.g.*, an untreated tumor).

[0598] The term “treatment” refers to the medical management of a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder. This term includes active treatment, that is, treatment directed specifically toward the improvement of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, pathological condition, or disorder. In addition, this term includes palliative treatment, that is, treatment designed for the relief of symptoms rather than the curing of the disease, pathological condition, or disorder; preventative treatment, that is, treatment directed to minimizing or partially or completely inhibiting the development of the associated disease, pathological condition, or disorder; and supportive treatment, that is, treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder.

[0599] The term “therapeutically effective” refers to the amount of the composition used is of sufficient quantity to ameliorate one or more causes or symptoms of a disease or disorder. Such amelioration only requires a reduction or alteration, not necessarily elimination.

[0600] The term “pharmaceutically acceptable” refers 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 problems or complications commensurate with a reasonable benefit/risk ratio.

[0601] The term “carrier” means a compound, composition, substance, or structure that, when in combination with a compound or composition, aids or facilitates preparation, storage, administration, delivery, effectiveness, selectivity, or any other feature of the compound or

composition for its intended use or purpose. For example, a carrier can be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject.

[0602] As used herein, the term "pharmaceutically acceptable carrier" refers to sterile aqueous or nonaqueous solutions, dispersions, suspensions or emulsions, as well as sterile powders for reconstitution into sterile injectable solutions or dispersions just prior to use. Examples of suitable aqueous and nonaqueous carriers, diluents, solvents or vehicles include water, ethanol, polyols (such as glycerol, propylene glycol, polyethylene glycol and the like), carboxymethylcellulose and suitable mixtures thereof, vegetable oils (such as olive oil) and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of coating materials such as lecithin, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants. These compositions can also contain adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microorganisms can be ensured by the inclusion of various antibacterial and antifungal agents such as paraben, chlorobutanol, phenol, sorbic acid and the like. It can also be desirable to include isotonic agents such as sugars, sodium chloride and the like. The injectable formulations can be sterilized, for example, by filtration through a bacterial-retaining filter or by incorporating sterilizing agents in the form of sterile solid compositions which can be dissolved or dispersed in sterile water or other sterile injectable media just prior to use. Suitable inert carriers can include sugars such as lactose.

[0603] Table 10

20

Abbreviation	IUPAC Name
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate
HOBt	1-hydroxybenzotriazole
PyAOP	(7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
PyOxim	[(E)-(1-cyano-2-ethoxy-2-oxoethylidene)amino]oxy-tripyrrolidin-1-ylphosphonium;hexafluorophosphate
Oxyma	Ethyl (2Z)-2-cyano-2-(hydroxyimino)acetate

HBTU	3-[Bis(dimethylamino)methylumyl]-3H-benzotriazol-1-oxide hexafluorophosphate
TBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate
COMU	(1-Cyano-2-ethoxy-2-oxoethylidenaminoxy)dimethylamino-morpholino-carbenium hexafluorophosphate
DEPBT	Diethyl 4-oxo-1,2,3-benzotriazin-3(4H)-yl phosphate

EXAMPLES

Example 1. Construction of a cell-penetrating peptide - antisense compound conjugate

- 5 **[0604] Oligonucleotide design.** An antisense compound (AC) is designed to bind to and block mRNA expression of a target protein of interest is constructed as a phosphorodiamidate morpholino oligomer (PMO) with a C6-thiol 5' modification.

- [0605] Cell penetrating peptide.** A cell-penetrating peptide is formulated using Fmoc chemistry and conjugated to the AC, for example, as described in International Application No.
 10 PCT/US20/66459, filed by Entrada Therapeutics, Inc., on December 21, 2021, entitled "COMPOSITIONS FOR DELIVERY OF ANTISENSE COMPOUNDS," the disclosure of which is hereby incorporated in its entirety herein. In embodiments, the CPP is cCPP12, which has an amino acid sequence of FfΦRrRr.

Example 2. Construction of a cell-penetrating peptide - antisense compound conjugate

- 15 **[0606] Oligonucleotide design.** An antisense compound (AC) is designed to bind to and block mRNA expression of a target protein of interest and is constructed as a phosphorodiamidate morpholino oligomer (PMO) composed exclusively of phosphorodiamidate morpholino bases.

- [0607] Cell penetrating peptide.** A cell-penetrating peptide containing arginine derivatives was formulated using Fmoc chemistry and conjugated to the AC, for example, as described in U.S.
 20 Provisional Application No. 63/171,860, filed by Ziqing Qian, on April 7, 2021, entitled "NOVEL CYCLIC CELL PENETRATING PEPTIDES," the disclosure of which is hereby incorporated in its entirety herein. In embodiments, the CPP has the sequence Acetyl-Pro-Lys-Lys-Lys-Arg-Lys-Val-PEG₂-Lys(cyclo[Phe-D-Phe-2-Nal-Cit-D-Arg-Cit-D-Arg-γ-Glu])-PEG₁₂-Lys(N₃)-NH₂.

Example 3. Construction of cell-penetrating peptide – antisense compound conjugates containing a modulatory peptide sequence

[0608] The compounds of Table A below were prepared as described. Compounds ENTR-0047, ENTR-0168, ENTR-0203 and ENTR-0207 each contain an exocyclic peptide (EP) sequence, indicated by “NLS”, while compounds ENTR-0006, ENTR-0070, ENTR-0059 and ENTR-0121 lack an NLS.

Table A. Compounds that target EGFP-654

Oligo#	Design	Sequence
ENTR-0006	PMO-654	5'-GCT ATT ACC TTA ACC CAG-3'
ENTR-0070	PMO654-NH2	5'-GCT ATT ACC TTA ACC CAG-3' -C6-NH2 (all PMO monomers)
ENTR-0059	PMO654-LSR	cyclooctyne-5'-GCT ATT ACC TTA ACC CAG-3' -Lissamine Rho (all PMO monomers)
ENTR-0123	CPP12-PMO654-LSR	(CPP12-PEG12)-5'-GCT ATT ACC TTA ACC CAG-3' -Lissamine Rho (all PMO monomers)
ENTR-0121	PMO654-CPP12	5'-GCT ATT ACC TTA ACC CAG-3' -NH2 (all PMO monomers)+CPP12-PEG12-TFP-Amide chemistry
ENTR-0047	PMO654-NLS-ss-CPP12	5'-GCT ATT ACC TTA ACC CAG-3-NLS-ss-CPP12 (all PMO monomers)
ENTR-0168	NLS-CPP12-PMO654-LSR	(Ac-NLS-Lys(CPP12)-PEG12-K(N3))-5'-GCT ATT ACC TTA ACC CAG-3' -Lissamine Rho (all PMO monomers)
ENTR-0203	PMO654-CPP12-NLS	5'-GCT ATT ACC TTA ACC CAG-3'-click-K-PEG12-Lys(CPP12)-NLS-Ac (all PMO monomers)
ENTR-0207	PMO654-NLS-CPP12	5'-GCT ATT ACC TTA ACC CAG-3'-click-Lys(N3)-miniPEG-NLS-ss-PEG12-CPP12 (all PMO monomers)

[0609] **Target gene design.** A missense EGFP gene ("EGFP-654") was designed with a mutated intron 2 of the human β globin gene interrupting the EGFP coding sequence. A mutation was

introduced at nucleotide 654 of intron 2, which activated aberrant splice sites and led to retention of the intron fragment in spliced, mature mRNA, thereby preventing proper translation of EGFP.

[0610] Oligonucleotide design. An antisense compound (AC) was designed to bind to and block the aberrant splice site of the target gene in order to correct pre-mRNA splicing and restore EGFP expression. The AC had the sequence "5'-GCTATTACCTTAACCCAG-3'" and was designed as a phosphorodiamidate morpholino oligomer (PMO) with a C6-thiol 5' modification (ENTR-0006, Table A).

[0611] Cell penetrating peptide. A cell-penetrating peptide comprising cyclo(Phe-D-Phe-SNal-Arg-D-Arg-Arg-D-Arg-γ-Glu)-b-Ala-b-Ala-Lys(Maleimide)-NH₂ ("CPP12-Maleimide") was formulated as a TFA salt. The peptide was synthesized using standard Fmoc chemistry according to the following procedure:

- a. 1. DCM added to the vessel containing Rink amide resin (1 mmol, 1 g, 1.0 mmol/g) and Fmoc-Lys(Trt)-OH (1 eq) with N₂ bubbling.
- b. 2. DIEA (4.0 eq) was added dropwise and mix for 4 hours.
- 15 c. 3. MeOH (0.2 mL) was added to the resin and mixed for 30 min.
- d. 4. Solution was drained and washed with DMF for 5 times.
- e. 5. 20% piperidine/DMF was added and reacted for 30 min.
- f. 6. The solution was drained and washed with DMF for 5 times.
- g. 7. Fmoc-amino acid solution was added to resin and mixed for 30 seconds, then activation
20 buffer was added and mixed with N₂ bubbling for about 1 hour.
- h. 8. Step 4 to 7 was repeated for the coupling of following amino acids. All couplings completions were confirmed by negative ninhydrin test. After couplings, resin was washed with DMF for 5 times.
- i. 9. Allyl protecting group was removed by adding PhSiH₃ (10eq) and Pd(PPh₃)₄ (0.1 eq)
25 in DCM, repeat twice.
- j. 10. Peptide was cyclized by HATU (1.0 eq) and DIEA (2 eq). The cyclization reaction was monitored by ninhydrin test. Resin was washed with DMF five times.

k. 11. The Trt protecting group was removed with 20% HFP/80% DCM and the resulting primary amine was reacted with OAt activated 3-Maleimidopropionic acid using HATU/DIPEA for 2 hours

l. 12. After coupling, the resin was washed with MeOH for 3 times and dried under reduce pressure.

[0612] The table below shows the materials used for solid-phase peptide synthesis and coupling reagents:

#	Materials	Coupling reagents
1	Fmoc-NH-PEG12-CH ₂ CH ₂ COOH (1.0 eq)	DIEA (4.0 eq)
2	Fmoc-Glu-OAll (1.5 eq)	HATU (1.45 eq) and DIEA (3.0 eq)
3	D-Fmoc-Arg(Pbf)-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
4	Fmoc-Arg(Pbf)-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
5	D-Fmoc-Arg(Pbf)-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
6	Fmoc-Arg(Pbf)-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
7	Fmoc-3-(2-Nal)-Ala-OH (1.5 eq)	HATU (1.45 eq) and DIEA (3.0 eq)
8	D-Fmoc-Phe-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
9	Fmoc-Phe-OH (3 eq)	HBTU (2.85 eq) and DIEA (6.0 eq)
10	Cyclization	HATU (1.00 eq) and DIEA (2.0 eq)

[0613] The peptide was cleaved from the solid-phase peptide synthesis resin and purified according to the following procedure:

- a. Add cleavage buffer (95 % TFA, 2.5 % TIPS, 2.5 % H₂O) to the flask containing the side chain protected peptide at room temperature and stir for 1 hour and repeat one time.
- b. Filter and collect the peptide solution.
- c. The peptide is concentrated under reduce press to give residue.
- d. The resulting solid is dissolved in CH₃CN and H₂O, then lyophilized to give crude peptide (2 g, 70.1% yield) as white solid.

[0614] **CPP-AC conjugate formation.** The steps of the conjugation process are shown in **FIG. 11**. Compound 1 was treated with 1M TCEP and MeCN/H₂O to reduce the 5' end resulting in compound 2. To this solution of compound 2 (~13 mg, from reduction of ~10 mg*2 PMO654-G) in H₂O/CH₃CN (3:2, 2 mL) was added **CPP12-Maleimide** (4.5 mg, 1.2 eq, previously dissolved

in H₂O/CH₃CN (3:2), 98.4 uL) in one portion. Then to the mixture was added PB buffer (PH=7, 1mL) and stirred for 1 hour at 25 °C. Another batch (~12 mg) processed in parallel. LCMS shown the compound **1** was consumed completely and the solution was directly injected and purified by C18 reverse phase column. The mixture was first eluted with TEEA condition (2 mM TEAA in water, CH₃CN) and then eluted with TFA condition (0.075% TFA in water, CH₃CN) to give **CPP12-Maleimide-PMO-654** (28 mg, 96.3% purity, total yield: 56.8%).

[0615] Purification Conditions:

Separation condition	
Dissolution condition	Dissolve in H ₂ O
Instrument	GX-281D
Mobile Phase 1	A: H ₂ O (2mM TEEA in H ₂ O)
	B: CH ₃ CN
Gradient 1	10-60%B 60min. Separate the impurity that could be elute with TEEA condition.
Mobile Phase 2	A: H ₂ O (0.075% TFA in H ₂ O)
	B: CH ₃ CN
Gradient 2	20-50%B 60min. Retention time: 14.5 min
Column	Luna 200*25 mm, C18, 10 um, 100Å+Gemini® 150*30 mm, C18, 5 um, 110Å
Flow Rate	20 mL/Min
Wavelength	220/254 nm
Oven Tem.	Room temperature

[0616] **PMO synthesis protocol and linker installation at the 3' end.** The following protocol was used to synthesize the PMO. For every step of the following synthesis protocol, it was ensured

- that the volume of reagent or solvent used completely covers resin (add more, if needed). Volumes listed below are estimates in number of milliliters per gram of resin used, which will increase during synthesis due to increasing resin size. All synthesis steps in the table were performed at room temperature. Prior to synthesis the resin is swelled in NMP for 1 hour. The resin was washed
- 5 two times with DCM, followed by washing resin 2 times with 30% TFE/DCM (15 mL/g of resin). The table below describes the PMO synthesis and linker installation protocol:

Entry	Step	Volume	Time	Mixture
1	Wash	7-24 mL/g	15 seconds	30% TFE/DCM
2	Deprotection	7-24 mL/g	15 minutes	CYTFA Solution
3	Deprotection	7-24 mL/g	15 minutes	CYTFA Solution
4	Wash	7-24 mL/g	15 seconds	DCM
5	Wash	7-24 mL/g	15 seconds	DCM
6	Neutralization	7-24 mL/g	5 minutes	Neutralization Solution
7	Neutralization	7-24 mL/g	5 minutes	Neutralization Solution
8	Wash	7-24 mL/g	15 seconds	DCM
9	Wash	7-24 mL/g	15 seconds	DCM or anhydrous DMI
10	Coupling	4.6-7.3 mL/g	4.25-5.0 hours	[0001] -6 eq monomer, 5-8 eq NEM in DMI (0.2-0.4 M monomer & 0.5 M NEM)
11	Wash	7-24 mL/g	15 seconds	DCM

12	Capping	7-24 mL/g	15 minutes	0.55 M Benzoic anhydride + 0.55 M NEM in NMP
13	Neutralization	7-24 mL/g	5 minutes	Neutralization Solution
14	Wash	7-24 mL/g	15 seconds	DCM
15	Wash	7-24 mL/g	15 seconds	30% TFE/DCM
16	Wash	7-24 mL/g	15 seconds	30% TFE/DCM

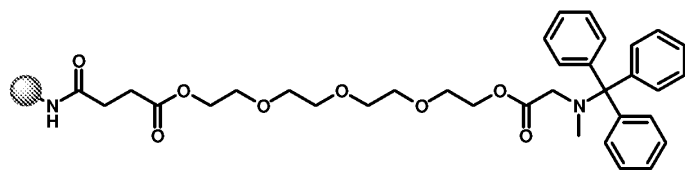
CYTFA Solution = 100 mM 4-cyanopyridine + 100 mM TFA in 4:1 DCM:TFE + 1% Ethanol.

[0617] Below are the solutions used for synthesis:

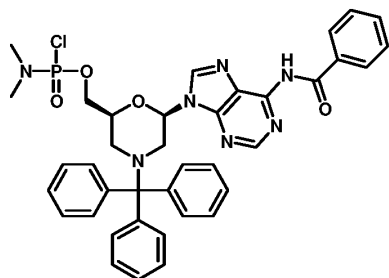
- a. Neutralization Solution = 5% DIPEA in 3:1 DCM:iPrOH.
- b. Coupling Solution = The number of equivalents and concentration will increase throughout synthesis using the following guide:
- c. Residues 1-10 = 3 eq morpholino monomer, 5 eq NEM in DMI (0.2 M morpholino monomer, 0.5 M NEM), rt, 4.25h.
- d. Residue 1: couple for 5h at rt.
- e. Residues 11-20 = 4 eq morpholino monomer, 6.5 eq NEM in DMI (0.3 M morpholino monomer, 0.5 M NEM), rt, 4.25h.
- f. Residues 21-25 = 5 eq morpholino monomer, 8 eq NEM in DMI (0.3 M morpholino monomer, 0.5 M NEM), rt, 4.25h.
- g. Guanine morpholino monomers: couple for 4.75h at rt.

Some couplings employed 6 eq morpholino monomer, 8 eq NEM in DMI (0.4 M morpholino monomer, 0.5 M NEM) at 45 °C for 4.75 hours.

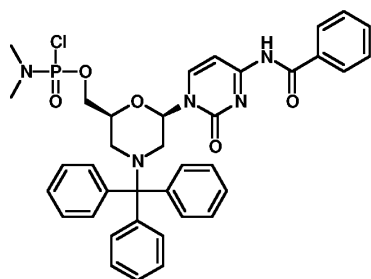
The PMO synthesis resin with linker has the following structure:



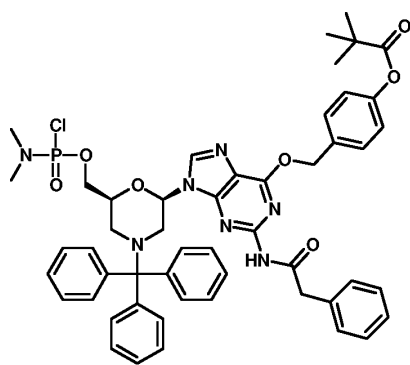
[0618] The morpholino monomers used during coupling have the following structures:



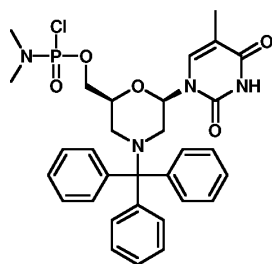
Adenine morpholino monomer



Cytosine morpholino monomer



Guanine morpholino monomer



Thymine morpholino monomer

[0619] Below is the protocol for PMO synthesis:

[0620] Deprotection: Resin was first washed with 30% TFE/DCM solution and was allowed to stand for 15 seconds before draining. CYTFA solution was then added to the drained resin and
5 reacted for 15 minutes. The resin was drained, then fresh CYTFA solution was added and reacted again for 15 minutes. The resin was drained and rinsed twice with DCM for 15 seconds before proceeding to neutralization.

[0621] Neutralization: Neutralization solution was added to the resin, stirred, and was allowed to stand for 5 minutes, then drained. A second wash of fresh neutralization solution was delivered to
10 the resin, stirred, and reacted for 5 minutes. The resin was washed once with DCM and once with either DCM or anhydrous DMI before coupling.

[0622] Coupling: Using the guide listed above, two coupling solutions were made: 1) PMO monomer dissolved in DMI, and 2) NEM dissolved in DMI. These two solutions were mixed immediately before adding to the resin. The resin was stirred and reacted for 4.25-5 hours. The
15 resin was washed one time with DCM.

[0623] Capping: A capping solution consisting of 0.55 M benzoic anhydride and 0.55 M NEM in NMP was added to the resin and reacted for 15 minutes. The resin was drained, and Neutralization solution was added to the resin to react for 5 minutes. The resin was drained again and was washed once with DCM, then twice with 30% TFE/DCM solution.

[0624] Post-synthesis: After the final coupling step, the resin-bound PMO can be stored until it is cleaved by washing the resin eight times with iPrOH, then drying the resin under vacuum at room temperature (Note: 3'-Trityl protecting group must still be on PMO for this). For PMO modifications at 3', the Trityl protecting group was removed, resin was neutralized, then appropriate bifunctional linker (TFA-protected amino or cyclooctyne) PFP ester (4 eq) in NMP

and DIPEA (8 eq) were added to the resin and reacted for 3 hours. Solution was drained and resin was washed once with DCM, then twice with 30% TFE/DCM solution.

[0625] Cleavage: The following options can be utilized to perform PMO cleavage:

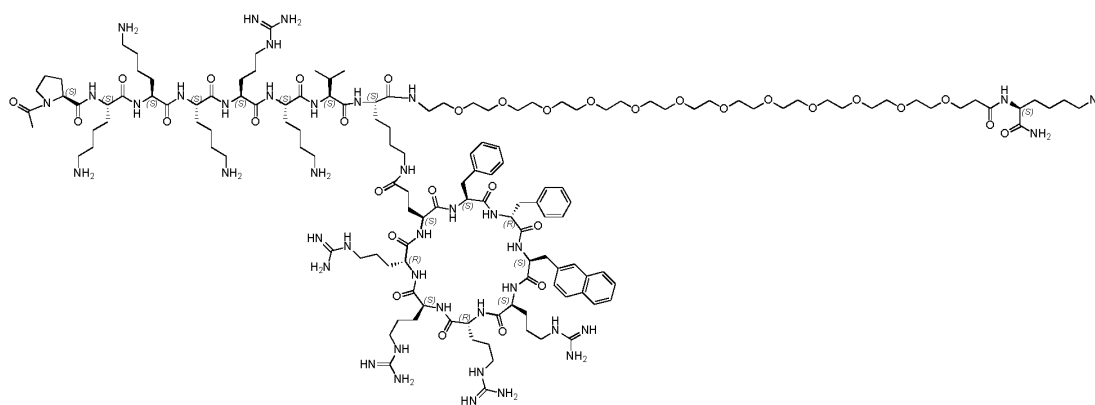
- a. PMO was cleaved from resin with a 1:1 solution of ammonium hydroxide (25% ammonia, aqueous) and methylamine (8 mL/g) at 65 °C for 15 minutes.
- b. PMO was cleaved from resin with ammonium hydroxide (25% ammonia, aqueous) at 65 °C for 16 hours.
- c. PMO was cleaved from resin with 7M Ammonia in methanol solution (8 mL/g of resin) at 65 °C for 16 hours.
- d. The deprotected PMO solution was then desalted prior to lyophilization and purification.
- e. Resulting PMO was purified by Clarity 5 µm, C18 oligo reverse phase (250 mm×30 mm), AXIA packed, with 10-30% gradient over 40 min, flow rate of 30 mL/min, solvent A as water with 0.05% TFA and solvent B as Acetonitrile.

[0626] Design and preparation of CPP-PMO654 conjugates. For 3' covalent conjugation of primary amine modified PMOs, a solution of desired peptide-TFP ester in DMF (4 eq, 5 mM) was added to a solution of PMO-3'-primary amine (1 equivalent, 2 mM) in PBS-10X. Reaction proceeded to completion in 4-8 hours at room temperature as confirmed by LCMS (Q-TOF), using BEH C18 column (130Å, 1.7 µm, 2.1mm×150 mm), buffer A: water, 0.1% FA), buffer B: acetonitrile, 0.1% FA), flow rate (0.3 mL/min), starting with 2% buffer B and ramping up to 70% for 11 min for a total of 20 min run. For 3' or 5' conjugation via click reaction, a solution of peptide-azide in nuclease-free water (1 mM) was added to the PMO-3'-cyclooctyne or cyclooctyne-5'-PMO solids. The mixture was vortexed to dissolve the peptide-PMO conjugate, centrifuged to settle the solution, and incubated at room temperature for 8-12 hours for completion as confirmed by LCMS (Q-TOF). For purifications, crude mixtures were diluted with DMSO, loaded onto a C18 reverse-phase column (150mm× 21.2 mm), flow rate of 20 mL/min and purified by an appropriate gradient over 20 min using water with 0.05% TFA and acetonitrile as solvents. Desired fractions were pooled, pH of the solution was adjusted to 5-6 by 1M NaOH and the solution underwent the lyophilization process, affording white lyophilized powder. For *in vitro* and *in vivo* formulations, the conjugates were reconstituted in appropriate amount of PBS or Saline for the desired concentration (2-10 mg/mL). Concentration of the non LSR labeled conjugates were measured by

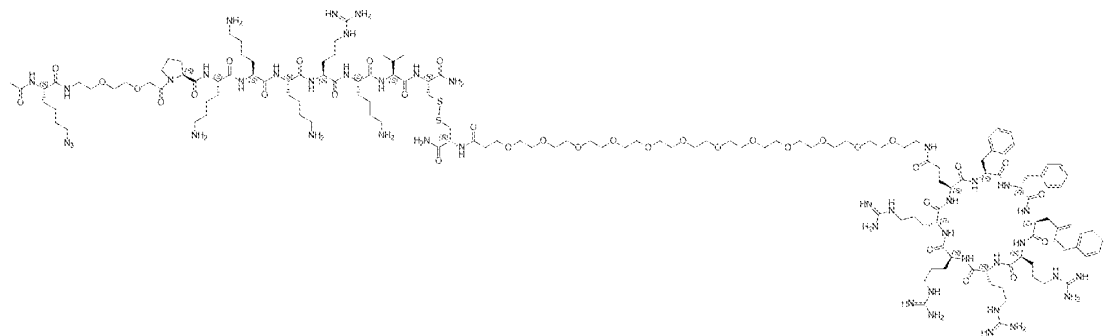
- preparing 10, 20, and 50-fold dilutions in formulated buffer and reading the absorbance at 260 nm or 280 nm using a nanodrop. Once the linear range of dilution achieved, the absorbance was measured in triplicates and concentration was calculated using the average absorbance and ϵ_{260} or ϵ_{280} . ϵ_{280} for conjugates were calculated by the following formula: $\epsilon_{280} = 100356 + (n \times 3550)$; n= number of CPP. For LSR modified PMOs, the concentrations were measured at 566 nm with $\epsilon_{566} = 100000 \text{ LMol}^{-1}\text{Cm}^{-1}$. The diluted samples were analyzed by LCMS (Q-TOF) for the conjugate identity confirmation. Table below summarizes the calculated MW and experimental MWs. All experimental MWs reasonably matched the calculated average MW with expected ± 6 Da assay variation.

Name	Calculated MW	Experimental Average MW	Purity
ENTR 0006	6038.19		
ENTR 0070	6151.27	6148.22	>99
ENTR 0059	7411.06	7409.78	99
ENTR 0121	7938.3	7933.15	97
ENTR 0047	9039.81	9034.81	99
ENTR 0168	10444.78	10447.3	95
ENTR 0203	9293.79	9295.2	92
ENTR 0207	9532.05	9532.8	>99

- 10 [0627] The structure of ENTR-0203 is below:



- [0628] The structure of ENTR-0207 is below:



Example 4. Use of cell-penetrating peptide coupled to an oligonucleotide and nuclear localization sequence for splicing correction of exon 23 of DMD in an MDX mouse model

5 **[0629] Purpose.** This study employs an MDX mouse model, a model of DMD, to study the effect of compositions comprising an AC, a CPP, and a nuclear localization sequence (NLS, or modulatory peptide, EP) on dystrophin expression and muscle fiber damage.

[0630] Preparation and design of CPP-PMO targeting murine DMD exon 23. Design of antisense compounds (AC) that target DMD exon 23 are shown below in Table B1. Design of
10 CPP-NLS-PMO constructs (ENTR-0164, ENTR-0165, ENTR-0201) are shown below in Table B2.

[0631] Table B1. Compounds that target DMD

Oligo#	Design	Target
ENTR-0013	5'-GGC CAA ACC TCG GCT TAC CTG AAA T-3' (all PMO monomers)	Murine DMD exon 23
ENTR-0017	5'-GGC CAA ACC TCG GCT TAC CTG AAA T-3'-C3-NH2 (all PMO monomers)	Murine DMD exon 23
ENTR-0066	Cyclooctyne-5'- GGC CAA ACC TCG GCT TAC CTG AAA T -3' (all PMO monomers)	Murine DMD exon 23
ENTR-0068	NH2-5'-C3- GGCCAAACCTCGGCTTACCTGAAAT -3'- C3-NH2 (all PMO monomers)	Murine DMD exon 23
ENTR-0149	5'- GGC CAA ACC TCG GCT TAC CTG AAA T -3'-C4-cyclooctyne (all PMO monomers)	Murine DMD exon 23

[0632] Table B2. CPP-NLS-PMO Compounds that target DMD

Oligo#	Design	Peptide Fusion
ENTR-0164	5'- GGC CAA ACC TCG GCT TAC CTG AAA T -3'-C3NH-PEG4-BCN+Lys(N3)- miniPEG-NLS-ss-bA-bA-CPP12 (all PMO monomers)	NLS: PKKKRKV
ENTR-0165	5'- GGC CAA ACC TCG GCT TAC CTG AAA T -3'-C3NH-PEG4-BCN+Ac-NLS- Lys(CPP12)-PEG12-K(N3) (all PMO monomers)	NLS: PKKKRKV
ENTR-0201	5'- GGC CAA ACC TCG GCT TAC CTG AAA T -3'-click-K-PEG12-Lys(CPP12)- NLS-Ac (all PMO monomers)	NLS: PKKKRKV

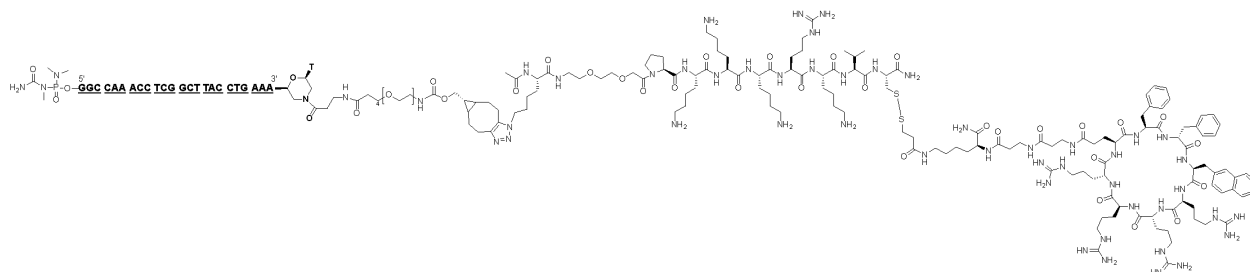
[0633] For 3' or 5' conjugation via click reaction (ENTR-0164, ENTR-0165, ENTR-0201), a solution of peptide-azide in nuclease-free water (1 mM) was added to the PMO-3'-cyclooctyne or cyclooctyne-5'-PMO solids. The mixture was vortexed to dissolve the peptide-PMO conjugate, centrifuged to settle the solution, and incubated at room temperature for 8-12 hours for completion as confirmed by LCMS (Q-TOF). For purifications, crude mixtures were diluted with DMSO, loaded onto a C18 reverse-phase column (150mm* 21.2 mm), flow rate of 20 mL/min and purified by an appropriate gradient using water with 0.05% TFA and acetonitrile as solvents. Desired fractions were pooled, pH of the solution was adjusted to 5-6 by 1M NaOH and the solution underwent the lyophilization process, affording white lyophilized powder. For *in vitro* and *in vivo* formulations, the conjugates were reconstituted in appropriate amount of PBS or Saline for the desired concentration (2-10 mg/mL). Concentration of the non LSR labeled conjugates were measured by preparing 10, 20, and 50-fold dilutions in formulated buffer and reading the absorbance at 260 nm or 280 nm using a nanodrop. Once the linear range of dilution achieved, the absorbance was measured in triplicates and concentration was calculated using the average absorbance and ϵ_{260} or ϵ_{280} . ϵ_{280} for conjugates were calculated by the following formula: $\epsilon_{280} = 138993 + (n \times 3550)$; n= number of CPP. The diluted samples were analyzed by LCMS (Q-TOF)

for the conjugate identity confirmation. Table below summarizes the calculated MW and experimental MWs. All experimental MWs reasonably matched the calculated average MW with expected ± 6 Da assay variation.

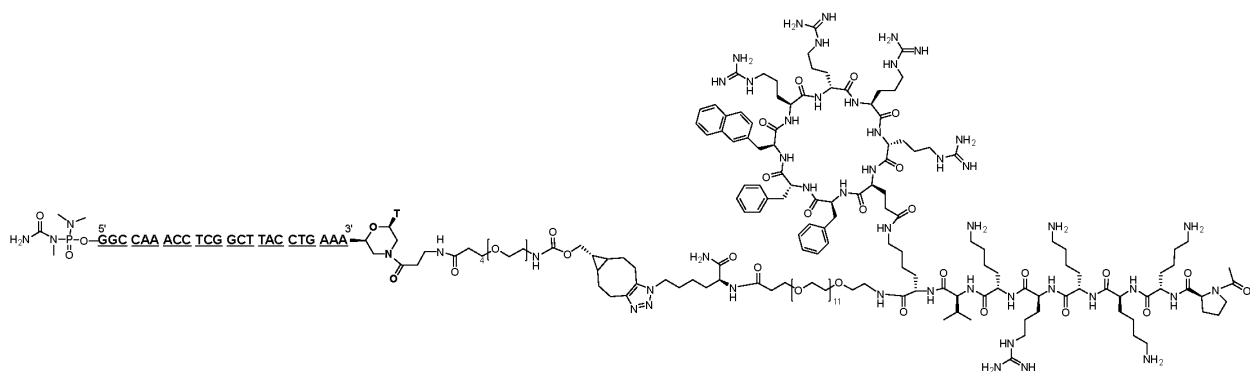
Name	Calculated MW	Experimental Average MW	Purity
ENTR-0013	8413.1	8410.02	99
ENTR-0017	8487	8484	95
ENTR-0066	9001.7	9003.8	95
ENTR-0068	8751.07	8748.02	85
ENTR-0149	8635.07	8635.8	82
ENTR-0164	11838.19	11836.3	97
ENTR-0165	11944.31	11942.3	99
ENTR-0201	11669.5	11669.5	>99

- 5 [0634] The structures of examples of compounds comprising antisense oligonucleotides (underlined) that target murine DMD and CPPs are shown below.

ENTR-0164:

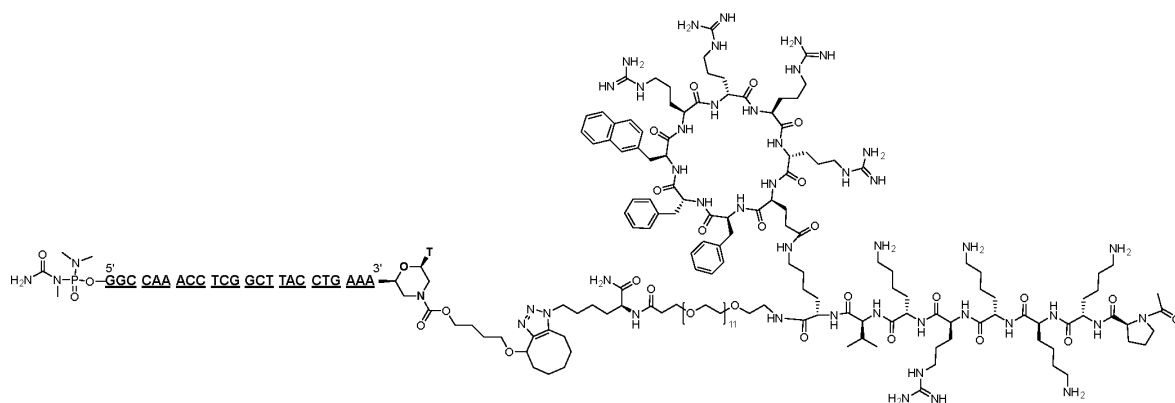


ENTR-0165:



5

ENTR-0201:



[0635] Study design. Compositions comprising an AC having a sequence of 5'-GGCCAAACCTCGGCTTACCTGAAAT-3', a cCPP12 (amino acid sequence is FfΦRrRr), and a nuclear localization sequence PKKKRKV (referred to herein as "ENTR-201") are applied to MDX mice to evaluate the ability of the compositions to skip exon 23 and thus treat DMD. A control composition lacks cCPP12 and a nuclear localization sequence. The sequence of the AC of the control composition is 5'-GGC CAA ACC TCG GCT TAC CTG AAA T-3'. The ENTR-201 composition is administered to the mice intravenously (IV) at a dose of 10 mg/kg once per week for four weeks or at a dose of 20 mpk once. The control composition is administered to mice intravenously (IV) at a dose of 20 mpk. Total RNA is extracted from tissue samples and analyzed by RT-PCR and protein is extracted from tissue sample and analyzed by Western Blot to visualize the efficiency of splicing correction and to detect dystrophin products. The percentage of exon 23 corrected products is evaluated. The dystrophin protein level is evaluated with respect to alpha-actinin (loading control) as well as in comparison to dystrophin expression in wild-type mice. Serum levels of creatine kinase, which is increased in DMD patients as a result of muscle fiber damage, were also evaluated by a commercially available kit purchased from Sigma Chemicals.

[0636] Evaluations of peptide fusions to the CPP12-PMO construct. To further increase the functional delivery of PMO, we also explored various peptides fusions for the CPP-PMO constructs. For one example, T9 peptide (SKTFNTHPQSTP) (Y. Seow et al. / Peptides 31 (2010) 1873–1877) and muscle specific peptide (MSP peptide, ASSLNIA (Gao et al. Molecular Therapy (2014) 22, 7: 1333-1341) which have been demonstrated to enhanced muscle targeting were also fused to CPP12 construct as in ENTR-0119 and ENTR-0163 respectively. Neither of these peptide fusions improved the activity in MDX mice.

[0637] Notably, CPP12 with a Nuclear Localization Sequence (PKKKRKV) outperformed CPP12 alone significantly (e.g. ENTR-164, ENTR-0165, and ENTR-0201, Table B2). One week after a single intravenous dose at 20 mpk, ENTR-164 yielded $59.5 \pm 2.2\%$, $29.1 \pm 0.8 \%$, and $61.0 \pm 13.7\%$ exon 23 skipping in Quad, heart, and diaphragm respectively. One week after a single intravenous dose at 20 mpk, ENTR-165 yielded $38.5 \pm 2.5\%$, $30.5 \pm 17.0 \%$, and $30.2 \pm 2.8\%$ exon 23 skipping in Quad, heart, and diaphragm respectively. One week after a single intravenous dose at 20 mpk, ENTR-201 yielded $73.5 \pm 8.4\%$, $60.5 \pm 17.2 \%$, and $79.0 \pm 6.8\%$ exon 23 skipping in Quad, heart, and diaphragm respectively. The preparation of these NLS fusions is described in

detail above. The structure of ENTR-164, ENTR-165 are shown above. Consistent with the exon skipping data, by ENTR-0164 and ENTR-0165 also significantly corrected the expression of dystrophin protein levels as analyzed by Western Blot. By comparing to the level in respective tissues from wild type (C57BL/10) and using the actinin as loading control, percentage of dystrophin protein correction is further quantified to that of wild type. Instead of using the modified PMO with 3' amide bond formation, we also tested the incorporation of cyclooctyne on solid support by modifying the morpholino amino group with a bifunctional linker comprised of a cyclooctyne moiety for click reaction to a CPP-azide and a PFP ester to form a carbamate with PMO and thus produced the precursor which can be used for synthesis of ENTR-201. Similar to that of ENTR-165, ENTR-201 also demonstrated high exon skipping activity across all the muscle groups.

[0638] Activity of ENTR-201 in MDX mice. The table below outlines the injection, sample collection and bioanalysis to study the duration of effects of ENTR-201 after single IV injection.

Group	N	Treatment	Dosage (mpk)	SAC time (p.i.)	Endpoints
1	3	PBS	0	1 week	- Exon skipping (RT-PCR) - Dystrophin expression (Western blot and immunohistochemistry)
2	3	ENTR-201	20	1 week	
3	3	ENTR-201	20	2 weeks	
4	3	ENTR-201	20	4 weeks	

[0639] After a single dose of ENTR-201 at 20 mg/kg on day 1, animals were sacrificed at 1 week, 2 weeks and 4 weeks post injection. Vehicle (PBS) only was used as a negative control. Heart, diaphragm, quadriceps and transverse abdominis were collected for RT-PCR to detect dystrophin exon 23 skipped product, and Western blot analysis to detect dystrophin protein expression (relative to alpha-actinin). Samples from 4 weeks post single IV injection of ENTR-201 at 20 mpk or PBS were also analyzed by immunohistochemistry staining to detect expression and distribution of dystrophin in various muscle tissues.

[0640] Treatment of mice with single dose 20 mg/kg ENTR-0201 resulted in splicing correction of dystrophin in the heart, diaphragm, quadriceps and transverse abdominis (TrA). ENTR-0201 delivers significant enhancements in exon skipping efficiency up to four weeks post single IV injection. The corresponding dystrophin protein levels were analyzed by Western Blot. Restored dystrophin protein sustained up to four weeks after single IV injection at 20 mpk in the heart, diaphragm, quadriceps and transverse abdominis (TrA). The level of protein correction is consistent with the RNA analysis. The tissue samples from the last injection were also analyzed by immunohistochemistry, which show that all the skeletal muscle fibers immunostained positive for dystrophin protein as visualized by brown color staining. The intensity of dystrophin expression was significant in the heart muscle tissue reaching near normal levels. Widespread uniform expression of dystrophin protein over multiple tissue sections within each of muscle group analyzed.

[0641] **Activity of ENTR-201 in MDX mice after repeated dosage.** The table below outlines the injection, sample collection and bioanalysis to study the activity of ENTR-201 after repeated dosage.

Group	N	Treatment	Dosage (mpk)	Dosage Frequency	Endpoints
1	3	PBS	0	QW x 4	• Ex
2	3	ENTR-013	20	QW x 4	• on skipping (RT-PCR) • Cr
3	3	ENTR-201	10	QW x 4	• eatine Kinase • Dy • dystrophin Expression (Western blot and immunohistochemistry)

[0642] Mice were treated with 10 mg/kg of ENTR-201 once every week for 4 weeks. ENTR-013 at 20 mg/kg (PMO only) and vehicle (PBS) only were used as control groups. All animals were sacrificed at 1 week post the last injection. Heart, diaphragm, quadriceps and transverse abdominis were collected for RT-PCR to detect dystrophin exon skipped products, Western blot analysis and

immunohistochemistry staining to detect dystrophin protein expression to detect expression and distribution of dystrophin in various tissues. Serum creatine kinase level was quantified as a muscle functional biomarker.

[0643] Treatment of MDX mice with 10 mg/kg ENTR-201 once per week for four weeks also
5 resulted in significant splicing correction of dystrophin mRNA and dystrophin protein levels in various muscle tissues (heart, diaphragm, quadricep, and transverse abdominis (TrA)). In comparison to treatment with 20 mpk PMO, treatment with ENTR-201 at 10 mg/kg results in a higher amount of both splicing correction and dystrophin protein in all four muscle tissues. Notably, the mRNA correction and dystrophin protein expression in the heart are only observed in
10 10 mpk ENTR-201 treated MDX mice, not in 20 mg/kg PMO treated MDX mice. The findings from IHC study was also consistent with RT-PCR and WB analysis. Treatment with 10 mpk ENTR-201 once per week for four weeks also normalized the serum creatine kinase level, which is a muscle damage biomarker, suggesting that Oligo 201 treatment reduces muscle fiber damage in a DMD mouse model. PMO (ENTR-0013) treatment alone in contract did not significantly
15 reduce the elevated serum CK level. Serum samples were collected one week after the last injection from repeated dosing study. Analysis of CK levels was performed using commercially available CK measurement kit (Millipore Sigma chemicals, MAK116) as per instructions from the manufacturer. Quantification of dystrophin protein showed nearly 40 % cells are positive for dystrophin in cardiac tissue compared to 5% or less in vehicle treated or PMO alone treated cardia
20 tissues.

[0644] Treatment of mice with 20 mg/kg Oligo 201 one time per week resulted in splicing correction of dystrophin in the heart and in the diaphragm. Treatment of mice with 10 mg/kg Oligo 201 or 5 mg/kg Oligo 201 four times per week also resulted in splicing correction of dystrophin in the heart and in the diaphragm. Treatment with Oligo 201 at 10 mg/kg results in a higher amount
25 of splicing correction than treatment with 5 mg/kg in the heart and diaphragm. Treatment with 5 mg/kg or 10 mg/kg Oligo 201 four times per week also resulted in a decrease in creatine kinase expression in comparison to control, suggesting that Oligo 201 treatment reduces muscle fiber damage in patients with DMD.

**Example 5. Tissue modulation in muscle using a cell-penetrating peptide coupled to an
30 oligonucleotide and nuclear localization sequence for splicing correction of exon 23 of DMD in an MDX mouse model**

[0645] The compounds of Table C below include additional non-limiting examples of NLS-containing compounds. Compounds were prepared as described in previous examples.

Table C:

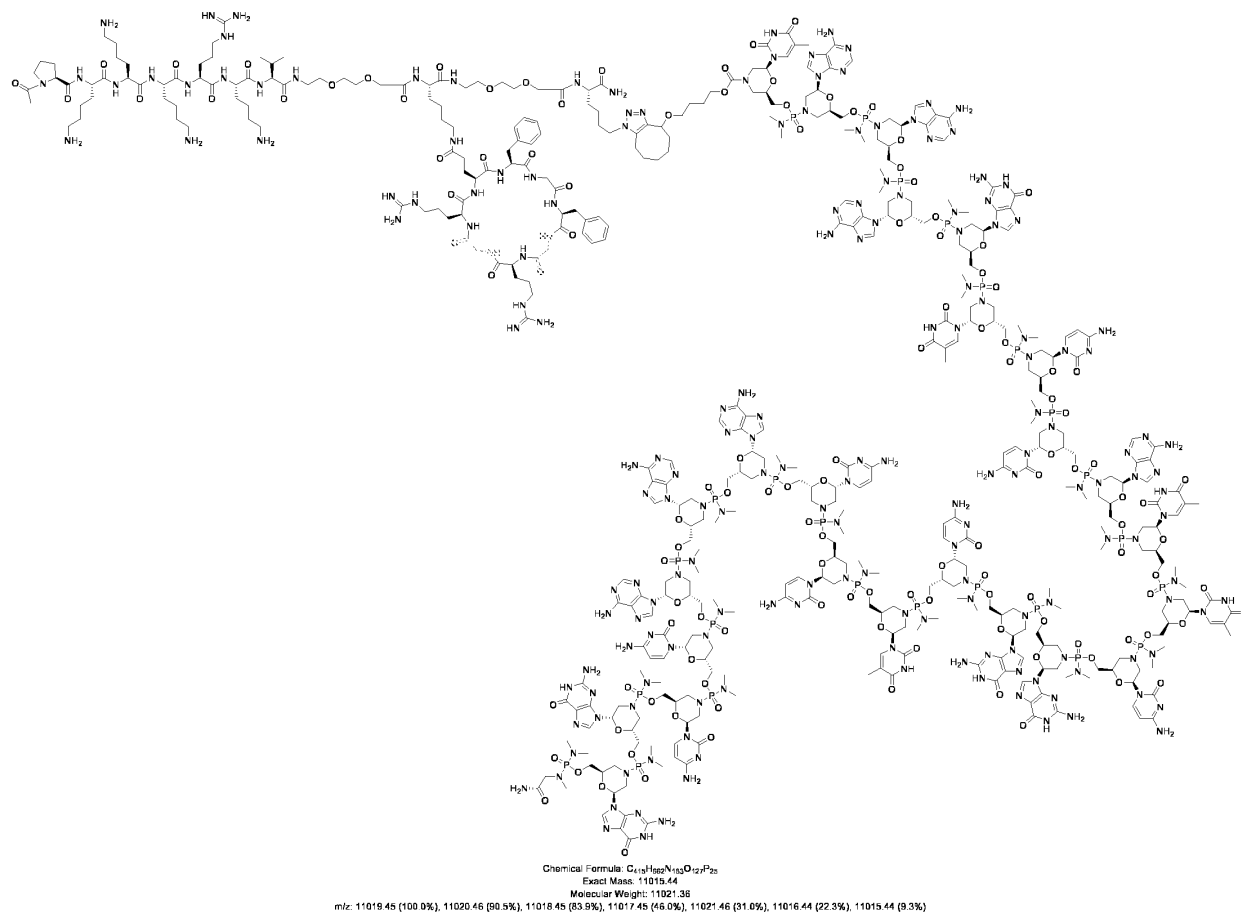
Sequence	Entrada ID
Ac-PKKKRKV-Lys(cyclo[FfΦR-cit-R-cit-Q])-PEG12-K(N3)	ETRD-965
Ac-rr-miniPEG2-Dap[cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-997
Ac-frr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-998
Ac-rfr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-999
Ac-rbfbr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1000
Ac-rrr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1001
Ac-rbr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1002
Ac-rbrbr-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1003
Ac-hh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1004
Ac-hbh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1005
Ac-hbhbh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1006
Ac-rbhbh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1007
Ac-hbrbh-PEG2-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-PEG12-OH	ETRD-1008
Ac-rr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1009
Ac-frr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1010
Ac-rfr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1011
Ac-rbfbr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1012
Ac-rrr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1013
Ac-rbr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1014
Ac-rbrbr-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1015
Ac-hh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1016
Ac-hbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1017
Ac-hbhbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1018
Ac-rbhbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1019
Ac-hbrbh-Dap(cyclo(FfΦ-Cit-r-Cit-rQ))-b-OH	ETRD-1020
Ac-KKKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1045

Ac-KGKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1046
Ac-KKGK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1047
Ac-KKK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1048
Ac-KK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1049
Ac-KGK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1050
Ac-KBK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1051
Ac-KBKBK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1052
Ac-KR-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1053
Ac-KBR-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1054
Ac-PKKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1055
Ac-PKKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1055
Ac-PGKKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1056
Ac-PKGKRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1057
Ac-PKKGRKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1058
Ac-PKKKGKV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1059
Ac-PKKKRGV-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1060
Ac-PKKKRKG-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1061
Ac-KKKRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1062
Ac-KKRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1063
Ac-KRK-miniPEG2-Lys(cyclo(Ff-Nal-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1064
Ac-KKKK-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1092
Ac-KBKBK-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1093
Ac-PKKKRKV-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1094
Ac-KGK-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1095
Ac-KBK-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1096
Ac-KGKK-miniPEG2-Lys(cyclo(FGFGRGRQ))-miniPEG2-K(N3)-NH2	ETRD-1097
Ac-KKKK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1098
Ac-KBKBK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1099

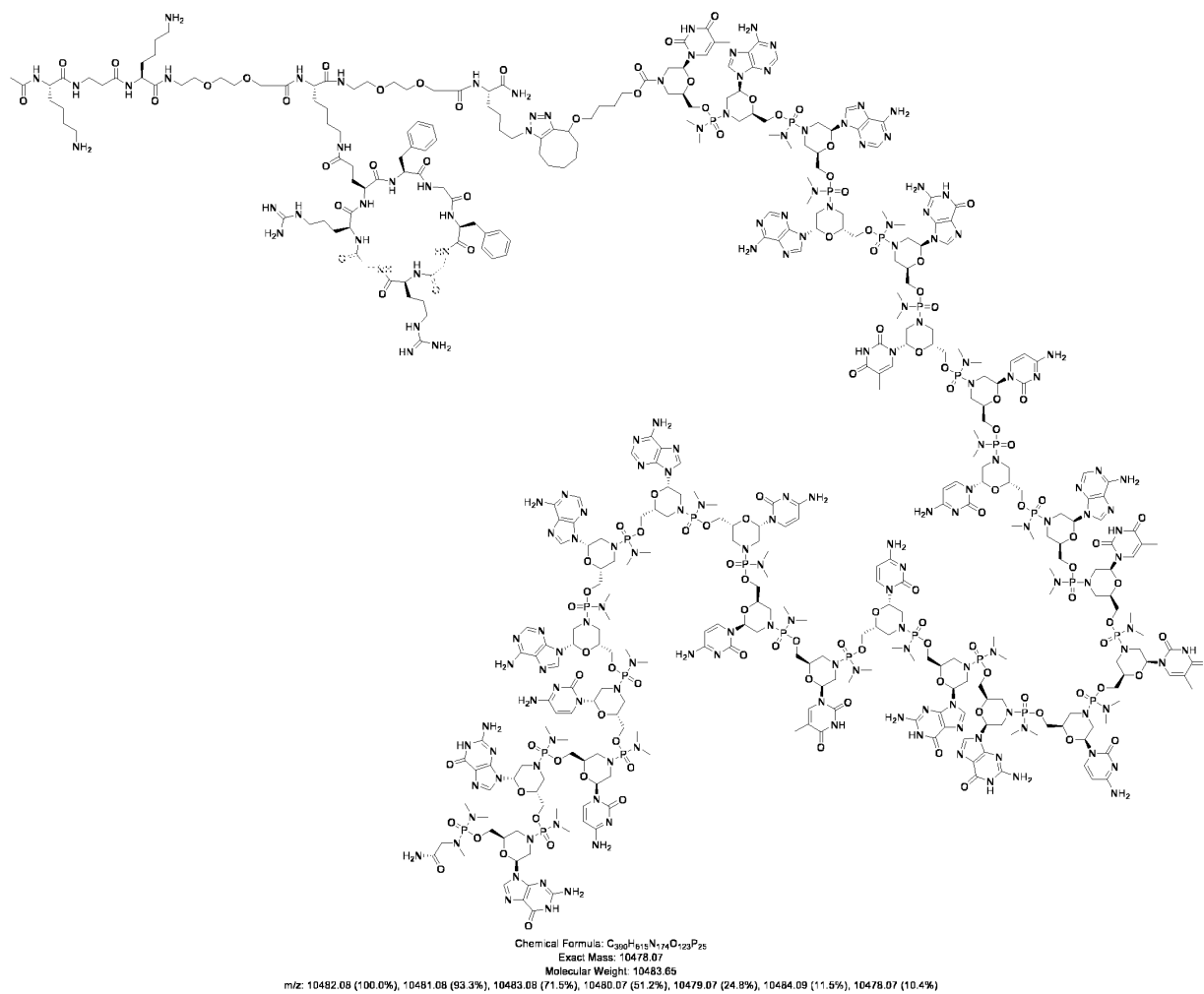
Ac-PKKKRKV-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1100
Ac-KGK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1101
Ac-KBK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1102
Ac-KGKK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1103
Ac-KKKK-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1104
Ac-KBKBK-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1105
Ac-PKKKRKV-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1106
Ac-KGK-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1107
Ac-KBK-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1108
Ac-KGKK-miniPEG2-Lys(cyclo(FfF-GrGrQ))-miniPEG2-K(N3)-NH2	ETRD-1109
Ac-PKKKRKV-miniPEG2-Lys(cyclo(FGFGRGRQ))-PEG12-K(N3)-NH2	ETRD-1131
Ac-PKKKRKV-miniPEG2-Lys(cyclo(GfF-GrGrQ))-PEG12-K(N3)-NH2	ETRD-1132
Ac-KKKK-miniPEG2-Lys(cyclo(FGFGRGRQ))-PEG12-K(N3)-NH2	ETRD-1133
Ac-KKKK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-PEG12-K(N3)-NH2	ETRD-1134
Ac-PKKKRKV-miniPEG2-Lys(cyclo(FGFGRGRQ))-PEG12-OH	ETRD-1135
Ac-PKKKRKV-miniPEG2-Lys(cyclo(GfF-GrGrQ))-PEG12-OH	ETRD-1136
Ac-KKKK-miniPEG2-Lys(cyclo(FGFGRGRQ))-PEG12-OH	ETRD-1137
Ac-KKKK-miniPEG2-Lys(cyclo(GfF-GrGrQ))-PEG12-OH	ETRD-1138

[0646] Non-limiting examples of CCP-AC chemical structures that further comprise a modulatory peptide (MP; or NLS) are shown:

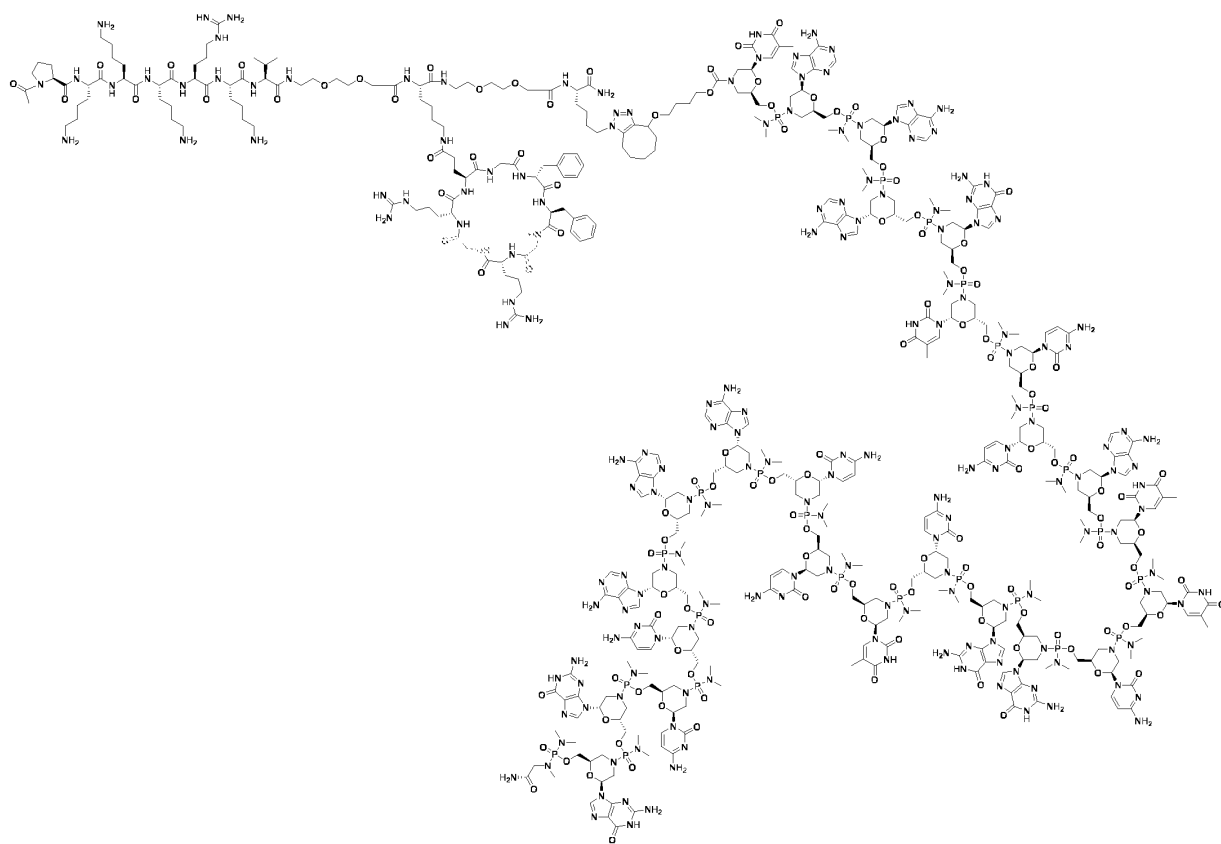
EEV-PMO-MDX23-1



EEV-PMO-MDX23-2



EEV-PMO-MDX23-3

Chemical Formula: $\text{C}_{415}\text{H}_{663}\text{N}_{135}\text{O}_{127}\text{P}_{35}$

Exact Mass: 11015.44

Molecular Weight: 11021.36

 m/z : 11019.45 (100.0%), 11020.46 (90.5%), 11018.45 (63.9%), 11017.45 (46.0%), 11021.46 (31.0%), 11016.44 (22.3%), 11015.44 (8.3%)

[0647] Exemplary compounds were tested in the MDX model described in Example 4. Mice were treated with a single intravenous dose at 20 mg/kg on day 1. On day 5 post-injection, animals were sacrificed and the specified tissues were harvested and flash-frozen. RNA was extracted and exon skipping was quantified by RT-PCR as previously described in diaphragm (**FIG. 12A**), heart (**FIG. 12B**), tibialis anterior (**FIG. 12C**) and triceps (**FIG. 12D**). The results indicated that treatment with NLS-containing compounds (ENTR-1491093, ENTR-1491094, ENTR-1491096) resulted in a lower level of exon skipping in heart tissue as compared to the three the types of muscle tissue examined. Additionally, on day 5 post-injection, dystrophin levels were quantified by Wester blot analysis in diaphragm (**FIG. 13A**), heart (**FIG. 13B**) and tibialis anterior (**FIG. 13C**). The results indicated that treatment with NLS-containing compounds (ENTR-1491093, ENTR-1491096) also resulted in lower levels dystrophin expression in heart tissue and tibialis anterior tissue as compared to the diaphragm tissue.

Example 6A. Use of cell-penetrating peptides conjugated to oligonucleotides for CD33 knockout in human macrophage cells.

[0648] The compounds of Table D were prepared. Exemplary experiments are described below and in Examples 6B and 6C.

Table D. Compounds that target CD33

Oligo#	Design	Note
PMO-CD33	5'-GTA ACT GTA TTT GGT ACT TCC-3'-primary amine (all PMO monomers)	PMO ^{CD33} Human CD33 exon 2 skipping
EEV-PMO- CD33-2	5'-GTA ACT GTA TTT GGT ACT TCC-3'-PEG12CPP12 (all PMO monomers)	CPP-PMO ^{CD33} Human CD33 exon 2 skipping

EEV-PMO-CD33-3	5'-CTG TAT TTG GTA CTT-3' + CPP12-PEG12-K(CPP12)-PEG12-TFP (all PMO monomers)	CPP-CPP-PMO ^{CD33} Human CD33 exon 2 skipping
EEV-PMO-CD33-4	5'-GTA ACT GTA TTT GGT ACT TCC-3' + CPP12-PEG12-K(CPP12)-PEG12-TFP (all PMO monomers)	CPP-CPP-PMO ^{CD33} Human CD33 exon 2 skipping
EEV-PMO-CD33-5	5'-GTA ACT GTA TTT GGT ACT TCC-3' (Ac-NLS-Lys(CPP12)-PEG12-K(N3)- (all PMO monomers)	CPP-NLS-PMO ^{CD33} Human CD33 exon 2 skipping

[0649] Cells. Differentiated THP-1 cells (human monocyte cells) and glioblastoma cells (human neuronal cells) were used in this study.

[0650] Study design. CD33 is implicated in diseases such as cancer and Alzheimer's Disease ("AD"). Targeting CD33 expression represents a treatment strategy for AD and cancer.

[0651] Targeting CD33 expression represents a treatment strategy for AD and cancer. Skipping of exon 2 of the gene expressing CD33 produces D2-CD33, a CD33 isoform that lacks a binding domain of sialic acid. In the absence of such a ligand binding domain, CD33 cannot inhibit microglial activation and phagocytosis of amyloid beta by microglial cells. Such a result is protective against AD. This example evaluated the efficacy of the platform described in Examples 1-5 for treating AD or cancer. Briefly, THP1 and glioblastoma cells were treated with AC having a nucleic acid sequence of 5'-GTAAGTGTATTTGGTACTTCC-3' ("PMO-CD33"), PMO-CD33 conjugated to a CPP ("EEV-PMO-CD33-2"), or PMO-CD33 conjugated to both CPP and NLS (EEV-PMO-CD33-5), in the presence of 10 % fetal bovine serum (FBS).

[0652] PMO sequence development and optimization. The nucleic acid sequence 5'-CTGTATTTGGTACTT-3' has previously been reported to induce human CD33 exon2 skipping

in THP1 cells (Bergeijk P. et al. Molecular and Cellular Biol. 2019). We first modified the oligonucleotide chemistry from 2'-MOE modified RNA to phosphorodiamidate morpholino oligomers (PMO) as in the conjugated construct ENTR-085 but with moderate success. To improve efficacy, we further developed a 21nt-long PMO, PMO-CD33, TABLE 5, 5'-GTAAGTGTATTTGGTACTTCC-3' and its CPP conjugate (iEEV-PMO-CD33-4) showed superior efficacy. Thus, PMO sequence PMO-CD33 was used for subsequent studies.

[0653] Detection of exon 2 skipping by RT-PCR and flow cytometry. Reverse Transcription followed by semi-quantitative PCR analysis revealed that treatment of THP1 cells for 48 hours in the presence of 10 % FBS with EEV-PMO-CD33-4 resulted in skipping of exon 2 and the production of D2-CD33, a CD33 isoform that lacks a ligand binding domain. Treatment of THP1 cells with PMO-CD33 alone resulted in a lower amount of exon skipping in comparison to treatment with EEV-PMO-CD33-4. Exon 2 skipping was dependent on the dose of EEV-PMO-CD33-4. Flow cytometry revealed reduced production of CD33 in cells treated with EEV-PMO-CD33-4 in comparison to untreated (NT) cells.

[0654] Dose-dependent Exon 2 skipping induced by EEV-PMO-CD33-5 in THP1 cells CD33 mRNA from differentiated THP1 cells (human monocyte cells) with various concentrations of EEV-PMO-CD33-5, PMO-CD33 with Endoport (6μL/mL) transfection reagent, or PMO-CD33 alone for 48 hours in the presence of 10% FBS were analyzed by RT-PCR. Result shows dose-dependent skipping of exon 2 CD33 by EEV-PMO-CD33-5 treatment, with significant improvement over transfection (over 2-fold) and over 1000-fold improvement compared to unconjugated PMO-CD33. CD33 mRNA of glioblastoma cells (human neuronal cell line, U-87 MG) treated with various concentrations of EEV-PMO-CD33-5 for 48 hours in the presence of 10% FBS were analyzed by RT-PCR. Result shows dose-dependent skipping of exon 2 CD33 by EEV-PMO-CD33-5 treatment.

[0655] Duration of effects of EEV-PMO-CD33 in differentiated THP1 cells. Differentiated THP1 cells (human monocyte cells) treated with EEV-PMO-CD33-4 for 1 day, and cells were continued to be cultured with full growth medium. 2-8 days post incubation, cells were harvested and the CD33 mRNA were analyzed. Result shows that uptaken EEV-PMO-CD33-4 can induce CD33 exon 2 skipping for a sustained period of time (> 8 days).

[0656] Exon 2 skipping induced by monovalent EEV-PMO-CD33 and bivalent EEV-PMO-CD33 CD33 mRNA of differentiated THP1 cells (human monocyte cells) treated by PMO-CD33,

monovalent EEV-PMO-CD33-2 and bivalent EEV-PMO-CD33-4 for 48 hours were analyzed by RT-PCR. Result shows effect of EEV-PMO-CD33-4 is more potent than EEV-PMO-CD33-2 in inducing CD33 exon 2 skipping.

Example 6B. Use of cell-penetrating peptides conjugated to oligonucleotides for CD33

5 knockout in human macrophage cells.

[0657] This example used the protocols of Example 6A to evaluate an AC having a nucleic acid sequence of 5'-GTAAGTGTATTTGGTACTTCC-3' ("PMO-CD33") or PMO^{CD33} conjugated to a CPP ("EEV-PMO-CD33-2") in the presence of 10 % fetal bovine serum (FBS). The CPP used was cCPP12, which has an amino acid sequence of FfΦRrRr.

10 [0658] **Detection of exon 2 skipping by RT-PCR and flow cytometry.** RT-PCR analysis revealed that treatment of THP1 cells for 48 hours in the presence of 10 % FBS with EEV-PMO-CD33-2 resulted in skipping of exon 2 and the production of D2-CD33, a CD33 isoform that lacks a ligand binding domain. Treatment of THP1 cells with PMO-CD33 alone resulted in a lower amount of exon skipping in comparison to treatment with EEV-PMO-CD33-2. Exon 2 skipping
15 was dependent on the dose of EEV-PMO-CD33-2. Flow cytometry revealed reduced production of CD33 in cells treated with EEV-PMO-CD33-2 in comparison to untreated (NT) cells.

Example 6C. Use of cell-penetrating peptides conjugated to oligonucleotides for production of D2-CD33 in nonhuman primates.

[0659] The EEV-PMO-CD33-2, and 5 described in Examples 6A-C is used animal studies, e.g. in
20 rodents, monkeys, and humans. Animals or humans are administered intravenously or intrathecally via various doses (0.5, 1, 2.5, 5, 10, 20, 40 mpk) of a EEV-PMO-CD33 or PMO-CD33 conjugates that targets exon skipping (exon 2 of human CD33 or exon 5 of monkey CD33) of the gene expressing CD33. The results show that the oligonucleotide therapeutics induce exon skipping of target CD33 gene, downregulate CD33 level and can treat AD.

25 [0660] **NHP Study design.** To explore the tolerability of EEV-PMO on non-human primate (NHP), two CPP-PMO constructs as well as PMO itself are dosed in staggered fashion at 2 mpk and 5 mpk through intravenous infusion on d0 and d3, respectively. Oligonucleotides include "PMO-CD33", PMO-CD33 conjugated to a CPP (EEV-PMO-CD33-2), or PMO-CD33 conjugated to both CPP and NLS (EEV-PMO-CD33-5), and are formulated in saline (0.9% w/v sodium chloride). The CPP used was CPP12, which has a cyclic amino acid sequence of FfΦRrRr. PBMCs
30 (peripheral blood mononuclear cells) were isolated at 1 day (d4) and 7 days (10) post 5 mpk

injection to detect splicing correction. Blood, serum and urine samples were collected at 1 day and or 7 days post each injection for hematology, clinical chemistry, coagulation, urinalysis, cytokine and histamine analysis.

[0661] Detection of exon exclusion by RT-PCR. Although human and non-human primates share high sequence homology of CD33 gene, the 5'-UTR and splicing pattern is different. The sequence coding for the IgV domain of CD33 is located in exon 2 for human and located in exon 5 for non-human primate CD33 gene. Therefore, the skipping of exon 2 in human CD33 (D2-CD33), resulting in Δ IgV-CD33 protein, corresponds to the skipping of exon 5 of non-human primate CD33. Monkey PBMC was collected at specified time points mentioned above. Total RNA was extracted, and RT-PCR was conducted using forward primer 5'-CTCAGACATGCCGCTGCT-3' and reverse primer 5'-TTGAGCTGGATGGTTCTCTCCG-3' resulting in full length CD33 mRNA (FL-CD33) at 700bp and Exon-5 skipped CD33 mRNA (D5-CD33) at 320bp. Reverse Transcription followed by semi-quantitative PCR analysis revealed that treatment of monkey PBMC cells showed that IV administration of EEV-PMO-CD33-2 and EEV-PMO-CD33-5, but not PMO, resulted in skipping of exon 5 of CD33 gene and the production of D2-CD33. And the activity of both EEV-PMO-CD33-2 and EEV-PMO-CD33-5 last for at least 7 days post treatment.

[0662] Example 7. Tissue modulation in the central nervous system using a cell-penetrating peptide coupled to an oligonucleotide and nuclear localization sequence administered intrathecally
Objective: The objective of this study was to evaluate the tolerability and CNS tissue distribution of PMO-containing test articles, including ones containing a modulatory peptide (MP; or NLS), administered to rats by intrathecal injection.

[0663] General Methodology: Fifteen (15) male Sprague Dawley rats with JVCs were obtained from Envigo. Animals were assigned into seven (7) treatment groups plus one (1) spare animal. All groups were dosed 50 μ L per animal by intrathecal injection. Group 1 received vehicle. Groups 2-1 and 2-2 received PMO-CD33 at 10 and 25 μ g per animal, respectively. Groups 3-1 and 3-2 received EEV-PMO-CD33-2 at 10 and 25 μ g per animal, respectively. Groups 4-1 and 4-2 received EEV-PMO-CD33-1 at 10 and 25 μ g per animal, respectively. All animals were dosed once on Day 0.

Interim bloods were collected at pre-dose, 0.5, 2, 6, 10, and 24 hours post dose administration. Blood was processed to plasma and stored frozen at nominally -70°C. Terminal procedures were performed on Day 1, approximately 24 hours after dosing. All animals were euthanized by CO₂ asphyxiation followed by thoracotomy and terminal blood collection via cardiac puncture.

5 Maximum obtainable volume of whole blood was collected into lithium heparin tubes and processed to plasma. Plasma was analyzed for clinical chemistries by the Testing Facility. Residual plasma was stored frozen. Following euthanasia, maximum obtainable volume of CSF was collected and stored frozen. Brain (dissected into cerebellum, cortex, hippocampus, hypothalamus, and olfactory bulbs), spinal cord, and DRGs were removed and frozen individually. All frozen
10 samples were stored at nominally -70°C.

[0664] Compound Synthesis: EEV-PMO-CD33-1 was synthesized according to the following procedure (see also **FIG. 14A**). PMO-CD33 with the following sequence (5'-GTA ACT GTA TTT GGT ACT TCC-3'-primary amine), was reacted with cyclooctyne-PEG4-PFP carbonate to obtain cyclooctyne modified PMO. Briefly, to a solution of PMO-CD33 in 100 mM NaHCO₃ (1 eq, 5
15 mM, 500 µL) was added a solution of cyclooctyne-PEG4-PFP in DMF (3-4 eq, 500 µL). Reaction mixture was vortexed, centrifuged and incubated for 2 hours at room temperature. Reaction was monitored by LCMS (Q-TOF), using BEH C18 column (130Å, 1.7 µm, 2.1mm×150 mm), buffer A: water, 0.1% FA), buffer B: acetonitrile, 0.1% FA), flow rate (0.3 mL/min), starting with 2% buffer B and ramping up to 70% for 11 min for a total of 20 min run. Upon completion, reaction
20 mixture was loaded onto a pre-equilibrated PD-10 desalting column and product was eluted using 1.5 mL nuclease free buffer. The azide-EEV (1.5 eq) was added to the PMO-PEG4COT solution as solids. The mixture was vortexed to dissolve the EEV-PMO-CD33-1, centrifuged to settle the solution, and incubated at room temperature for 8-12 hours for completion as confirmed by LCMS (Q-TOF). The crude mixture was diluted with DMSO, loaded onto a C18 reverse-phase
25 column (150mm* 21.2 mm), flow rate of 20 mL/min and purified by a gradient of 5-25% over 20 min using water with 0.1% TFA and acetonitrile as solvents. Desired fractions were pooled, pH of the solution was adjusted to 7 by 1M NaOH and the solution underwent the lyophilization process, affording white lyophilized powder. Product underwent salt exchange to chloride ion by reconstitution of EEV-PMO-CD33-1 in 1M NaCl in water. Solution was transferred into a pre-
30 equilibrated amicon tube 3KD and centrifuged at 3500 rpm for 20 min. This process was repeated twice followed by the three Saline treatments (0.9% NaCl, sterile, endotoxin-free) until the

conductivity of last filtrate reached to conductivity of reference saline solution. The solution was further diluted with saline to the desired formulation concentration and sterile filtered in a biosafety cabinet. The concentration of each formulation was remeasured post filtration according to the table below. The purity and identity of each formulation was assessed by LCMS (QTOF).

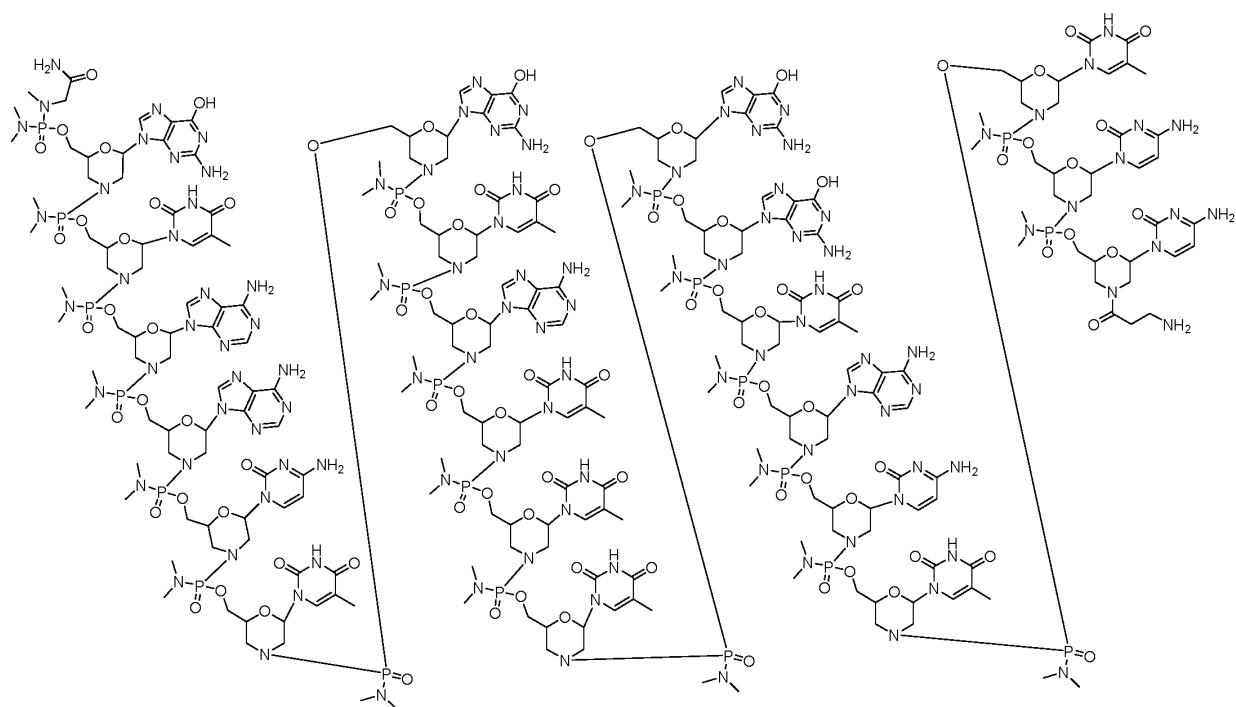
- 5 Formulations were further assayed for their endotoxin amount, residual free peptide, TFA content and pH.

[0665] EEV-PMO-CD33-2 was synthesized according to the following procedure (see also **FIG. 14B**). PMO-CD33 with the following sequence (5'-GTA ACT GTA TTT GGT ACT TCC-3'-primary amine), was reacted with EEV using PYAOP as coupling reagent. Briefly, EEV(1.5 eq), PYAOP (1.5 eq) and DIPEA (3 eq) were dissolved in NMP and reacted for 2 minutes. The solution was then added to the PMO-CD33 (1 eq) in DMSO with the final PMO concentration of 50 mg/mL. Reaction mixture was vortexed, centrifuged and incubated for 2 hours at room temperature. Reaction was monitored by LCMS (Q-TOF), using BEH C18 column (130Å, 1.7 µm, 2.1mm×150 mm), buffer A: water, 0.1% FA), buffer B: acetonitrile, 0.1% FA), flow rate (0.3 mL/min), starting with 2% buffer B and ramping up to 70% for 11 min for a total of 20 min run. Upon completion, reaction mixture loaded onto a C18 reverse-phase column (150mm* 21.2 mm), flow rate of 20 mL/min and purified by a gradient of 5-25% over 20 min using water with 0.1% TFA and acetonitrile as solvents. Desired fractions were pooled, pH of the solution was adjusted to 7 by 1M NaOH and the solution underwent the lyophilization process, affording white lyophilized powder.

15

20 Product underwent salt exchange to chloride ion and formulated as described above.

[0666] The chemical structures of PMO-CD33, EEV-PMO-CD33-1 and EEV-PMO-CD33-2 are shown below:

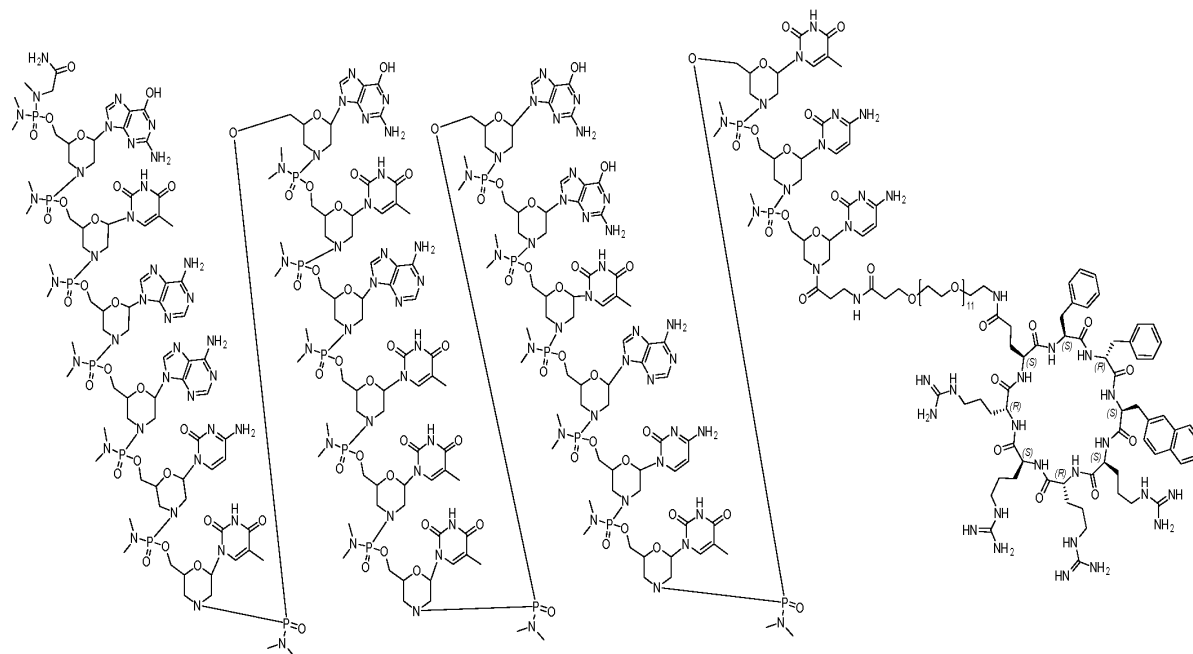
**PMO-CD33**

Molecular Weight: 7171.06



EEV-PMO-CD33-1

Molecular Weight: 10332.99



EEV-PMO-CD33-2

Molecular Weight: 8998.32

[0667] Experimental Design: Animals were treated as shown in the table below:

Group	Animal No.	Test Article	[0002] Dose Level μg/An	Dose Volume (μL/An)	Dosing Regimen	Terminal Time Point
1	2	Vehicle	0	50	Day 0 by IT Injection	[0003] Day 1 (24 Hours Post Dose)
2-1	2	PMO-CD33	10			
2-2	2	PMO-CD33	25			
3-1	2	EEV-PMO-CD33-2	10			
3-2	2	EEV-PMO-CD33-2	25			
4-1	2	EEV-PMO-CD33-1	10			
4-2	2	EEV-PMO-CD33-1	25			

No = Number; An = Animal; IT = Intrathecal

5

[0668] Bioanalytical Sample Analysis: Tissues were thawed, weighed, and homogenized (w/v, 1/5) with RIPA buffer spiked with 1x protease inhibitor cocktail. The tissue homogenates were centrifuged at 5000 rpm for 5minutes at 4°C. The supernatants were precipitated with a mixture of H₂O, Acetonitrile and MeOH, and centrifuged at 15000 rpm for 15minutes at 4°C. The supernatants were transferred to an injection plate for LC-MS/MS analysis. The dynamic range of the LC-MS/MS assay was 25 to 50,000 ng/g tissue.

10

[0669] Results: After administration of PMO (antisense compound (AC) alone) and EEV-PMO, individual rat body weight was measured immediately as well as 24 hours post dose. No significant reduction in body weight was observed post EEV-PMO-CD33-1, EEV-PMO-CD33-2and PMO-CD33 administration compared to Vehicle control. At 30min and 8 hours post dose, rat clinical features were observed, and body temperature was measured. No severe adverse effect was observed. Clinical chemistries measuring liver and kidney toxicity (Albumin, Albumin-Globulin ratio, Alkaline Phosphatase, Alanine Aminotransferase, Aspartate Aminotransferase, Blood Urea nitrogen, Calcium, Cholesterol, Creatine kinase, and Creatinine) were also evaluated 24 hours post IT injection. No significant toxicity was detected by clinical chemistry evaluation in EEV-PMO-CD33-1, EEV-PMO-CD33-2and PMO-CD33 treated rats compared to Vehicle control.

15

20

[0670] After treatment, various rat brain tissue section (cerebellum, cortex, hippocampus, olfactory bulb) was collected and frozen 24 hours post IT injection. Rat brain tissues were pooled

together and homogenized. The tissue homogenate was analyzed using LC-MS/MS to quantify the amount of PMO, EEV-PMO-CD33-2, and EEV-PMO-CD33-1 detected in the various tissue sections of the rat brain. Both EEV-PMO-CD33-2 and EEV-PMO-CD33-1 showed increase uptake in brain tissue compared to PMO-CD33 alone (**FIG. 15**). EEV-PMO-CD33-1 was detected at higher concentration in rat brain compared to PMO-CD33 alone (12-fold increase in cerebellum, 4-fold increase in cortex, 15-fold increase in hippocampus, and 11-fold increase in olfactory bulb) as well as EEV-PMO-CD33-2 (9-fold increase in cerebellum, 1.5-fold increase in hippocampus, 6.7-fold increase in olfactory bulb). In particular, EEV-PMO-CD33-1 exhibited much higher expression in the cerebellum and the olfactory bulb as compared to PMO alone or EEV-PMO-CD33-2

5 [0671] Rat spinal Cord, dorsal root ganglion (DRG), and cerebrospinal fluid (CSF) were also collected 24 hours post IT injection. Concentration of PMO-CD33, EEV-PMO-CD33-2, and EEV-PMO-CD33-1 in rat spinal cord, DRG, and CSF were measured by LC-MS/MS. EEV-PMO-CD33-1 was detected at higher concentration compared to equal dose of PMO-CD33 alone and EEV-PMO-CD33-2 in spinal cord, DRG, and CSF 24 hours post IT injection (**FIG. 16A-C**). EEV-PMO was

10 detected at higher concentration in both spinal cord (**FIG. 16A**, 40x increase) and DRG (**FIG. 16B**, 60x increase) compared to PMO alone 24 hours post injection. In particular, the NLS-containing EEV-PMO-CD33-1 exhibited much higher expression in the spinal cord, DRG and CSF as compared to PMO alone or an EEV-PMO lacking the NLS at a comparable dose.

[0672] Taken together, the LC-MS data demonstrated superior delivery efficiency into CNS for

20 EEV-PMO-CD33-1, which contains a modulatory peptide (MP; or NLS) as compared to both PMO alone and EEV-PMO-CD33-2, which does not contain a modulatory peptide.

[0673] **Conclusions:** In conclusion, these data indicate that these conjugated PMOs administered to rats by intrathecal injection were well tolerated and well retained within the CNS. Significantly, EEV-PMO-CD33-1 is delivered in the central nervous system, including cortex, hippocampus,

25 olfactory bulb, cerebellum, spinal cord and DRG. EEV-conjugation enhanced PMO delivery into CNS compared to PMO alone or EEV-conjugates lacking an NLS.

CLAIMS

1. A compound comprising:
 - (a) a cell penetrating peptide (CPP);
 - 5 (b) a therapeutic moiety (TM); and
 - (c) an exocyclic peptide (EP), wherein tissue distribution or retention of the compound is modulated as compared to a compound comprising the CPP and TM but lacking the EP.
- 10 2. The compound of claim 1, wherein the EP is conjugated to the CPP.
3. The compound of claim 1, wherein the EP is conjugated to the TM.
4. The compound of any one of claims 1-3, wherein the CPP is a cyclic cell penetrating
15 peptide (cCPP).
5. The compound of any one of claims 1-4, wherein the EP comprises at least one positively charged amino acid residue.
- 20 6. The compound of claim 5, wherein the EP comprises at least one lysine residue and/or at least one arginine residue.
7. The compound of claim 5, wherein the EP comprises at least one lysine residue.
- 25 8. The compound of claim 5, wherein the EP comprises the sequence PKKKRKV.
9. The compound of claim 5, wherein the EP comprises a sequence selected from the group consisting of KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKBK, KKKRKV,
30 PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV and PKKKRKG.

10. The compound of any one of claims 1-4, wherein the EP is a nuclear localization signal (NLS).
11. The compound of any one of claims 1-10, wherein the therapeutic moiety is a protein, a polypeptide, a small molecule or an oligonucleotide other than an antisense compound.
12. The compound of any one of claims 1-10, wherein the therapeutic moiety is an antisense compound (AC).
13. The compound of claim 12, wherein the AC comprises at least one modified nucleotide or nucleic acid selected from a phosphorothioate (PS) nucleotide, a phosphorodiamidate morpholino nucleotide, a locked nucleic acid (LNA), a peptide nucleic acid (PNA), a nucleotide comprising a 2'-O-methyl (2'-OMe) modified backbone, a 2'-O-methoxy-ethyl (2'-MOE) nucleotide, a 2',4' constrained ethyl (cEt) nucleotide, and a 2'-deoxy-2'-fluoro-beta-D-arabinonucleic acid (2'F-ANA), and wherein hybridization of the AC with the target sequence reduces or prevents splicing, inhibits or regulates translation, mediates degradation, or blocks expansions of nucleotide repeats.
14. The compound of claim 12, wherein the AC comprises small interfering RNA (siRNA), microRNA (miRNA), ribozymes, immune stimulating nucleic acids, antisense, antagomir, antimir, microRNA mimic, supermir, U1 adaptor, aptamer, or a CRISPR gene-editing machinery.
15. The compound of any one of claims 12-14, wherein the cyclic peptide is conjugated to the 3' end of the AC.
16. The compound of any one of claims 12-14, wherein the cyclic peptide is conjugated to the 5' end of the AC.
17. The compound any one of claims 12-14, wherein the cyclic peptide is conjugated to the backbone of the AC.

18. The compound of claims any one of claims 12-17, comprising a linker, which conjugates the cyclic peptide to the AC.
19. The compound of claim 18, wherein the linker is covalently bound to the 5' end of the AC.
20. The compound of claim 18, wherein the linker is covalently bound to the 3' end of the AC.
21. The compound of claims 18, wherein the linker is covalently bound to the backbone of the AC.
22. The compound of any one of claims 18-21, wherein the linker is covalently bound to the side chain of an amino acid residue on the cyclic peptide.
23. The compound of any one of claims 18-22, wherein the linker comprises a bivalent or trivalent C₁-C₅₀ alkylene, wherein 1-25 methylene groups are optionally and independently replaced by -N(H)-, -N(C₁-C₄ alkyl)-, -N(cycloalkyl)-, -O-, -C(O)-, -C(O)O-, -S-, -S(O)-, -S(O)₂-, -S(O)₂N(C₁-C₄ alkyl)-, -S(O)₂N(cycloalkyl)-, -N(H)C(O)-, -N(C₁-C₄ alkyl)C(O)-, -N(cycloalkyl)C(O)-, -C(O)N(H)-, -C(O)N(C₁-C₄ alkyl), -C(O)N(cycloalkyl), aryl, heteroaryl, cycloalkyl, or cycloalkenyl.
24. The compound of any one of claims 18-22, wherein the linker comprises:
- (i) one or more D or L amino acid residues, each of which is optionally substituted;
 - (ii) optionally substituted alkylene;
 - (iii) optionally substituted alkenylene;
 - (iv) optionally substituted alkynylene;
 - (v) optionally substituted carbocyclyl;
 - (vi) optionally substituted heterocyclyl;
 - (vii) -(J-R¹)_z-, wherein each R¹ is independently alkylene, alkenylene, alkynylene, carbocyclyl, or heterocyclyl, each J is independently NR³, -NR³C(O)-, S, or O, wherein each R³ is

independently H, alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, wherein each alkyl, alkenyl, alkynyl, carbocyclyl, and heterocyclyl is optionally substituted, and z is independently an integer from 1 to 20;

(viii) $-(J-R^2)_x-$, wherein each R^2 at each instance, is independently alkylene, alkenylene, alkynylene, carbocyclyl, or heterocyclyl, each J, at each instance, is independently NR^3 , $-NR^3C(O)-$, S, or O, wherein each R^3 is independently H, alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, wherein each alkyl, alkenyl, alkynyl, carbocyclyl, and heterocyclyl is optionally substituted, and x is independently an integer from 1 to 20; or
(ix) a combination thereof.

10

25. The compound of claim 24, wherein the linker comprises:

(i) a residue of lysine, glycine, b-alanine, 4-aminobutyric acid, 5-aminopentanoic acid, 6-aminopentanoic acid, or a combination thereof;

(ii) $-(J-R^1)_z$;

15 (iii) $-(J-R^2)_x$, or

(iv) a combination thereof.

20

26. The compound of claim 24 or 25, wherein each R^1 and R^2 are independently alkylene, each J is O, each x is independently an integer from 1 to 20, and each z is independently an integer from 1 to 20.

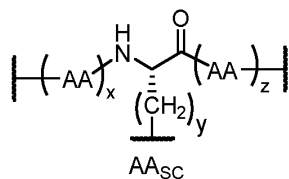
27. The compound of claim 18, wherein the linker comprises:

(i) a $-(OCH_2CH_2)_z-$ subunit, wherein z is an integer from 2 to 20;

25 (ii) a residue of glycine, b-alanine, 4-aminobutyric acid, 5-aminopentanoic acid, 6-aminopentanoic acid, or a combination thereof; or

(iii) combinations of (i) and (ii).

28. The compound of any one of claims 18-27, wherein the linker has a structure:



wherein:

each AA is independently a residue of glycine, b-alanine, 4-aminobutyric acid, 5-aminopentonic acid or 6-aminopentanoic acid;

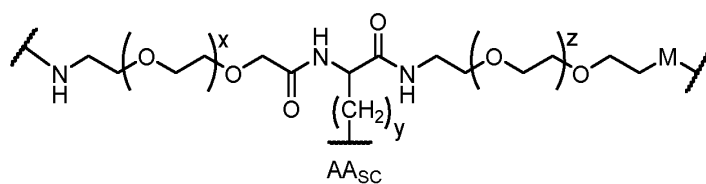
- 5 AA_{SC} is a side chain of an amino acid on the cyclic peptide;

x is an integer from 1 to 10;

y is an integer from 1 to 5; and

z is an integer from 1 to 10.

- 10 29. The compound of any one of claims 18-27, wherein the linker has the structure:



wherein:

x is an integer from 1 to 20;

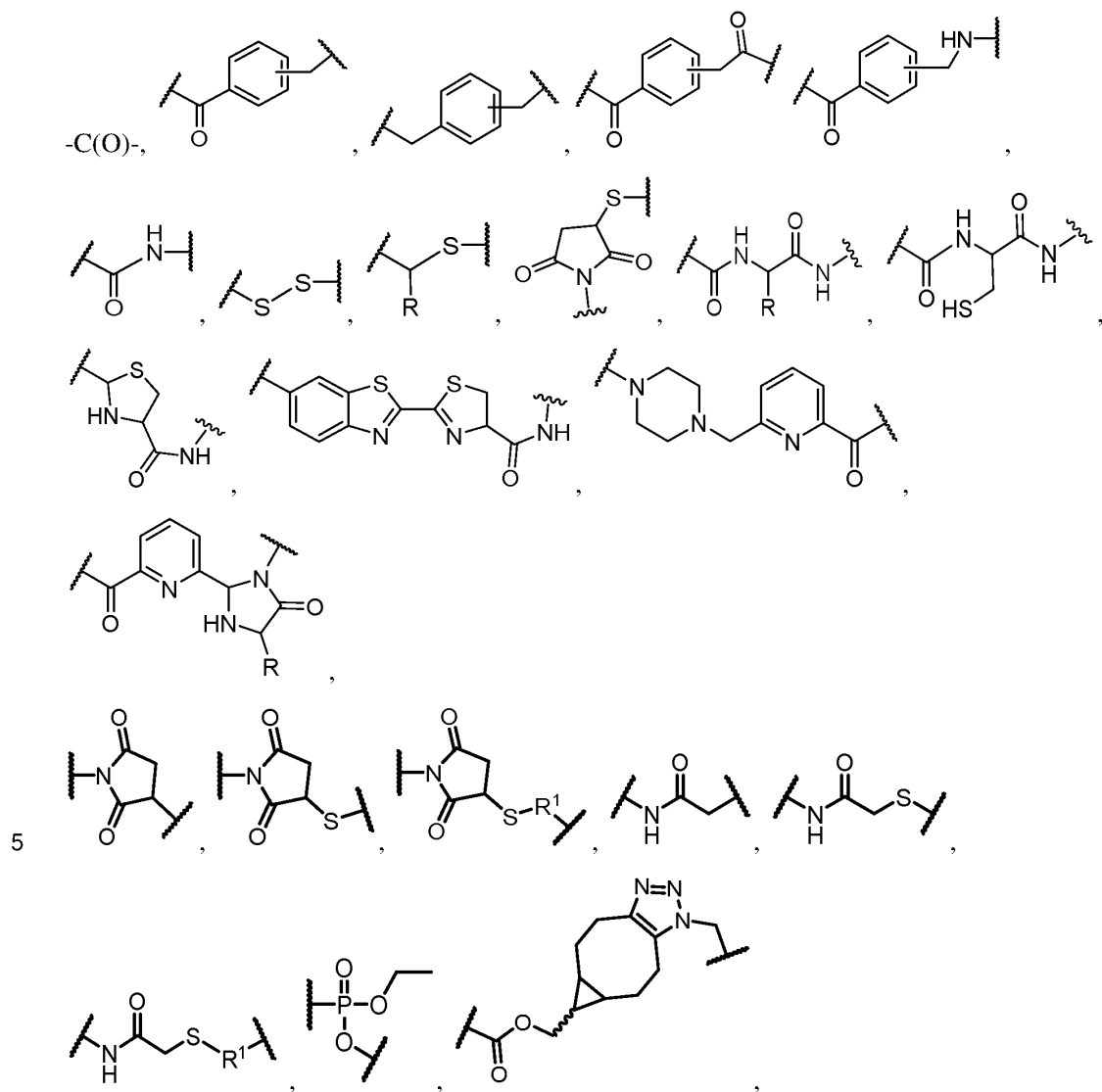
- 15 y is an integer from 1 to 5;

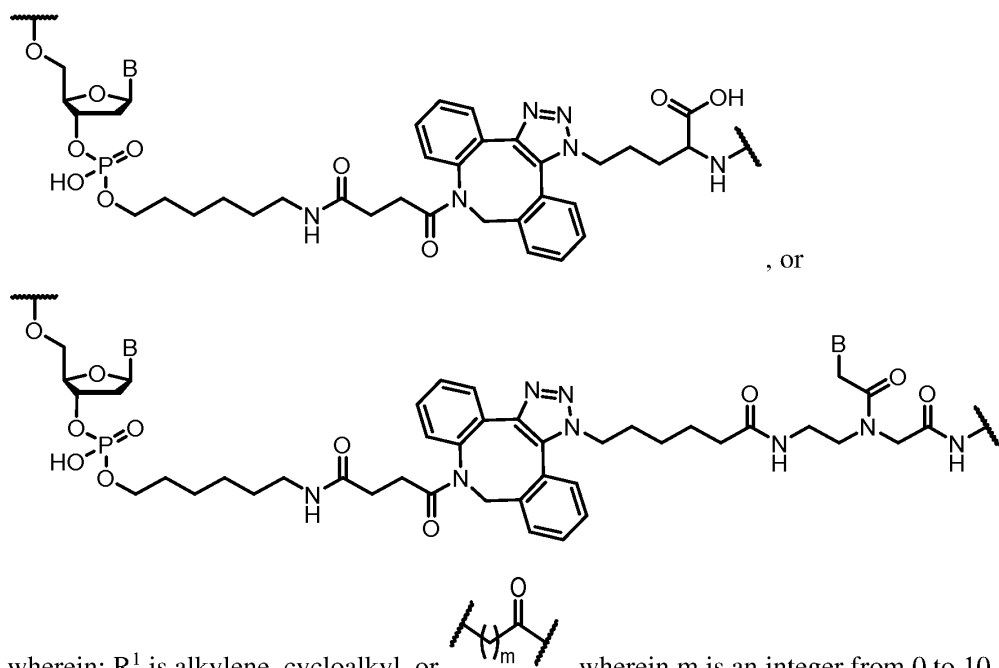
z is an integer from 2 to 20;

M is a bonding moiety; and

AA_{SC} is a side chain of an amino acid on the cyclic peptide.

- 20 30. The compound of claim 29, wherein M is:

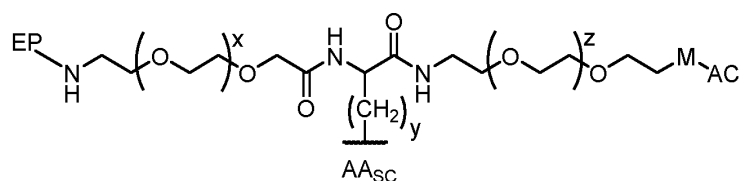




wherein: R^1 is alkylene, cycloalkyl, or $(CH_2)_m$, wherein m is an integer from 0 to 10, wherein each R is independently an alkyl, alkenyl, alkynyl, carbocyclyl, or heterocyclyl, and wherein
 5 each B is independently selected from a nucleobase.

31. The compound of claim 30, wherein M is $-C(O)$.
32. The compound of any one of claims 29-31, wherein z is 11.
- 10 33. The compound of any one of claims 29-32, wherein x is 1.
34. The compound of any one of the preceding claims, comprising an exocyclic peptide (EP) conjugated to the cyclic peptide of any one of the claims or the linker of the claims.
- 15 35. The compound of any one of claims 29-34, comprising an exocyclic peptide (EP) conjugated to an amino group of the linker.
36. The compound of claim 34 or 35, wherein the EP comprises from 2 to 10 amino acid
 20 residues.

37. The compound of any one of claims 34-36, wherein the EP comprises from 4 to 8 amino acid residues.
- 5 38. The compound of any one of claims 34-37, wherein the EP comprises 1 or 2 amino acids comprising a side chain comprising a guanidine group, or a protonated form thereof.
39. The compound of any one of claims 34-38, wherein the EP comprises 1, 2, 3, or 4 lysine residues.
- 10 40. The compound of claim 39, wherein the amino group on the side chain of each lysine residue is substituted with a trifluoroacetyl (-COCF₃) group, allyloxycarbonyl (Alloc), 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde), or (4,4-dimethyl-2,6-dioxocyclohex-1-ylidene-3)-methylbutyl (ivDde) group.
- 15 41. The compound of any of claims 34-40, wherein the EP comprises at least 2 amino acid residues with a hydrophobic side chain.
42. The compound of claim 41, wherein the amino acid residue with a hydrophobic side
20 chain is valine, proline, alanine, leucine, isoleucine, or methionine.
43. The compound of any one of claims 34-42, wherein the exocyclic peptide comprises one of the following sequences: PKKKRKV; KR; RR; KKK; KGK; KBK; KBR; KRK; KRR; RKK; RRR; KKKK; KKRK; KRKK; KRRK; RKKR; RRRR; KGKK; KKGK; KKKKK; KKKRK;
25 KBK BK; KKKRKV; PGKKRKV; PKGKRKV; PKKGRKV; PKKKGKV; PKKKRGV; or PKKKRKG.
44. The compound of any one of claims 34-43, wherein the exocyclic peptide has the structure: Ac-P-K-K-K-R-K-V-.
- 30 45. The compound of any one of the preceding claims comprising the following structure:



wherein:

x is an integer from 1 to 20;

5 y is an integer from 1 to 5;

z is an integer from 1 to 20;

EP is an exocyclic peptide;

M is a bonding moiety;

AC is an antisense compound; and

10 AA_{SC} is an amino acid residue of the cyclic peptide.

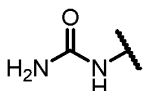
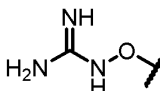
46. The compound of any one of claims 1-45, wherein the cyclic peptide comprises from 4 to 20 amino acid residues in the cyclic peptide, wherein at least two amino acid residues comprise a guanidine group, or a protonated form thereof, and at least two amino acid residues

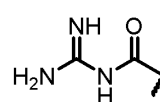
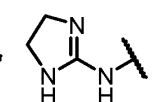
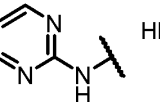
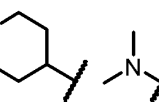
15 independently comprise hydrophobic side chains.

47. The compound of any one of claims 1-46, wherein the cyclic peptide comprises 2, 3, or 4 acid residues comprising a guanidine group, or a protonated form thereof.

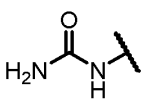
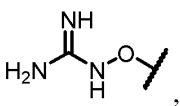
20 48. The compound of claim 46 or 47, wherein the cyclic peptide comprises 2, 3, or 4 amino acid residues comprising hydrophobic side chains.

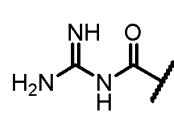
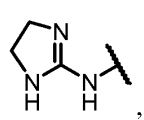
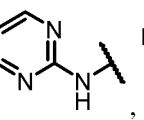
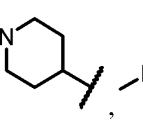
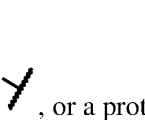
49. The compound of any one of claims 46-48, wherein the cyclic peptide comprises at least

one amino acid comprising a side chain selected from , ,

25 , , , , or a protonated form thereof.

50. The compound of any one of claims 46-49, wherein the cyclic peptide comprises 1, 2, 3,

or 4 amino acids comprising a side chain selected from , ,

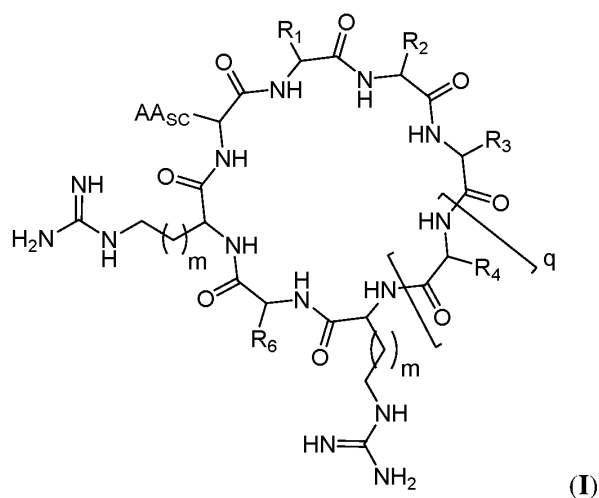
, , , , , or a protonated form thereof.

5

51. The compound of any one of claims 46-50, wherein the cyclic peptide comprises at least one glycine residue.

52. The compound of any one of claims 46-51, wherein the cyclic peptide comprises 1, 2, 3,
10 or 4 glycine residues.

53. The compound of any one of the preceding claims, wherein the cyclic peptide has a structure of Formula (I):



15 or a protonated form thereof,

wherein:

R₁, R₂, and R₃ can each independently be H or an amino acid residue having a side chain comprising an aromatic group;

at least one of R_1 , R_2 , and R_3 is an aromatic or heteroaromatic side chain of an amino acid;

R_4 and R_6 are independently H or an amino acid side chain;

AA_{sc} is an amino acid side chain;

q is 1, 2, 3 or 4; and

5 each m is independently an integer 0, 1, 2, or 3.

54. The compound of claim 53, wherein R_4 is H or an amino acid residue comprising a side chain comprising an aromatic group.

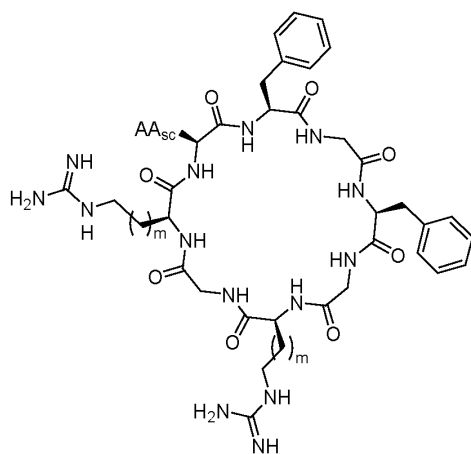
10 55. The compound of claim 53 or 54, wherein the amino acid residue comprising a side chain comprising an aromatic group is phenylalanine.

56. The compound of any one of claims 53-55, wherein two of R_1 , R_2 , R_3 , and R_4 comprise phenylalanine.

15

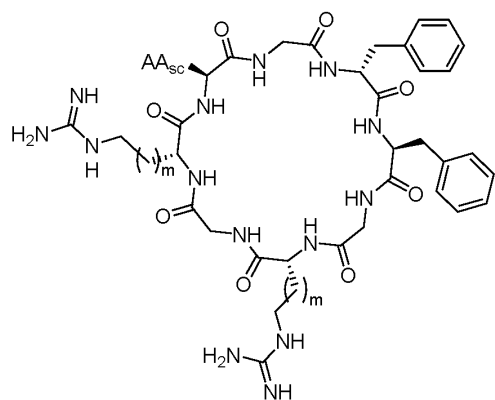
57. The compound of any one of claims 53-56, wherein two of R_1 , R_2 , R_3 , and R_4 are H.

58. The compound of any one of claims 53-57, wherein the cyclic peptide has a structure of Formula (I-1), (I-2), (I-3), or (I-4):

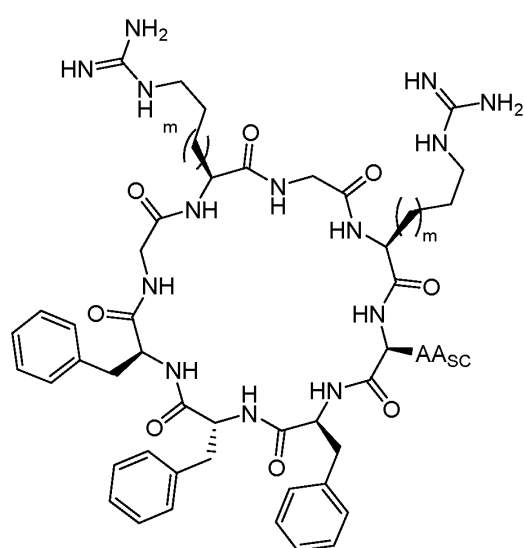


20

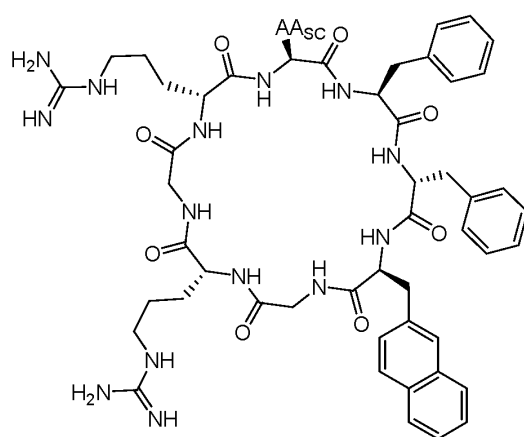
(I-1),



(I-2),



(I-3)



(I-4)

or a protonated form thereof,

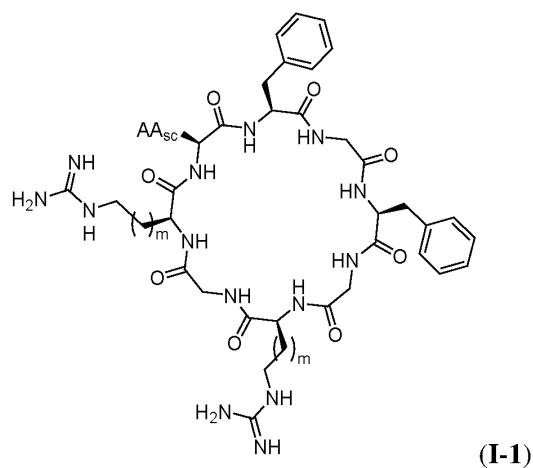
wherein:

AA_{SC} is an amino acid side chain; and

each m is independently an integer from 0 to 3.

5

59. The compound of any one of claims 53-58, wherein the cyclic peptide has a structure of Formula (V-1):



or a protonated form thereof,

10 wherein:

AA_{SC} is an amino acid side chain ; and

each m is independently an integer from 0 to 3.

60. The compound of any one of claims 53-59, wherein AA_{SC} comprises a side chain of an asparagine residue, aspartate residue, glutamine residue, a glutamate residue, homoglutamate residue, or a homoglutamine residue.

61. The compound of any one of claims 53-59, wherein AA_{SC} comprises a side chain of a glutamine residue.

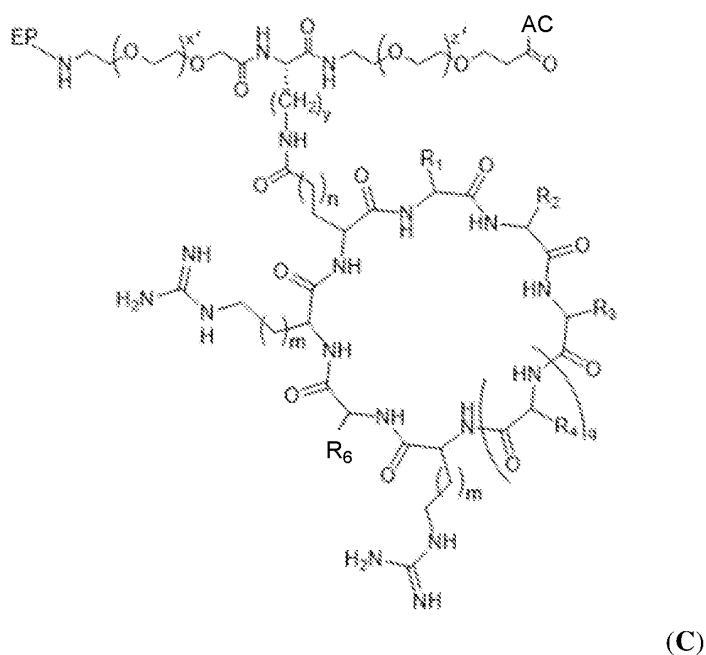
20

62. The compound of any one of claims 1-61, wherein the cyclic peptide comprises FGFGGRGR.

63. The compound of any one of claims 1-61, wherein the cyclic peptide comprises
5 GfFGrGr.

64. The compound of any one of claims 1-61, wherein the cyclic peptide comprises FfΦGRGR.

10 65. The compound of any one of claims 1-64, comprising a structure of Formula (C):



or a protonated form thereof,

wherein:

15 R₁, R₂, and R₃ can each independently be H or an amino acid residue having a side chain comprising an aromatic group;

R_4 and R_6 are independently H or an amino acid side chain;

EP is an exocyclic peptide as defined herein;

each m is independently an integer from 0-3;

n is an integer from 0-2;

5 x' is an integer from 2-20;

y is an integer from 1-5;

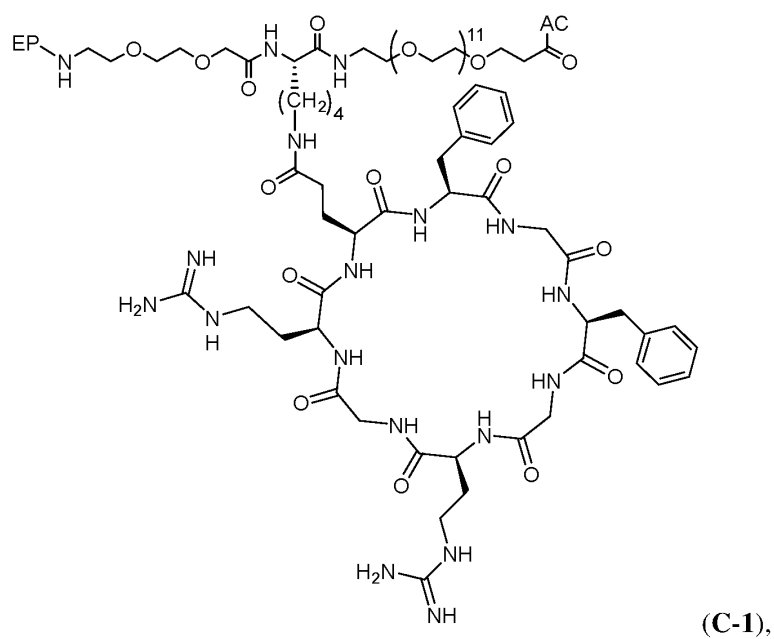
q is an integer from 1-4;

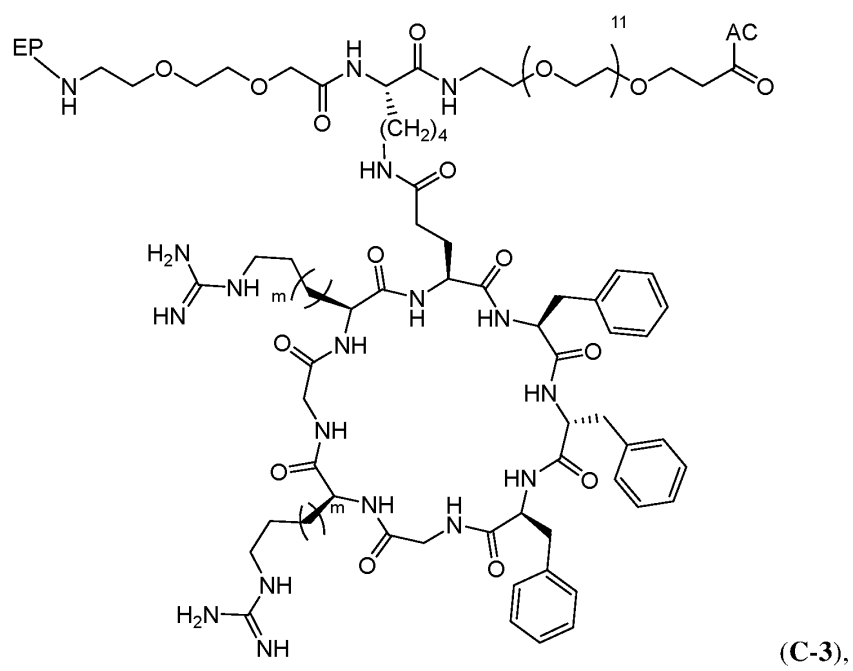
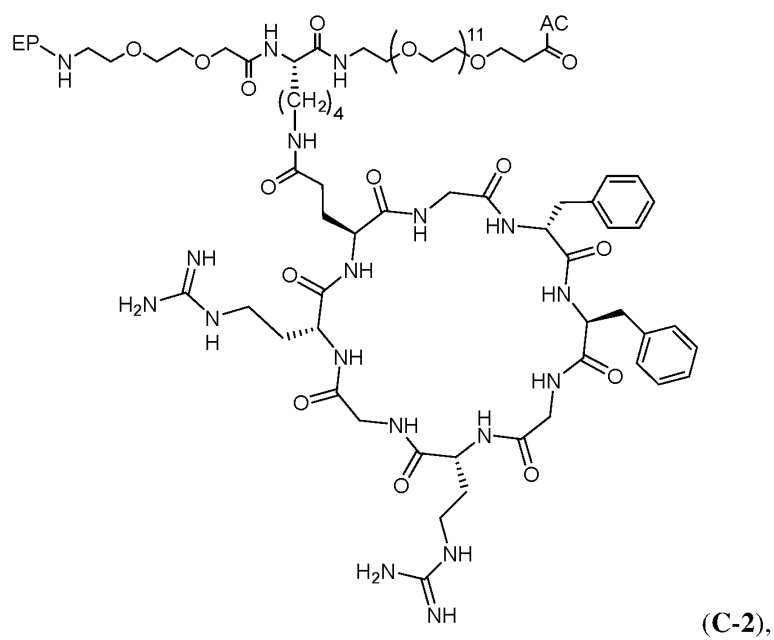
z' is an integer from 2-20;

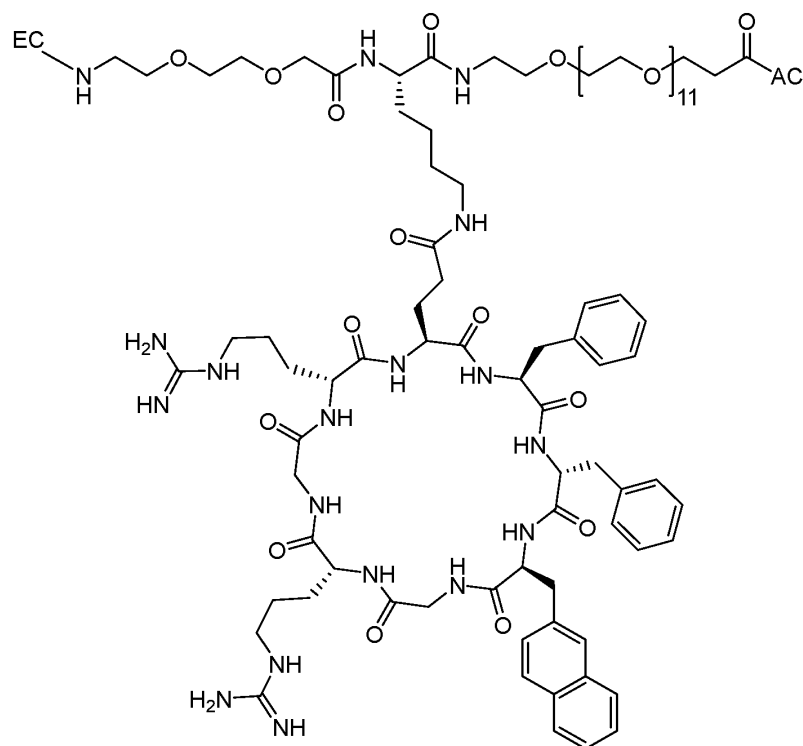
wherein the AC is an oligonucleotide.

10

66. The compound of any one of claims 1-64, having the structure of Formula (C-1), (C-2), (C-3), or (C-4):







(C-4)

or a protonated form thereof;

67. The compound of any one of the preceding claims, wherein the compound modulates
 5 splicing of exon 2, 8, 11, 17, 19, 23, 29, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 51, 52, 53, 55, and 59 of DMD.

68. The compound of any one of the preceding claims, wherein the compound modulates
 10 splicing of exon 2, 8, 11, 23, 43, 44, 45, 50, 51, 53, and 55 of DMD.

69. The compound of any one of the preceding claims, wherein the compound modulates
 splicing of exon 2, 23, 44, or 51 of DMD.

70. The compound of any one of the preceding claims, wherein the compound modulates
 15 splicing of exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 7a, and exon 7b of CD33.

71. The compound of any one of the preceding claims, wherein the C-terminus of the EP sequence is conjugated to the CPP.

72. The compound of claim 1, which is selected from the group consisting of EEV-PMO-MDX1,2,3, EEV-PMO-CD33-1, and the compounds shown in Table C of Example 5.

73. A pharmaceutical composition comprising the compound of any one of claims 1-72.

74. A cell comprising a compound of any one of claims 1-72.

10

75. A method of modulating tissue distribution or retention of a therapeutic agent in a subject in need thereof, comprising administering a compound of any one of claims 1-72.

76. The method of claim 75, wherein the compound is administered to the subject intrathecally and the compound modulates tissue distribution or retention of the therapeutic agent in tissues of the central nervous system (CNS).

15

77. The method of claim 75, wherein the compound modulates tissue distribution or retention of the therapeutic agent in muscle tissue.

20

78. A method of treating a disease or disorder in a subject in need thereof, comprising administering a compound of any one of claims 1-72.

79. The method of claim 78, wherein the therapeutic agent is an antisense compound (AC) and administration of the compound modulates splicing or expression of a target gene, degrades mRNA, stabilizes mRNA, or sterically blocks mRNA.

25

80. The method of claim 78, wherein administration of the compound modulates splicing of the target pre-mRNA.

30

81. The method of any one of claims 78-80, wherein the disease or disorder is a central nervous system disorder, a neuromuscular disorder, or a musculoskeletal disorder.
82. The method of any one of claims 78-80, wherein the disease or disorder is Duchenne muscular dystrophy, β thalassemia, dystrophin Kobe, osteogenesis imperfect, cystic fibrosis, Merosin-deficient congenital muscular dystrophy type 1A, or spinal muscular atrophy.
83. The method of claim 82, wherein the disease is Duchenne muscular dystrophy.
84. The method of any one of claims 78-80, wherein the disease or disorder is Fragile X, Friedreich's ataxia (FRDA), Huntington's Disease (HD), Myotonic dystrophy type 1 (DM1), Myotonic dystrophy type 2 (DM2), Spinal and bulbar muscular atrophy (SBMA), Spinal cerebellar ataxia type 1 (SCA1), Spinal cerebellar ataxia type 2 (SCA2), or Spinal cerebellar ataxia type 3 (SCA3).
85. The method of claim 84, wherein the disease or disorder is Fragile X, Friedreich's ataxia (FRDA), Huntington's Disease (HD), Myotonic dystrophy type 1 (DM1).
86. The method of any one of claims 78-80, wherein the disease or disorder comprises a neuroinflammatory disease.
87. The method of claim 86, wherein the neuroinflammatory disease comprises Alzheimer's disease or Parkinson's disease.
88. A method of modulating tissue distribution or retention of a therapeutic agent in the central nervous system (CNS) of a subject, the method comprising:
administering intrathecally to the subject a compound comprising:
(a) a cell penetrating peptide (CPP);
(b) a therapeutic moiety (TM) comprising the therapeutic agent; and
(c) an exocyclic peptide (EP) comprising at least one positively-charged amino acid residue,

wherein the amount, expression, function or activity of the therapeutic agent is modulated at least 10% in at least one tissue of the CNS of the subject as compared to a second tissue of the CNS of the subject.

5 89. The method of claim 88, wherein the therapeutic agent is a CD33-targeted antisense compound.

90. The method of claim 89, wherein the CPP is a cyclic CPP.

10 91. The method of claim 88, wherein the EP is selected from the group consisting of KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KK RK, KRKK, KRRK, RKKR, RRRR, KGGK, KKGK, KKKKK, KKKRK, KBKBK, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV and PKKKRKG.

15 92. The method of claim 88, wherein the subject has a central nervous system disease or disorder or a neuroinflammatory disease or disorder.

93. The method of claim 92, wherein the subject has Alzheimer's disease or Parkinson's disease.

20

94. A method of modulating tissue distribution or retention of a therapeutic agent in the muscular system of a subject, the method comprising:
administering to the subject a compound comprising:

(a) a cell penetrating peptide (CPP);

25 (b) a therapeutic moiety (TM) comprising the therapeutic agent; and

(c) an exocyclic peptide (EP) comprising at least one positively-charged amino acid residue,
wherein the amount, expression, function or activity of the therapeutic agent is modulated at least 10% in at least one tissue of the muscular system of the subject as compared to a second tissue of the muscular system of the subject.

30

95. The method of claim 94, wherein the therapeutic agent is a DMD-targeted antisense compound.

96. The method of claim 94, wherein the CPP is a cyclic CPP.

5

97. The method of claim 94, wherein the EP is selected from the group consisting of KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKBK, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV and PKKKRKG.

10

98. The method of claim 94, wherein the subject has a neuromuscular disorder or a musculoskeletal disorder.

99. The method of claim 98, wherein the subject has Duchenne muscular dystrophy.

15