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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
05 October 2023 (05.10.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/192642 A3

(51) International Patent Classification:
A61P 25/00 (2006.01) *A61K 31/4178* (2006.01)
A61K 31/4164 (2006.01) *C07D 403/14* (2006.01)

(21) International Application Number:
PCT/US2023/017201

(22) International Filing Date:
31 March 2023 (31.03.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/326,625 01 April 2022 (01.04.2022) US
63/482,670 01 February 2023 (01.02.2023) US

(71) Applicant: **DESIGN THERAPEUTICS, INC.** [US/US]; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US).

(72) Inventors: **ZHANG, Chengzhi**; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US). **BHAT, Abhijit**; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US). **WAYBRIGHT, Jarod**; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US). **NOURIE, Hannah**; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US). **YANG, Fei**; 6005 Hidden Valley Road, Suite 110, Carlsbad, California 92011 (US).

(74) Agent: **OTRUBOVA, Katerina**; WILSON SONSINI GOODRICH & ROSATI, 650 Page Mill Road, Palo Alto, California 94304 (US).

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

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

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

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

(88) Date of publication of the international search report:
30 November 2023 (30.11.2023)

(54) Title: METHODS AND COMPOUNDS FOR MODULATING HUNTINGTON'S DISEASE

(57) Abstract: The present disclosure relates to transcription modulator molecules having a first terminus, a second terminus, and an oligomeric backbone and methods for treating Huntington's disease.

WO 2023/192642 A3

METHODS AND COMPOUNDS FOR MODULATING HUNTINGTON'S DISEASE

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Application No. 63/326,625, filed April 1, 2022, and U.S. Application No. 63/482,670, filed February 1, 2023, each of which are hereby incorporated by reference in their entirety.

FIELD OF THE DISCLOSURE

[0002] Disclosed herein are new chimeric heterocyclic polyamide compounds and compositions and their application as pharmaceuticals for the treatment of disease. Methods to modulate the expression of a target gene comprising the CAG trinucleotide repeat sequence in a subject are also provided for the treatment of diseases such as Huntington's disease ("HD").

BACKGROUND OF THE DISCLOSURE

[0003] Huntington's disease ("HD") was first identified in the late 19th century as an autosomal dominant, neurodegenerative disorder. The symptoms of HD, which include a range of movement, cognitive and psychiatric disorders, generally appear in adulthood. HD is associated with the presence of the CAG trinucleotide repeat sequence in the *Htt* gene, which codes for a protein termed huntingtin. Subjects with more than about 36 trinucleotide repeat sequences generally present with symptoms of HD, with a larger number of trinucleotide repeat sequences associated with an earlier onset of symptoms. Pathology stems from a cascade of steps: production of poly-Q huntingtin, followed by fragmentation of the elongated huntingtin into smaller peptides, which bind together and accumulate in neurons. The effects of this cascade are pronounced in the basal ganglia and cortex of the brain.

[0004] Huntington's disease-like syndrome refers to a group of ailments whose symptoms are similar to those of Huntington's disease, but which lack the characteristic mutation in the *Htt* gene. Huntington's disease-like 2 syndrome ("HDL2") is associated with a count of about 40 or more CAG trinucleotide repeat sequences in the junctophilin 3 (*Jph3*) gene. HDL2 is a genetic disorder that has been seen in subjects with African lineage. Age of onset is inversely correlated with the number of trinucleotide repeat sequences. Symptoms of this syndrome include dystonia and chorea (uncontrolled movements), emotional disruptions, dysarthria, bradykinesia, inability to incorporate new learning, and difficulty in making decisions. Life expectancy can range from a few years post diagnosis to over a decade. The current theory holds that a poly-Q protein that is coded by the *Jph3* gene forms aggregates in neuronal cells that is responsible for the pathology of the disease. However, evidence suggesting toxic gain-of-function of mRNA has also been uncovered, indicating a possible dual pathway for pathology.

[0005] In some embodiments, the mechanism set forth above provides opportunity for an effective treatment for a disease or disorder which is characterized by the presence of an excessive count of CAG trinucleotide repeat sequences in a target gene. In some embodiments, the pathology of the disease or

disorder is due to the presence of mRNA containing an excessive count of CAG trinucleotide repeat sequences. In some embodiments, the pathology of the disease or disorder is due to the presence of a translation product containing an excessive count of glutamine amino acid residues. In some embodiments, the pathology of the disease or disorder is due to a loss of function in the translation product. In some embodiments, the pathology of the disease or disorder is due to a gain of function in the translation product. In some embodiments, the pathology of the disease or disorder can be alleviated by increasing the rate of transcription of the defective gene. In some embodiments, the pathology of the disease or disorder can be alleviated by decreasing the rate of transcription of the defective gene.

SUMMARY OF THE DISCLOSURE

[0006] This disclosure utilizes regulatory molecules present in cell nuclei that control gene expression. Eukaryotic cells provide several mechanisms for controlling gene replication, transcription, and/or translation. Regulatory molecules that are produced by various biochemical mechanisms within the cell can modulate the various processes involved in the conversion of genetic information to cellular components. Several regulatory molecules are known to modulate the production of mRNA and, if directed to the target gene (such as, *Htt)*, would modulate the production of the target gene mRNA that causes diseases such as, for example, Huntington's disease or Huntington's disease-like syndrome, and thus reverse the progress of these diseases.

[0007] Provided herein are compounds and methods for recruiting a regulatory molecule into close proximity to the target gene comprising a CAG trinucleotide repeat sequence. The compounds disclosed herein contain: (a) a DNA binding moiety that will selectively bind to the target gene, linked to (b) a recruiting moiety that will bind to a regulatory molecule. Without being bound by theory, the compounds may counteract the expression of defective target gene in the following manner:

(1) The DNA binding moiety can bind selectively the characteristic CAG trinucleotide repeat sequence of the target gene;

(2) The recruiting moiety, linked to the DNA binding moiety, can thus be held in proximity to the target gene;

(3) The recruiting moiety, now in proximity to the target gene, can recruit the regulatory molecule into proximity with the gene; and

(4) The regulatory molecule can modulate expression of the target gene and therefore counteract the expression of defective mRNA, by direct interaction with the gene.

[0008] The DNA binding moiety can bind selectively the characteristic CAG trinucleotide repeat sequence of for example, *Htt.* The recruiting moiety, linked to the DNA binding moiety, will thus be held in proximity to the target gene; will recruit the regulatory molecule into proximity with the gene; and the regulatory molecule will modulate expression, and therefore counteract the production of defective target gene by direct interaction with the target gene. This mechanism may provide an effective treatment for HD, which is caused by the expression of defective *Htt*, where correction of the expression of the defective target gene thus represents an effective method for the treatment for these diseases.

[0009] The disclosure further provides for DNA binding moieties that selectively bind to one or more copies of the CAG trinucleotide repeat that are characteristic of the defective target gene. Selective binding of the DNA binding moiety to the target gene, made possible due to the high CAG count associated with the defective target gene, directs the recruiting moiety into proximity of the gene, and recruits the regulatory molecule into position to modulate gene transcription.

[0010] The DNA binding moiety comprises a polyamide segment that will bind selectively to the target CAG sequence. Polyamides designed by for example Dervan (U.S. Patent Nos. 9,630,950 and 8,524,899) and others can selectively bind to selected DNA sequences. These polyamides sit in the minor groove of double helical DNA and form hydrogen bonding interactions with the Watson-Crick base pairs. Polyamides that selectively bind to particular DNA sequences can be designed by linking monoamide building blocks according to established chemical rules. One building block is provided for each DNA base pair, with each building block binding noncovalently and selectively to one of the DNA base pairs: A/T, T/A, G/C, and C/G. Following this guideline, trinucleotides bind to molecules with three amide units, i.e. tri-amides. In general, these polyamides can orient in either direction of a DNA sequence.

[0011] In principle, longer DNA sequences can be targeted with higher specificity and/or higher affinity by combining a larger number of monoamide building blocks into longer polyamide chains. Ideally, the binding affinity for a polyamide would simply be equal to the sum of each individual monoamide/DNA base pair interaction. In practice, however, due to the geometric mismatch between the fairly rigid polyamide and DNA structures, longer polyamide sequences do not bind to longer DNA sequences as tightly as would be expected from a simple additive contribution. The geometric mismatch between longer polyamide sequences and longer DNA sequences induces an unfavorable geometric strain that subtracts from the binding affinity that would be otherwise expected.

[0012] The disclosure provides for transcription modulator molecules that comprise a DNA binding moiety (for example a polyamide comprising multi-amine subunits) that is connected by a spacer (for example a linker moiety or oligomeric backbone) to the protein binding moiety. The spacer can alleviate the geometric strain that would otherwise decrease binding affinity of a larger polyamide sequence.

[0013] Disclosed herein are compounds that comprise a polyamide moiety that can bind to one or more copies of the CAG trinucleotide repeat sequence, and can modulate the expression of a target gene comprising a CAG trinucleotide repeat sequence. Treatment of a subject with these compounds will modulate expression of the defective target gene, and this can reduce the occurrence, severity, or frequency of symptoms associated with disease. Certain compounds disclosed herein will provide higher binding affinity and selectivity than has been observed previously for this class of compounds.

[0014] It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments, are given by way of illustration only, since various changes and modifications within the spirit and scope of the instant disclosure will become apparent to those skilled in the art from this detailed description.

INCORPORATION BY REFERENCE

[0015] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

DETAILED DESCRIPTION

[0016] The disclosure provides for transcription modulator molecules that comprise a DNA binding moiety (for example, a polyamide comprising multi-amine subunits) connected by spacers (for example, a linker moiety or oligomeric backbone) to the protein binding moiety. The spacers can alleviate the geometric strain that would otherwise decrease binding affinity of a larger polyamide sequences.

[0017] Treatment of a subject with these compounds will modulate the expression of the defective target gene, and this can reduce the occurrence, severity, or frequency of symptoms associated with genetic disease (such as for example HD). The compounds described herein recruit the regulatory molecule to modulate the expression of the defective *target* gene and effectively treat and alleviate the symptoms associated with diseases.

Compounds

[0018] The compounds disclosed herein are transcription modulator molecules. They possess useful activity for modulating the transcription of a target gene having one or more CAG repeats (e.g., *Htt*), and may be used in the treatment or prophylaxis of a disease or condition in which the target gene plays an active role. Thus, in broad aspects, some embodiments also provide pharmaceutical compositions comprising one or more compounds disclosed herein together with a pharmaceutically acceptable carrier, as well as methods of making and using the compounds and compositions.

[0019] In an aspect, provided herein is a transcription modulator molecule having a first terminus, a second terminus, and a linker moiety, wherein:

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus.

First terminus – DNA binding moiety

[0020] The first terminus interacts and binds with the gene, particularly with the minor grooves of the CAG sequence. In an aspect, the molecules disclosed herein provide a polyamide sequence for interaction of a single polyamide subunit to each base pair in the CAG repeat sequence. In some embodiments, the molecules provide a turn component (e.g., aliphatic amino acid moiety), in order to enable hairpin binding of the molecule to the CAG, in which each nucleotide pair interacts with two subunits of the polyamide. In some embodiments, one or more of the polyamide backbone carbonyl groups (C=O), is replaced with an

oxetane. In some embodiments, at least one of the polyamide backbone carbonyl groups is replaced with an oxetane.

[0021] In some embodiments, each subunit comprises a moiety that is independently chosen from a heterocycle and an aliphatic chain.

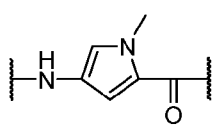
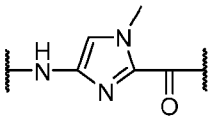
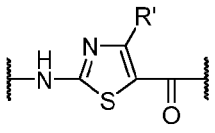
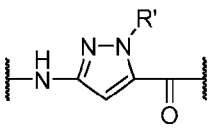
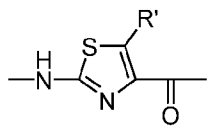
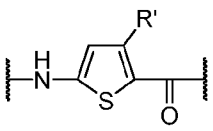
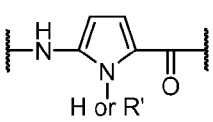
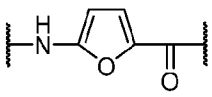
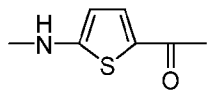
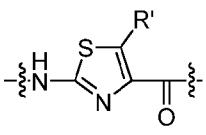
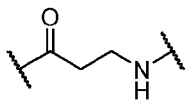
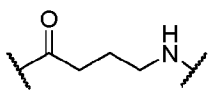
[0022] In some embodiments, the aliphatic chain is a C_1-C_6 straight chain aliphatic chain. In some embodiments, the aliphatic chain has structural formula $-(CH_2)_m-$, for m chosen from 1, 2, 3, 4, and 5. In some embodiments, the aliphatic chain is $-CH_2CH_2-$.

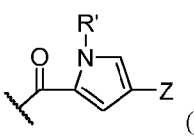
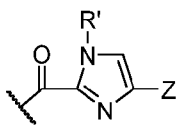
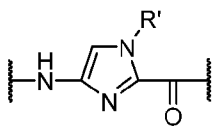
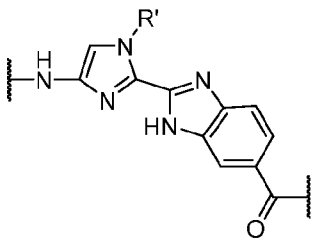
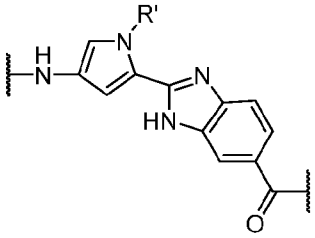
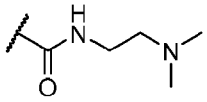
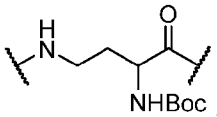
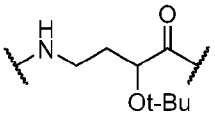
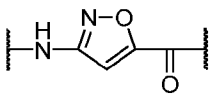
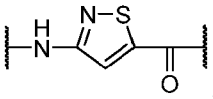
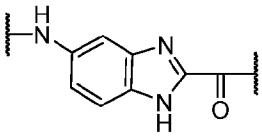
[0023] In some embodiments, the heterocycle is a monocyclic heterocycle. In some embodiments, the heterocycle is a monocyclic 5-membered heterocycle. In some embodiments, each heterocycle contains a heteroatom independently chosen from N, O, or S. In some embodiments, each heterocycle is independently chosen from pyrrole, imidazole, thiazole, oxazole, thiophene, and furan.

[0024] In some embodiments, the DNA-binding moiety comprises –NH-Q-C(O)-, wherein Q is an optionally substituted C_6-C_{10} arylene, optionally substituted 4 to 10-membered heterocyclene, optionally substituted 5 to 10-membered heteroarylene group, or an optionally substituted alkylene group.

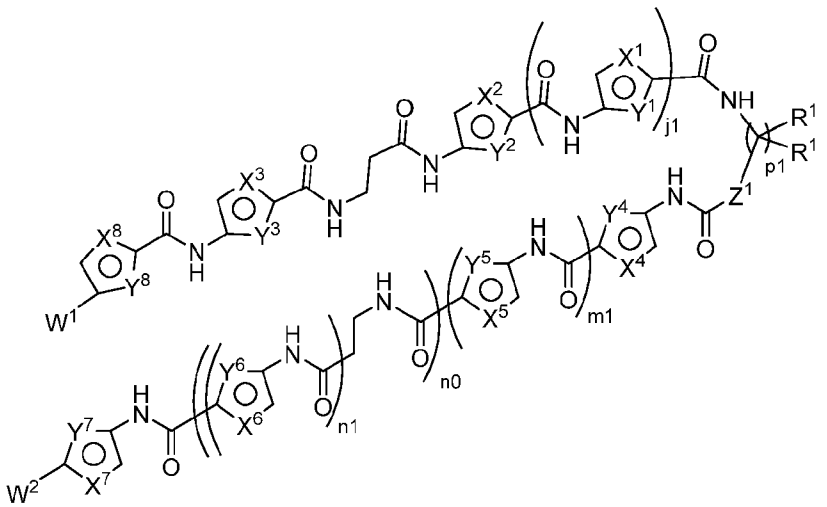
[0025] In some embodiments, the DNA-binding moiety comprises at least three aromatic carboxamide moieties selected to correspond to the nucleotide repeat sequence CAG and at least one aliphatic amino acid residue chosen from the group consisting of glycine, β-alanine, γ-aminobutyric acid, 2,4-diaminobutyric acid, and 5-aminovaleric acid. In some embodiments, the DNA-binding moiety comprises one or more subunits selected from the group consisting of optionally substituted N-methylpyrrole, optionally substituted N-methylimidazole, β-alanine (β), and γ-aminobutyric acid. In some embodiments, the DNA-binding moiety comprises at least one γ-aminobutyric acid.

[0026] In some embodiments, the DNA-binding moiety comprises a polyamide of one or more of the following subunits selected from


(Py),

(Im),

(Th),

(Pz),

(Nt),

(Tn),

(Nh),

(Fr),

(Tp),

(iNt),

(beta or β),

(gAB),


(PyT),

(ImT),

(iIm),

(ImBi),

(PyBi),

(Dp),

,

,

,

, -NH-benzopyrazinylene-C(O)-, -NH-phenylene-C(O)-, -NH-pyridinylene-C(O)-, -NH-piperidinylene-C(O)-, -NH-pyrimidinylene-C(O)-, -NH-anthracenylene-C(O)-, -NH-quinolinylene-C(O)-, and

, wherein each R' is independently hydrogen, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino; and Z is H, NH_2, C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 alkyl-NH_2.

[0027] In some embodiments, the first terminus is a DNA-binding moiety that comprises a structure of Formula (A-1):



Formula (A-1),

or a pharmaceutically acceptable salt thereof, wherein:

Z^1 is absent, -O- or -NH-;

each X^1, X^2, X^3, X^4, X^5, X^6, X^7, and X^8 is independently O, S, or NR^2;

each Y^1, Y^2, Y^3, Y^4, Y^5, Y^6, Y^7, and Y^8 is independently CH or N;

W^1 is deuterium, hydrogen, optionally substituted C_1-C_6 alkyl, (azaneylidene)methanediamine, (azaneylidene)-N,N,N',N'-tetramethylmethanediamine, $-C(O)-NR^{1A}R^{1B}$, $-NR^{1A}-C(O)-NR^{1A}R^{1B}$, $-Z_B-P(O)(OR^{1A})_2$, $-Z_B-(CH_2)_{p3}-PO(OR^{1A})_2$, $-Z_B-(CH_2)_{p3}-O-P(O)(OR^{1A})_2$, wherein

Z_B is -N- or -O-;

p_3 is an integer from 1-10;

W^2 is an optionally substituted C_1-C_6 alkyl or $-C(O)-NR^{1A}R^{1B}$;

each R^1 is independently hydrogen, deuterium, halogen, amino, cyano, optionally C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, or $-NC(O)R^{1A}$; or

two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3 to 6-membered carbocyclic ring or optionally substituted 3 to 6-membered heterocyclic ring;

each R^2 is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino;

each R^{1A} is independently hydrogen, deuterium, or optionally substituted C_1-C_{20} alkyl;

each R^{1B} is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_2-C_{10} heteroalkyl, optionally substituted 5-membered heteroaryl, or $(AA)_{p2}$, wherein

each AA is an amino acid;

p_2 is an integer from 1-10;

j_1 is 0 or 1;

n_0 is 1 or 0;

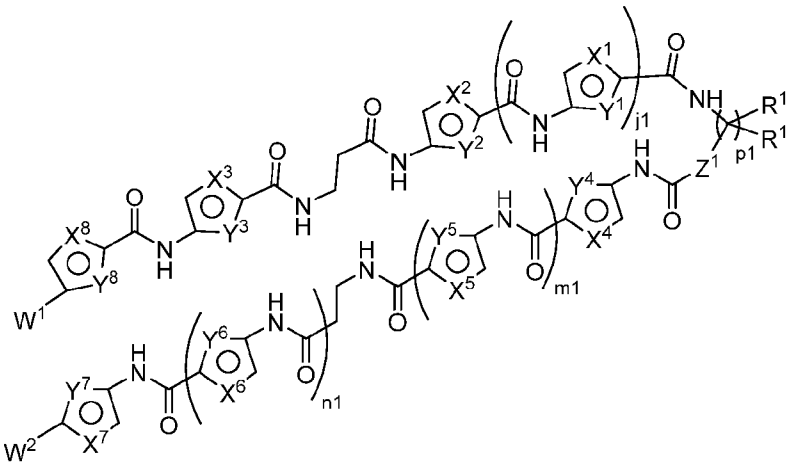
m_1 and n_1 are each independently an integer from 0-3;

p_1 is 2 or 3, provided that when Z^1 is -O- or -NH- then p_1 is 2, and when Z^1 is absent then p_1 is 3; and

wherein one of W^1 or W^2 is attached to the oligomeric backbone.

[0028] In some embodiments, n_0 is 1.

[0029] In some embodiments, the first terminus is a DNA-binding moiety that comprises a structure of Formula (A-2):



Formula (A-2),

or a pharmaceutically acceptable salt thereof, wherein:

Z^1 is absent, -O-, or -NH-;

each X^1, X^2, X^3, X^4, X^5, X^6, X^7, and X^8 is independently O, S, or NR^2;

each Y^1, Y^2, Y^3, Y^4, Y^5, Y^6, Y^7, and Y^8 is independently CH or N;

W^1 is deuterium, hydrogen, optionally substituted C_1-C_6 alkyl, (azaneylidene)methanediamine, (azaneylidene)-N,N,N',N'-tetramethylmethanediamine, -C(O)-$NR^{1A}R^{1B}$, -NR^{1A}-C(O)-$NR^{1A}R^{1B}$, -Z_B-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_{p3}$-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_{p3}$-O-$P(O)(OR^{1A})_2$, wherein

Z_B is -N- or -O-;

p_3 is an integer from 1-10;

W^2 is an optionally substituted C_1-C_6 alkyl or -C(O)-$NR^{1A}R^{1B}$;

each R^1 is independently hydrogen, deuterium, halogen, amino, cyano, optionally C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, or -NC(O)R^{1A}; or

two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3 to 6-membered carbocyclic ring or optionally substituted 3 to 6-membered heterocyclic ring;

each R^2 is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino;

each R^{1A} is independently hydrogen, deuterium, or optionally substituted C_1-C_{20} alkyl;

each R^{1B} is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_2-C_{10} heteroalkyl, optionally substituted 5-membered heteroaryl, or $(AA)_{p2}$, wherein

each AA is an amino acid;

p_2 is an integer from 1-10;

j_1 is 0 or 1;

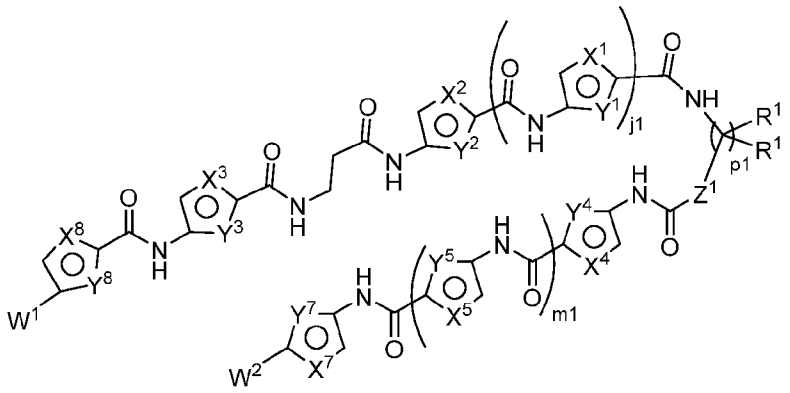
m_1 and n_1 are each independently an integer from 0-3;

p_1 is 2 or 3, provided that when Z^1 is -O -or -NH- then p_1 is 2, and when Z^1 is absent then p_1 is 3; and

wherein one of W^1 or W^2 is attached to the oligomeric backbone.

[0030] In some embodiments, n_0 is 0.

[0031] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-3), or a pharmaceutically acceptable salt thereof:



Formula (A-3),

or a pharmaceutically acceptable salt thereof, wherein:

Z^1 is absent, -O-, or -NH-;

each X^1, X^2, X^3, X^4, X^5, X^6, X^7, and X^8 is independently O, S, or NR^2;

each Y^1, Y^2, Y^3, Y^4, Y^5, Y^6, Y^7, and Y^8 is independently CH or N;

W^1 is deuterium, hydrogen, optionally substituted C_1-C_6 alkyl, (azaneylidene)methanediamine, (azaneylidene)-N,N,N',N'-tetramethylmethanediamine, -C(O)-$NR^{1A}R^{1B}$, -NR^{1A}-C(O)-$NR^{1A}R^{1B}$, -Z_B-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_{p3}$-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_{p3}$-O-$P(O)(OR^{1A})_2$, wherein

Z_B is -N- or -O-;

p_3 is an integer from 1-10;

W^2 is an optionally substituted C_1-C_6 alkyl or -C(O)-$NR^{1A}R^{1B}$;

each R^1 is independently hydrogen, deuterium, halogen, amino, cyano, optionally C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, or -NC(O)R^{1A}; or

two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3 to 6-membered carbocyclic ring or optionally substituted 3 to 6-membered heterocyclic ring;

each R^2 is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino;

each R^{1A} is independently hydrogen, deuterium, or optionally substituted C_1-C_{20} alkyl;

each R^{1B} is independently hydrogen, deuterium, optionally substituted 5-membered heteroaryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_2-C_{10} heteroalkyl, or $(AA)_{p2}$, wherein

each AA is an amino acid;

p_2 is an integer from 1-10;

j_1 is 0 or 1;

m_1 is an integer from 0-3;

p_1 is 2 or 3, provided that when Z^1 is -O- or -NH-, then p_1 is 2; and when Z^1 is absent then p_1 is 3; and

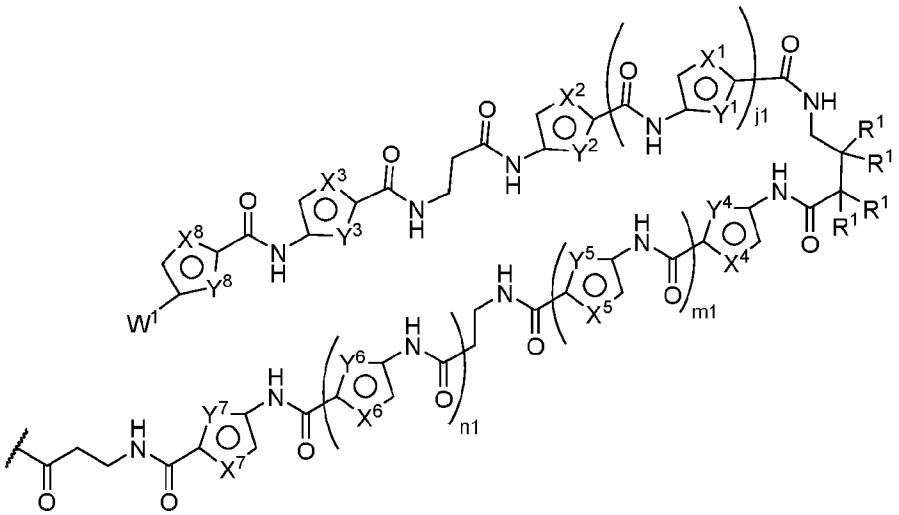
wherein one of W^1 or W^2 is attached to the oligomeric backbone.

[0032] In some embodiments, Z^1 is absent. In some embodiments, Z^1 is -O- or -NH-.

[0033] In some embodiments, W^2 is $-C(O)NR^{1A}R^{1B}$, wherein W^2 is attached to the oligomeric backbone. In some embodiments, R^{1A} is hydrogen and R^{1B} is AA, wherein AA is beta alanine. In some embodiments, W^2 is -C(O)NH-(beta-alanine)-. In some embodiments, W^2 is $-C(O)NR^{1A}R^{1B}$, wherein R^{1A} is hydrogen and R^{1B} is alkyl optionally substituted with an oxo (=O). In some embodiments, W^2 is $-C(O)NH(CH_2)_2C(O)-$. In some embodiments, W^2 is -C(O)NH-.

[0034] In some embodiments, the DNA-binding moiety is connected to the oligomeric backbone through W^2. In some embodiments, the oligomeric backbone is a linker moiety. In some embodiments, the DNA-binding moiety is not connected to the oligomeric backbone through W^2. In some embodiments, W^2 is $-C(O)NH(CH_2)_2C(O)-**$, wherein the linker moiety is attached at **. In some embodiments, W^2 is $-C(O)O(CH_2)_2C(O)-**$, wherein the linker moiety is attached at **. In some embodiments, W^2 is -C(O)-NH-**, wherein the linker moiety is attached at **. In some embodiments, W^2 is -C(O)OH-**, wherein the linker moiety is attached at **. In some embodiments, W^2 is -C(O)-**, wherein the linker moiety is attached at **.

[0035] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-4), or a pharmaceutically acceptable salt thereof:



Formula (A-4).

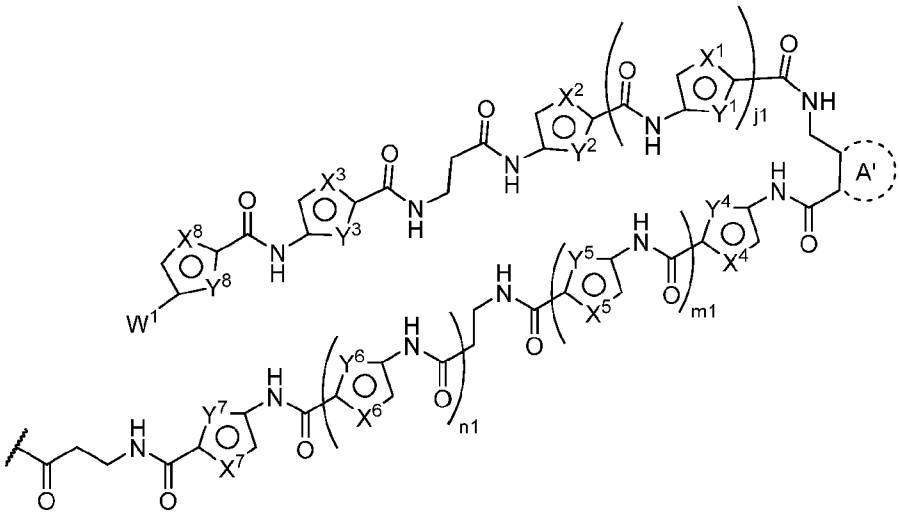
[0036] In some embodiments, each R^1 is independently halogen, amino, cyano, optionally C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl. In some embodiments, R^1 is C_1-C_{20} alkyl or C_1-C_{20} heteroalkyl. In some embodiments, each R^1 is independently $-NHC(O)R^{1A}$, wherein R^{1A} is alkyl, aryl, or heteroaryl. In some embodiments, each R^1 is independently $-NH_2$, $-NHCH_3$, or $-NHC(O)CH(CH_3)_3$. In some embodiments, each R^1 is hydrogen.

[0037] In some embodiments, two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3-6 membered carbocyclic ring or 3-6 membered heterocyclic ring. In some embodiments, two R^1 on the same carbon atom combine together to

form an optionally substituted 3-6 membered carbocyclic ring or 3-6 membered heterocyclic ring. In some embodiments, the two R^1 on the same carbon atom combine together to form an optionally substituted 3-6 membered carbocyclic ring. In some embodiments, the carbocyclic ring is a cyclopropyl, cyclobutyl, cyclopentyl, or cyclohexyl ring. In some embodiments, two R^1 on the same carbon atom combine together to form an optionally substituted 3-6 membered heterocyclic ring, optionally containing 1-2 heteroatoms selected from N, O, or S. In some embodiments, the heterocyclic ring is an oxetane, tetrahydrofuran, or tetrahydro-2H-pyran.

[0038] In some embodiments, two R^1 on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3-6 membered carbocyclic ring or 3-6 membered heterocyclic ring. In some embodiments, two R^1 on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3-6 membered carbocyclic ring. In some embodiments, two R^1 on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3-6 membered heterocyclic ring. In some embodiments, the cyclization occurs between the α and the β carbon atoms or between the β and the δ carbon atoms.

[0039] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-5):

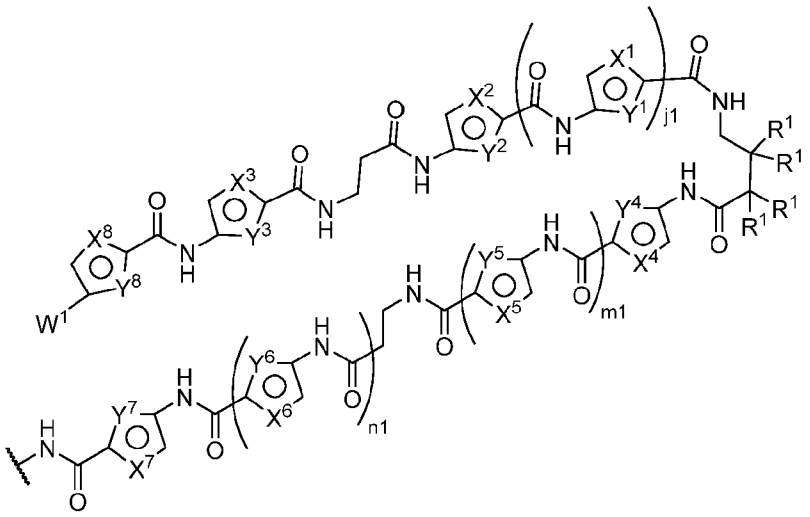


Formula (A-5),

or a pharmaceutically acceptable salt thereof, wherein:

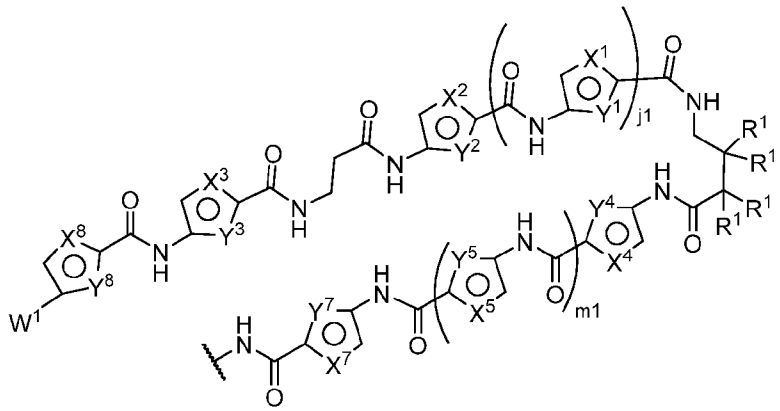
Ring A' is an optionally substituted C_3-C_6 carbocyclic ring or optionally substituted 3 to 6-membered heterocyclic ring.

[0040] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-6), or a pharmaceutically acceptable salt thereof:



Formula (A-6).

[0041] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-7), or a pharmaceutically acceptable salt thereof:



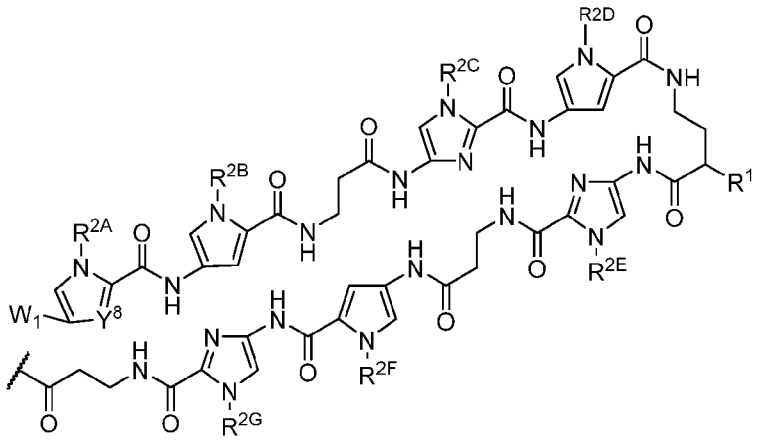
Formula (A-7).

[0042] In some embodiments, Y^2, Y^4, and Y^7 are each independently N; and Y^1 and Y^3 are each independently CH.

[0043] In some embodiments, Y^6 is CH. In some embodiments, Y^6 is N.

[0044] In some embodiments, X^1, X^2, X^3, X^4, X^5, X^6, and X^7 are each independently -NR^2.

[0045] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-8):



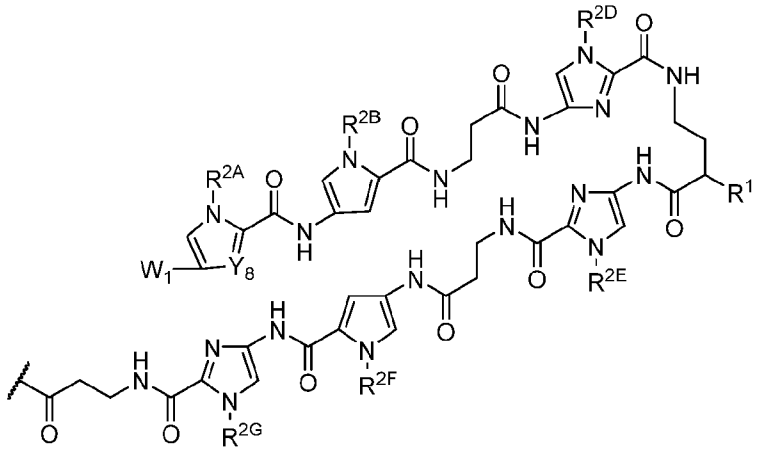
Formula (A-8),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

[0046] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-9):



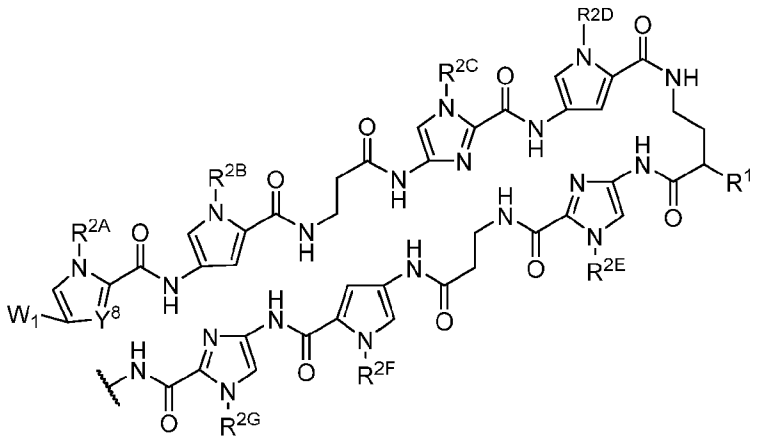
Formula (A-9),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

R^{2A}, R^{2B}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

[0047] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-10):



Formula (A-10),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

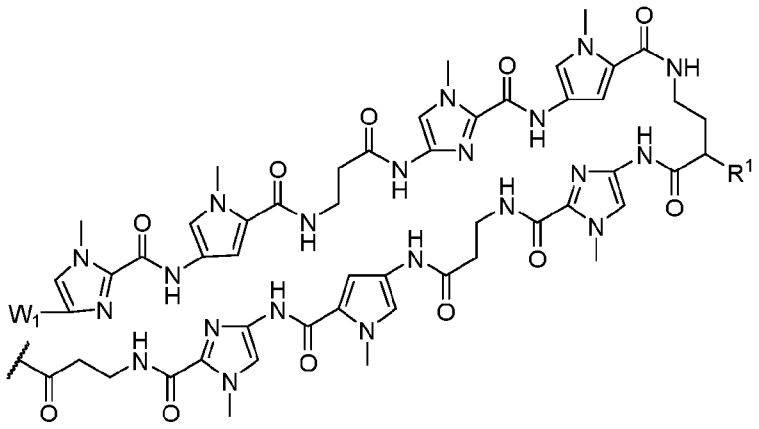
R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

[0048] In some embodiments, Y^8 is N. In some embodiments, Y^8 is CH.

[0049] In some embodiments, R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} heteroalkyl.

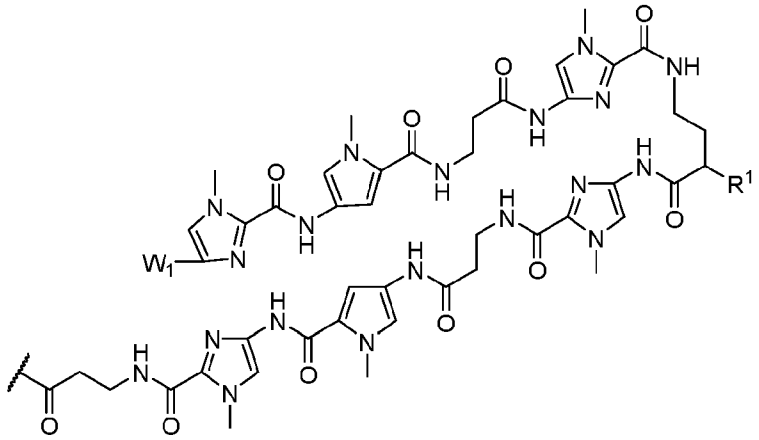
[0050] In some embodiments, R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently an optionally substituted C_1-C_{20} alkyl. In some embodiments, R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently a straight chain or branched C_1-C_{20} alkyl. In some embodiments, R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently an optionally substituted methyl, ethyl, isopropyl, n-butyl, isobutyl, sec-butyl, or tert-butyl. In some embodiments, R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently methyl, ethyl, or tert-butyl.

[0051] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-11), or a salt thereof:



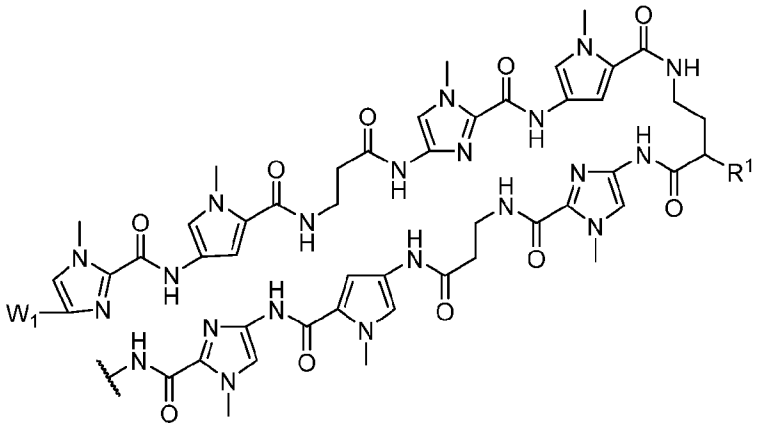
Formula (A-11).

[0052] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-12), or a salt thereof:



Formula (A-12).

[0053] In some embodiments, the DNA-binding moiety comprises the structure of Formula (A-13), or a salt thereof:

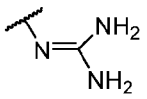
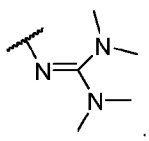


Formula (A-13).

[0054] In some embodiments, W^1 is $-C(O)-NR^{1A}R^{1B}$ or $-NR^{1A}-C(O)-NR^{1A}R^{1B}$.

[0055] In some embodiments, W^1 is hydrogen.

[0056] In some embodiments, W^1 is $-Z_B-P(O)(OR^{1A})_2$, $-Z_B-(CH_2)_{p3}-P(O)(OR^{1A})_2$, $-Z_B-(CH_2)_{p3}-O-P(O)_2(OR^{1A})_2$, wherein Z_B is O or N, and p_3 is an integer from 1-10.

[0057] In some embodiments, W^1 is (azaneylidene)methanediamine or (azaneylidene)-N,N,N',N'-tetramethylmethanediamine. In some embodiments, W^1 is . In some embodiments, W^1 is .

[0058] In some embodiments, the DNA-binding moiety is connected to the oligomeric backbone through W^1. In some embodiments, the oligomeric backbone is a linker moiety. In some embodiments, the DNA-binding moiety is not connected to the oligomeric backbone through W^1.

[0059] In some embodiments, m_1 is 0. In some embodiments, m_1 is 1. In some embodiments, m_1 is 2. In some embodiments, m_1 is 3.

[0060] In some embodiments, p_1 is 2. In some embodiments, p_1 is 3.

[0061] In some embodiments, m_1 is 0 or 1 and p_1 is 2.

[0062] In some embodiments, n_1 is 0. In some embodiments, n_1 is 1. In some embodiments, n_1 is 2. In some embodiments, n_1 is 3.

[0063] In some embodiments, j_1 is 0. In some embodiments, j_1 is 1.

[0064] The binding affinity between the polyamide and the target gene can be adjusted based on the composition of the polyamide. In some embodiments, the polyamide is capable of binding the DNA with an affinity of less than about 600 nM, about 500 nM, about 400 nM, about 300 nM, about 250 nM, about 200 nM, about 150 nM, about 100 nM, or about 50nM. In some embodiments, the polyamide is capable of binding the DNA with an affinity of less than about 300 nM. In some embodiments, the polyamide is capable of binding the DNA with an affinity of less than about 200 nM. In some embodiments, the polyamide is capable of binding the DNA with an affinity of greater than about 200 nM, about 150 nM, about 100 nM, about 50 nM, about 10 nM, or about 1 nM. In some embodiments, the polyamide is capable of binding the DNA with an affinity in the range of about 1-600 nM, 10-500 nM, 20-500 nM, 50-400 nM, or 100-300 nM.

[0065] In some embodiments, the first terminus is capable of binding the DNA with an affinity of less than 500 nM.

[0066] The binding affinity between the polyamide and the target DNA can be determined using a quantitative footprint titration experiment. The experiment involves measuring the dissociation constant K_d of the polyamide for the target sequence at either 24 ºC or 37 °C, and using either standard polyamide assay solution conditions or approximate intracellular solution conditions.

[0067] The binding affinity between the regulatory protein and the ligand on the second terminus can be determined using an assay suitable for the specific protein. The experiment involves measuring the dissociation constant K_d of the ligand for the protein and using either standard protein assay solution conditions or approximate intracellular solution conditions.

[0068] The first terminus DNA binding moiety in the molecules described herein has a high binding affinity to a sequence having multiple repeats of CAG and binds to the target nucleotide repeats preferentially over other nucleotide repeats or other nucleotide sequences. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAG than to a sequence having repeats of CGG. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAG than to a sequence having repeats of CCG. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAG than to a sequence

having repeats of CCTG. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAG than to a sequence having repeats of TGGAA. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAG than to a sequence having repeats of GGGGCC. In some embodiments, the first terminus has a higher binding affinity to a sequence having multiple repeats of CAGCTG than to a sequence having repeats of GAA.

[0069] Due to the preferential binding between the first terminus and the target nucleotide repeat, the transcription modulation molecules described herein become localized around regions having multiple repeats of CAG. In some embodiments, the local concentration of the first terminus of the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of CGG. In some embodiments, the local concentration of the first terminus of the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of CCG. In some embodiments, the local concentration of the first terminus of the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of CCTG. In some embodiments, the local concentration of the first terminus or the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of TGGAA. In some embodiments, the local concentration of the first terminus of the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of GGGGCC. In some embodiments, the local concentration of the first terminus of the molecules described herein is higher near a sequence having multiple repeats of CAG than near a sequence having repeats of GAA.

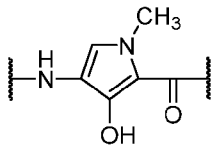
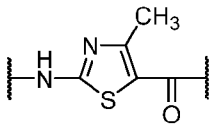
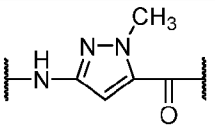
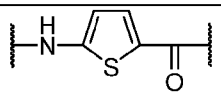
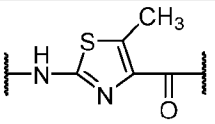
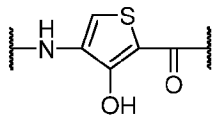
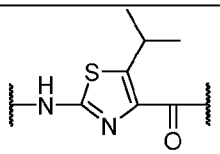
[0070] The first terminus DNA binding moiety in the molecules described herein is localized to a sequence having multiple repeats of CAG and binds to the target nucleotide repeats preferentially over other nucleotide repeats. In some embodiments, the sequence has at least 2, 3, 4, 5, 8, 10, 12, 15, 20, 25, 30, 40, 50, 100, 200, 300, 400, or 500 repeats of CAG. In some embodiments, the sequence comprises at least 1000 nucleotide repeats of CAG. In certain embodiments, the sequence comprises at least 500 nucleotide repeats of CAG. In certain embodiments, the sequence comprises at least 200 nucleotide repeats of CAG. In some embodiments, the sequence comprises at least 100 nucleotide repeats of CAG. In some embodiments, the sequence comprises at least 50 nucleotide repeats of CAG. In some embodiments, the sequence comprises at least 20 nucleotide repeats of CAG.

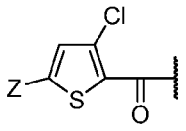
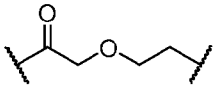
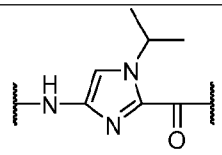
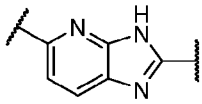
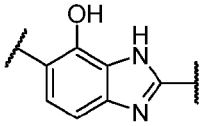
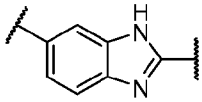
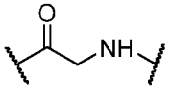
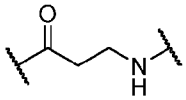
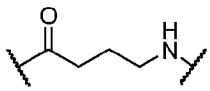
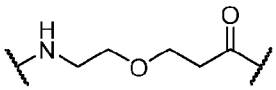
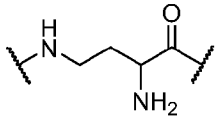
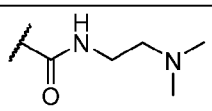
[0071] The polyamide composed of a pre-selected combination of subunits can selectively bind to the DNA in the minor groove. In their hairpin structure, antiparallel side-by-side pairings of two aromatic amino acids bind to DNA sequences, with a polyamide ring packed specifically against each DNA base. N-Methylpyrrole (Py) favors T, A, and C bases, excluding G; N-methylimidazole (Im) is a G-reader; and 3-hydroxyl-N-methylpyrrol (Hp) is specific for thymine base. The nucleotide base pairs can be recognized using different pairings of the amino acid subunits using the pairing principle shown in Table 1A and 1B below. For example, an Im/Py pairing reads G·C by symmetry, a Py/Im pairing reads C·G, an Hp/Py pairing

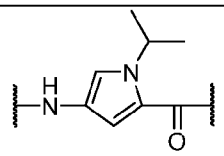
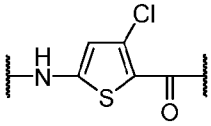
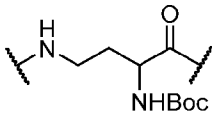
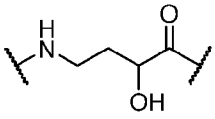
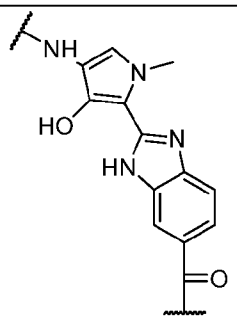
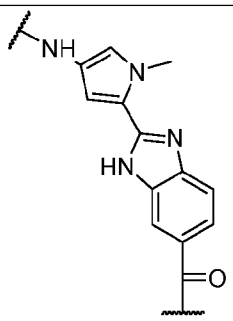
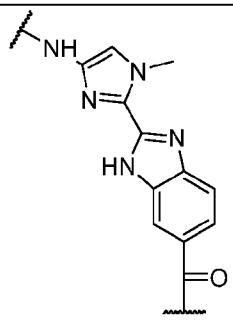
can distinguish T·A from A·T, G·C, and C·G, and a Py/Py pairing nonspecifically discriminates both A·T and T·A from G·C and C·G.

[0072] In some embodiments, the first terminus comprises Im corresponding to the nucleotide G; Im or Nt corresponding to the nucleotide pair G; Py corresponding to the nucleotide C, wherein Im is N-alkyl imidazole, Py is N-alkyl pyrrole, Hp is 3 -hydroxy N-methyl pyrrole, and β-alanine. In some embodiments, the first terminus comprises Im/Py to correspond to the nucleotide pair G/C, Py/Im to correspond to the nucleotide pair C/G, and wherein Im is N-alkyl imidazole (e.g., N-methyl imidazole), Py is N-alkyl pyrrole (e.g., N-methyl pyrrole), and Hp is 3-hydroxy N-methyl pyrrole.

Table 1A. Base pairing for single amino acid subunit (Favored (+), disfavored (-)).

Subunit	G	C	A	T
Py	-	+	+	+
Im	+	-	-	
 (Hp)	-	-	-	+
 (Th)	-	-	+	+
 (Pz)	-	-	+	+
 (Tp)	-	-	+	+
 (Nt)	+	-	-	-
 (Ht)	-	-	-	+
 (iPTA)	+	-	-	-

(CTh) 	-	-	-	+
(PEG) 	-	+	+	+
(iIm) 	+	-	-	-
(Ip) 	+	-	-	-
(Hz) 	-	-	-	+
(Bi) 	-	-	-	+
(Gly) 	-	-	-	-
(Beta or β) 	-	-	+	+
(gAB) 	-	-	+ (as a part of the turn)	+ (as a part of the turn)
(Alx) 	-	+	-	-
(Da) 	-	-	+	+
(Dp) 	-	-	+	+

 (iPP)	-	-	+	+
 (CTh)	+	+	-	-
 (Dab)	-	-	+	+
 (gAH)	-	-	+	+
 (HpBi)	WW* (bind to two nucleotides with same selectivity as Hp-Py)			
 (PhBi)	WW* (bind to two nucleotides with same selectivity as Py-Py)			
 (ImBi)	GW* (bind to two nucleotides with same selectivity as Im-Py)			

*The subunit HpBi, ImBi, and PyBi function as a conjugate of two monomer subunits and bind to two nucleotides. The binding property of HpBi, ImBi, and PyBi corresponds to Hp-Py, Im-Py, and Py-Py respectively.

Table 1B. Representative base pairing for hairpin polyamide.

	G·C	C·G	T·A	A·T
Im/β	+	-	-	-
β/Im	-	+	-	-
Py/ β	-	-	+	+
β/Py	-	-	+	+
β/ β	-	-	+	+
Py/Py	-	-	+	+
Im/Im	-	-	-	-
Im/Py	+	-	-	-
Py/Im	-	+	-	-
Th/Py	-	-	+	-
Py/Th	-	-	-	+
Th/Im	+	-	-	-
Im/Th	-	+	-	-
β/Th	-	-	+	-
Th/β	-	-	-	+
Hp/Py,	-	-	+	-
Py/Hp,	-	-	-	+
Hp/Im	+	-	-	-
Im/Hp	-	+	-	-
Tn/Py	-	-	+	+
Py/Tn,	-	-	+	+
Ht/Py,	-	-	+	+
Py/Ht,	-	-	+	+
Bi/Py,	-	-	+	+
Py/Bi,	-	-	+	+
β/Bi	-	-	+	+
Bi/β	-	-	+	+
Bi/Im,	-	+	-	-
Im/Bi,	+	-	-	-
Tp/Py,	-	-	+	+
Py/Tp,	-	-	+	+
β/Tp	-	-	+	+
Tp/β	-	-	+	+
Tp/Im,	-	+	-	-
Im/Tp	+	-	-	-
Tp/Tp	-	-	+	+
Tp/Tn	-	-	+	+
Tn/Tp	-	-	+	+
Hz/Py,	-	-	+	-
Py/Hz,	-	-	-	+
Ip/Py	+	-	-	-
Py/Ip,	-	+	-	-
Bi/Hz,	-	-	+	+

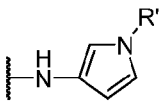
Hz/Bi,	-	-	+	+
Bi/Bi	-	+	+	+
Th/Py,	-	-	+	+
Py/Th	-	-	+	+
Im/gAB	+	-	-	-
gAB/Im	-	+	-	-
Py/ gAB	+	-	-	-
gAB/Py	-	+	-	-
gAB/ β	-	-	+	+
β/gAB	-	-	+	+
Im/Dp	+	-	-	-
Dp/Im	-	+	-	-
Py/ Dp	-	-	+	+
Dp/Py	-	-	+	+
Dp/β	-	-	+	+
Each of HpBi, ImBi, and PyBi can bind to two nucleotides and have binding properties corresponding to Hp-Py, Im-Py, and Py-Py respectively. HpBi, ImBi, and PyBi can be paired with two monomer subunits or with themselves in a hairpin structure to bind to two nucleotide pairs.				

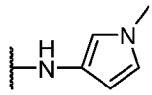
[0073] The monomer subunits of the polyamide can be strung together based on the pairing principles shown in Table 1A and Table 1B. The monomer subunits of the polyamide can be strung together based on the pairing principles shown in Table 1C.

[0074] Table 1C shows an example of the monomer subunits that can bind to the specific nucleotide. The first terminus can include a polyamide described as having several monomer subunits strung together, with a monomer subunit selected from each row. For example, the polyamide can include Py-Py-Im that binds to CAG, where Py is selected from the C column, Py is selected from the A column, and Im selected from the first G column. The polyamide can be any combinations of the subunits of CAGCAG, with a subunit selected from each column in Table 1C, wherein the subunits are strung together following the CAG binding order.

[0075] In addition, the polyamide can also include a partial or multiple sets of the five subunits, such as 1.5, 2, 2.5, 3, 3.5, or 4 sets of the three subunits. The polyamide can include 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, and 16 monomer subunits.

[0076] The polyamide can include monomer subunits that bind to 2, 3, 4, or 5 nucleotides of CAG. For example, the polyamide can bind to CA, CAG, AGC, CAGC, CAGCA, CAGCAG. The polyamide can include monomer subunits that bind to 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides of CAG repeat.

[0077] The monomer subunit, when positioned as a terminal unit, does not have an amine, carbonyl, or a carboxylic acid group at the terminal. The carboxylic acid group in the terminal is replaced by a hydrogen. For example, Py, when used as a terminal unit, is understood to have the structure of  (e.g.,


); and Im, when positioned as a terminal unit, is understood to have the structure of

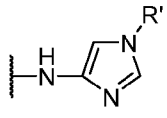
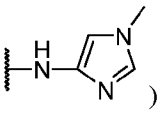

(e.g., 
).

Table 1C. Examples of monomer subunits in a linear polyamide that binds to CAG.

Nucleotide	C	A	G
Subunit that selectively binds to nucleotide	Py	Py	Im or ImT
	iIm or iImT	Th	iIm or iImT
	PEG	Pz	PEG
	CTh	Tp	CTh
	Alx	PEG	Nt
		β	iPTA
		iPP	Ip
		Da	CTh
		Dp	
		Dab	
		gAH	

[0078] Recognition of a nucleotide repeat or DNA sequence by two antiparallel polyamide strands depends on a code of side-by-side aromatic amino acid pairs in the minor groove, usually oriented N to C with respect to the 5' to 3' direction of the DNA helix. Enhanced affinity and specificity of polyamide nucleotide binding is accomplished by covalently linking the antiparallel strands. The "hairpin motif" connects the N and C termini of the two strands with a gamma-aminobutyric acid unit (gamma-turn)) to form a folded linear chain. The "H-pin motif" connects the antiparallel strands across a central or near central ring/ring pairs by a short, flexible bridge.

Second Terminus – Regulatory Protein Binding Moiety

[0079] In some embodiments, the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat.

[0080] In some embodiments, the second terminus comprises a bromodomain binding moiety.

[0081] In some embodiments, the second terminus comprises a moiety capable of binding to a bromodomain and extra terminal domain (BET) family member.

[0082] In some embodiments, the BET family member is BRD2, BRD3, BRD4, or BRDT. In some embodiments, the BET family member is BRD2. In some embodiments, the BET family member is BRD3. In some embodiments, the BET family member is BRD4. In some embodiments, the BET family member is BRD3. In some embodiments, the BET family member is BRDT.

[0083] In some embodiments, the bromodomain is CBP/p300, PCAF (P300/CBP-Associated Factor), CECR2 (cat eye syndrome chromosome region candidate 2), BRPF (bromodomain and PHD finger-containing protein), ATAD2/ATAD2B (chromatin remodeling proteins), TRIM24 (Tripartite motif-containing 24), BAZ2 (Bromodomain Adjacent to Zinc finger), or TAF1 (TBP associated factors).

[0084] In some embodiments, the bromodomain is CBP/p300.

[0085] In some embodiments, the bromodomain is PCAF (P300/CBP-Associated Factor).

[0086] In some embodiments, the bromodomain is CECR2 (cat eye syndrome chromosome region candidate 2).

[0087] In some embodiments, the bromodomain is BRPF (bromodomain and PHD finger-containing protein).

[0088] In some embodiments, the bromodomain is a ATAD2 or ATAD2B chromatin remodeling protein.

[0089] In some embodiments, the bromodomain is BAZ2 (Bromodomain Adjacent Zinc Finger.

[0090] In some embodiments, the bromodomain is TAF1 (TBP associated factor).

[0091] In some embodiments, the bromodomain is TRIM24 (tripartite motif-containing 24).

[0092] In some embodiments, the regulatory molecule modulates the rearrangement of histones.

[0093] In some embodiments, the regulatory molecule modulates the glycosylation, phosphorylation, alkylation, or acylation of histones.

[0094] In some embodiments, the regulatory molecule is a transcription factor.

[0095] In some embodiments, the regulatory molecule is an RNA polymerase.

[0096] In some embodiments, the regulatory molecule is a moiety that regulates the activity of RNA polymerase.

[0097] In some embodiments, the recruiting moiety binds to the regulatory molecule but does not inhibit the activity of the regulatory molecule. In some embodiments, the recruiting moiety binds to the regulatory molecule and inhibits the activity of the regulatory molecule. In some embodiments, the recruiting moiety binds to the regulatory molecule and increases the activity of the regulatory molecule.

[0098] In some embodiments, the recruiting moiety binds to the active site of the regulatory molecule. In certain embodiments, the recruiting moiety binds to a regulatory site of the regulatory molecule.

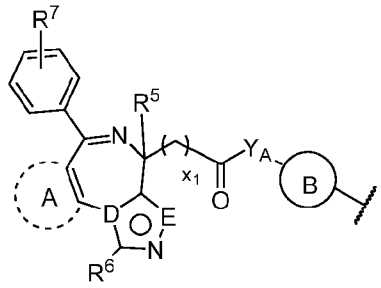
[0099] The binding affinity between the regulatory protein and the second terminus can be adjusted based on the composition of the molecule or type of protein. In some embodiments, the second terminus binds the regulatory molecule with an affinity of less than about 600 nM, about 500 nM, about 400 nM, about 300 nM, about 250 nM, about 200 nM, about 150 nM, about 100 nM, or about 50nM. In some embodiments, the second terminus binds the regulatory molecule with an affinity of less than about 300 nM. In some embodiments, the second terminus binds the regulatory molecule with an affinity of less than about 200 nM.

[00100] In some embodiments, the second terminus comprises a diazine or diazepine ring, wherein the diazine or diazepine ring is fused with a C_6-C_{10} aryl or a 5 to 10-membered heteroaryl ring comprising one or

more heteroatoms selected from S, N and O. In some embodiments, the second terminus comprises an optionally substituted bicyclic or tricyclic structure.

[00101] In some embodiments, the second terminus has a triazolodiazepine structure. In some embodiments, the second terminus has a thiazolodiazepine structure.

[00102] In some embodiments, the second terminus comprises the structure of Formula (2-A):



Formula (2-A),

or a pharmaceutically acceptable salt thereof, wherein:

Ring A is an optionally substituted aryl or optionally substituted 5 to 6-membered heteroaryl;

Ring B is absent or an optionally substituted 6-membered monocyclic aryl or heteroaryl;

D is C or N;

E is O or N;

Y_A is -NH- or -O-;

R^5 is hydrogen, deuterium, or C_1-C_6 alkyl;

R^6 is selected from hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

R^7 is selected from hydrogen, deuterium, halogen, -NO_2, -CN, optionally substituted aryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

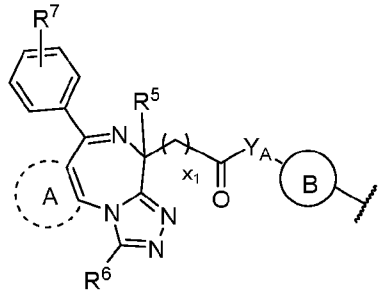
or R^7 is -$NR^{7A}R^{7B}$, wherein

R^{7A} and R^{7B} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl; and

x_1 is an integer from 1-6.

[00103] In some embodiments, D is N and E is N. In some embodiments, D is C and E is O.

[00104] In some embodiments, the second terminus comprises the structure of Formula (2-B):



Formula (2-B),

or a pharmaceutically acceptable salt thereof, wherein:

Ring A is an optionally substituted aryl or optionally substituted 5 to 6-membered heteroaryl;

Ring B is absent or an optionally substituted 6-membered monocyclic aryl or heteroaryl;

Y_A is -NH- or -O-;

R^5 is hydrogen, deuterium, or C_1-C_6 alkyl;

R^6 is selected from hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

R^7 is selected from hydrogen, deuterium, halogen, -NO_2, -CN, optionally substituted aryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

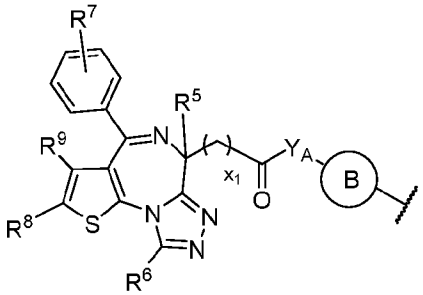
or R^7 is -$NR^{7A}R^{7B}$, wherein

R^{7A} and R^{7B} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl; and

x_1 is an integer from 1-6.

[00105] In some embodiments, Ring A is an optionally substituted aryl ring. In some embodiments, Ring A is an optionally substituted phenyl. In some embodiments, Ring A is an optionally substituted 5 membered heteroaryl. In some embodiments, Ring A is an optionally substituted oxazolyl. In some embodiments, Ring A is an optionally substituted furanyl. In some embodiments, Ring A is an optionally substituted thiophenyl.

[00106] In some embodiments, the second terminus comprising the structure of Formula (2-C):



Formula (2-C),

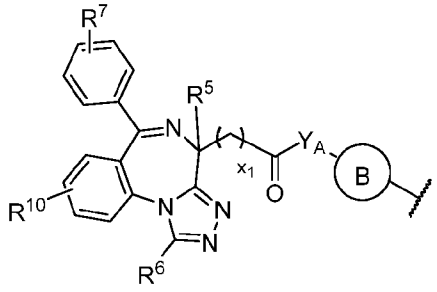
or a pharmaceutically acceptable salt thereof, wherein:

R^8 and R^9 are each independently selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

[00107] In some embodiments, R^8 and R^9 are each independently selected from optionally substituted C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 hydroxyalkyl. In some embodiments, R^8 and R^9 are each independently selected from optionally substituted C_1-C_6 alkyl. In some embodiments, R^8 and R^9 are each independently methyl, ethyl, or propyl. In some embodiments, R^8 and R^9 are each independently methyl. In some

embodiments, R^8 and R^9 are each independently ethyl. In some embodiments, R^8 and R^9 are each independently propyl.

[00108] In some embodiments, the second terminus comprising the structure of Formula (2-D):



Formula (2-D),

or a pharmaceutically acceptable salt thereof, wherein:

R^{10} is selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

[00109] In some embodiments, R^5 is C_1-C_6 alkyl. In some embodiments, R^5 is methyl or ethyl. In some embodiments, R^5 is methyl. In some embodiments, R^5 is ethyl. In some embodiments, R^5 is hydrogen.

[00110] In some embodiments, R^7 is selected from hydrogen, halogen, optionally substituted C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 hydroxyalkyl. In some embodiments, R^7 is halogen. In some embodiments, R^7 is Br, Cl, or F. In some embodiments, R^7 is Cl. In some embodiments, R^7 is F. In some embodiments, R^7 is Br.

[00111] In some embodiments, R^7 is $-NR^{7A}R^{7B}$, wherein R^{7A} and R^{7B} are each independently hydrogen or optionally substituted C_1-C_6 alkyl.

[00112] In some embodiments, R^{10} is selected from optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl. In some embodiments, R^{10} is selected from optionally substituted C_1-C_6 alkyl. In some embodiments, R^{10} is methyl, ethyl, or propyl. In some embodiments, R^{10} is methyl. In some embodiments, R^{10} is optionally substituted $C_{1\text{-}6}$ hydroxyalkyl. In some embodiments, R^{10} is -OMe.

[00113] In some embodiments, R^6 is selected from optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl. In some embodiments, R^6 is an optionally substituted C_1-C_6 alkyl. In some embodiments, R^6 is methyl, ethyl, or propyl. In some embodiments, R^6 is methyl. In some embodiments, R^6 is ethyl. In some embodiments, R^6 is propyl. In some embodiments, R^6 is hydrogen.

[00114] In some embodiments, Y_A is -NH-. In some embodiments, Y_A is -O-.

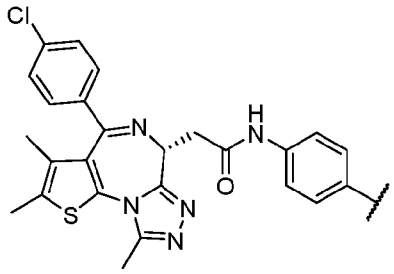
[00115] In some embodiments, Y_A is NH and x_1 is 1.

[00116] In some embodiments, x_1 is an integer from 1-5, 1-4, 1-3, or 1-2. In some embodiments, x_1 is 1. In some embodiments, x_1 is 2.

[00117] In some embodiments, Ring B is an optionally substituted 6-membered monocyclic aryl or heteroaryl, each of which is optionally substituted with alkyl, amino, halogen, hydroxy, hydroxyalkyl, or

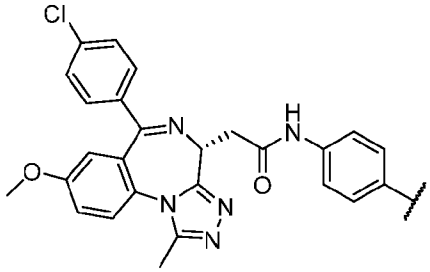
PEG. In some embodiments Ring B is phenyl. In some embodiments, Ring B is 6-membered monocyclic heteroaryl. In some embodiments, Ring B is pyridine or pyrimidine. In some embodiments, ring B is absent.

[00118] In some embodiments, the second terminus comprises the structure of Formula (2-E), or a pharmaceutically acceptable salt thereof:



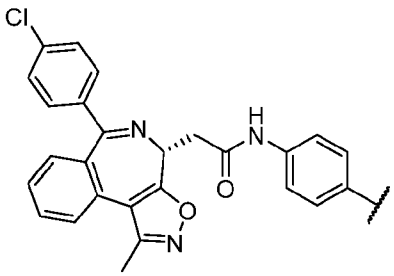
Formula (2-E).

[00119] In some embodiments, the second terminus comprises the structure of Formula (2-F), or a pharmaceutically acceptable salt thereof:



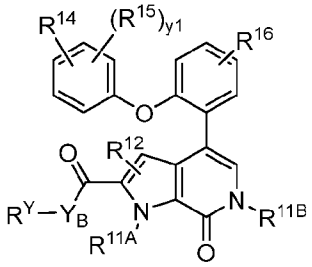
Formula (2-F).

[00120] In some embodiments, the second terminus comprises the structure of Formula (2-G), or a pharmaceutically acceptable salt thereof:



Formula (2-G).

[00121] In some embodiments, the second terminus comprises the structure of Formula (3-A):



Formula (3-A),

or a pharmaceutically acceptable salt thereof, wherein,

Y_B is -CH_2NH-, -CH_2O-, -NH-, or -O-;

R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, halogen, -OH, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{14} and R^{15} are each independently hydrogen, deuterium, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is -NR^AR^B;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, C_1-C_6 hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$;

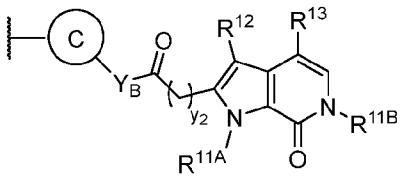
R^Y is hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted 5-6-membered monocyclic aryl or heteroaryl;

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is 1-3; and

wherein the attachment to the linker is at either R^{14} or at R^Y.

[00122] In some embodiments, the second terminus comprises the structure of Formula (3-B):



Formula (3-B),

or a pharmaceutically acceptable salt thereof, wherein,

Ring C is absent, optionally substituted 5 to 6-membered monocyclic aryl or heteroaryl or 4 to 8-membered heterocycle;

Y_B is -NH-, -CH_2NH-, -CH_2O-, or -O-;

R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, C(O)R^A, or C(O)NR^AR^B; wherein

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl;

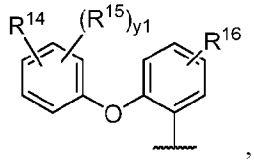
R^{13} is hydrogen, substituted aryl, substituted heteroaryl, or substituted oxydibenzene; and

y_2 is an integer from 0-2.

[00123] In some embodiments, y_2 is 0. In some embodiments, y_2 is 1. In some embodiments, y_2 is 2.

[00124] In some embodiments, R^{13} is substituted aryl or substituted heteroaryl. In some embodiments, R^{13} is hydrogen.

[00125] In some embodiments, R^{13} is substituted oxydibenzene.



[00126] In some embodiments, R^{13} is , wherein

R^{14} and R^{15} are each independently hydrogen, deuterium, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

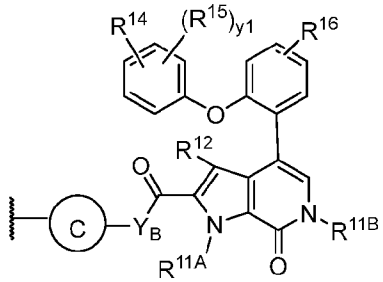
or R^{14} is -NR^AR^B;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, C_1-C_6 hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$;

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is 1-3.

[00127] In some embodiments, the second terminus comprises the structure of Formula (3-C):



Formula (3-C),

or a pharmaceutically acceptable salt thereof, wherein:

Ring C is absent, optionally substituted 5 to 6-membered monocyclic aryl or heteroaryl or 4 to 8-membered heterocycle;

Y_B is -NH-, -CH_2NH-, -CH_2O-, or -O-;

R^{14} and R^{15} are each independently hydrogen, deuterium, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is -NR^AR^B, wherein

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, C_1-C_6 hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$; and

y_1 is an integer from 1-3.

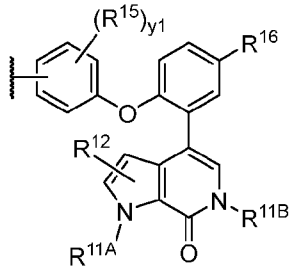
[00128] In some embodiments, Y_B is -NH-. In some embodiments, Y_B is -CH_2NH-. In some embodiments, Y_B is -CH_2O-. In some embodiments, Y_B is -O-.

[00129] In some embodiments, Ring C is an optionally substituted 5 or 6-membered monocyclic aryl or heteroaryl, each of which is optionally substituted with alkyl, amino, halogen, hydroxy, hydroxyalkyl, or PEG.

[00130] . In some embodiments, Ring C is phenyl. In some embodiments, Ring C is a 6-membered heteroaryl. In some embodiments, Ring C is pyridine, pyrazine, or triazine. In some embodiments, Ring C is pyridine. In some embodiments, Ring C is pyrazine. In some embodiments, Ring C is triazine. In some embodiments, Ring C is a 5-membered heteroaryl. In some embodiments, Ring C is a pyrazole. In some embodiments, Ring C is a triazole, pyrrole, imidazole, oxazole, oxadiazole, thiazole, or thiadiazole. In some embodiments, Ring C is a triazole. In some embodiments, Ring C is an imidazole or pyrrole. In some embodiments, an oxazole or oxadiazole. In some embodiments, Ring C is a thiazole or thiadiazole.

[00131] In some embodiments, Ring C is absent.

[00132] In some embodiments, the second terminus comprises the structure of Formula (3-D), or a pharmaceutically acceptable salt thereof:



Formula (3-D),

or a pharmaceutically acceptable salt or solvate thereof, wherein:

R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, $C(O)R^A$, or $C(O)NR^AR^B$;

each R^{15} is independently hydrogen, deuterium, halogen, -CN, $-NO_2$, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, optionally substituted C_1-C_6 hydroxyalkyl, $-SO_2R^A$-, or $-NHSO_2R^A$; wherein

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is an integer from 1-3.

[00133] In some embodiments, R^{11A} and R^{11B} are each independently optionally substituted C_1-C_6 alkyl. In some embodiments, R^{11A} and R^{11B} are each independently methyl, ethyl, propyl, or tert-butyl. In some embodiments, R^{11A} and R^{11B} are each independently methyl. In some embodiments, R^{11A} and R^{11B} are each independently hydrogen.

[00134] In some embodiments, R^{11A} is C_1-C_6 alkyl, optionally substituted with haloalkyl or phosphorous hydroxide. In some embodiments, R^{11A} is C_1-C_6 alkyl substituted with $-OP(O)(OH)_2$. In some embodiments, R^{11A} is unsubstituted C_1-C_6 alkyl. In some embodiments, R^{11A} is methyl, ethyl, or tert-butyl. In some embodiments, R^{11A} is methyl. In some embodiments, R^{11A} is hydrogen.

[00135] In some embodiments, R^{12} is optionally substituted C_1-C_6 alkyl. In some embodiments, R^{12} is hydrogen.

[00136] In some embodiments, R^{12} is $C(O)R^A$ or $C(O)NR^AR^B$. In some embodiments, R^{12} is $C(O)NR^AR^B$, wherein R^A and R^B are each independently hydrogen or optionally substituted C_1-C_6 alkyl.

[00137] In some embodiments, R^{14} and R^{15} are each independently hydrogen, -CN, or $-NO_2$. In some embodiments, R^{14} and R^{15} are each independently halogen or optionally substituted C_1-C_6 alkyl. In some embodiments, R^{14} and R^{15} are each independently Br, Cl, F, methyl, or ethyl. In some embodiments, R^{14} and R^{15} are each independently F or methyl.

[00138] In some embodiments, R^{16} is optionally substituted optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, or optionally substituted C_1-C_6 hydroxyalkyl, each of which is optionally substituted with amido, alkyl, alkynyl, azido, amino, halogen, haloalkyl, hydroxy, nitro, oxo (=O), phosphorous hydroxide, or PEG.

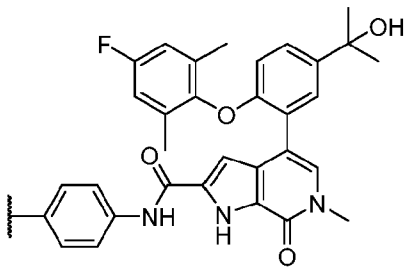
[00139] In some embodiments, R^{16} is optionally substituted optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, or optionally substituted C_1-C_6 hydroxyalkyl. In some embodiments, R^{16} is C_1-C_6 alkyl or C_1-C_6 heteroalkyl, each or which optionally substituted with -CN, $-NH_2$, $-N_3$, -OH, CF_3, or $-OP(O)(OH)_2$.

[00140] In some embodiments, R^{16} is $-SO_2R^A$, wherein R^A is C_1-C_6 alkyl. In some embodiments, R^{16} is $-SO_2Et$. In some embodiments, R^{16} is $-SO_2Me$.

[00141] In some embodiments, R^{16} is $-NHSO_2R^A$, wherein R^A is C_1-C_6 alkyl. In some embodiments, R^{16} is $-NHSO_2Et$. In some embodiments, R^{16} is $-NHSO_2Me$.

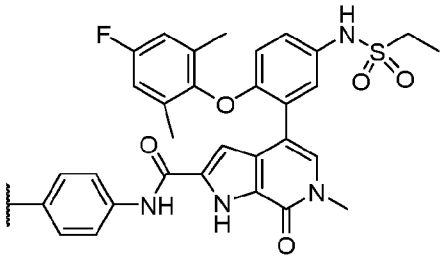
[00142] In some embodiments, y_1 is 1. In some embodiments, y_1 is 2. In some embodiments, y_1 is 3.

[00143] In some embodiments, the second terminus comprises the structure of Formula (3-E), or a pharmaceutically acceptable salt thereof:



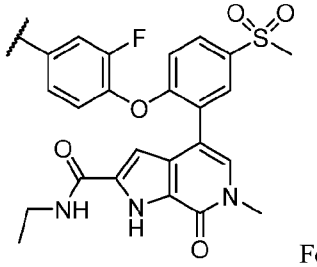
Formula (3-E).

[00144] In some embodiments, the second terminus comprising the structure of Formula (3-F), or a pharmaceutically acceptable salt thereof:

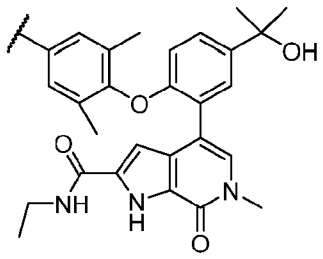


Formula (3-F).

[00145] In some embodiments, the second terminus comprising the structure of Formula (3-G) or Formula (3-H), or a pharmaceutically acceptable salt thereof:

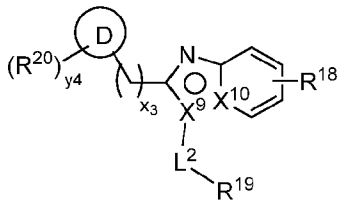


Formula (3-G) or



Formula (3-H).

[00146] In some embodiments, the second terminus comprises the structure of Formula (4-A):



Formula (4-A),

or a pharmaceutically acceptable salt thereof, wherein;

Ring D is absent, phenyl, or 5 to 6-membered heteroaryl;

X^9 and X^{10} are each independently C or N, wherein one of X^9 or X^{10} is N;

L^2 is absent, optionally substituted alkylene, -O-, or $-NR^D-$, wherein

R^D is hydrogen, deuterium, or optionally substituted C_1-C_3 alkyl;

R^{18} is an optionally substituted 5 to 6-membered heteroaryl;

R^{19} is an optionally substituted C_3-C_8 cycloalkyl or optionally substituted 4 to 7-membered heteroaryl;

each R^{20} is independently hydrogen, deuterium, halogen, -CN, $-NO_2$, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

x_3 is an integer from 1-3; and

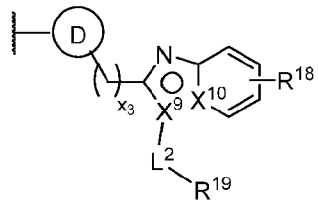
y_3 is an integer from 1-4; and

wherein the attachment to the linker is at either R^{19} or at one of R^{20}.

[00147] In some embodiments, attachment to the linker is at R^{19}.

[00148] In some embodiments, attachment to the linker is at one of R^{20}.

[00149] In some embodiments, the second terminus comprises the structure of Formula (4-B):



Formula (4-B),

or a pharmaceutically acceptable salt thereof, wherein;

Ring D is absent, optionally substituted phenyl, or optionally substituted 5 to 6-membered heteroaryl;

X^9 and X^{10} are each independently C or N, wherein one of X^9 or X^{10} is N;

L^2 is absent, optionally substituted alkylene, -O-, or -NR^D-, wherein

R^D is hydrogen, deuterium, or optionally substituted C_1-C_3 alkyl;

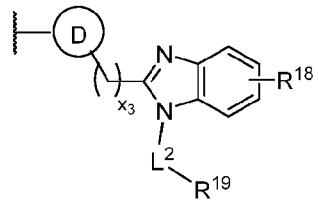
R^{18} is an optionally substituted 5 to 6-membered heteroaryl;

R^{19} is an optionally substituted C_3-C_8 cycloalkyl or optionally substituted 4 to 7-membered heteroaryl; and

x_3 is an integer from 1-3.

[00150] In some embodiments, X^9 is N; and X^{10} is C. In some embodiments, X^9 is C; and X^{10} is N.

[00151] In some embodiments, the second terminus comprises the structure of Formula (4-C), or a pharmaceutically acceptable salt thereof:

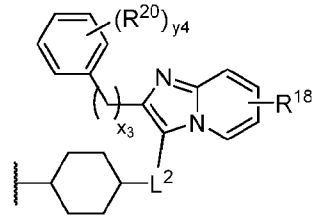


Formula (4-C).

[00152] In some embodiments, Ring D is an optionally substituted monocyclic 6-membered aryl or 5 to 6-membered heteroaryl. In some embodiments, Ring D is an optionally substituted monocyclic 6-membered aryl. In some embodiments, Ring D is an optionally substituted phenyl.

[00153] In some embodiments, R^{19} is an optionally substituted C_3-C_8 cycloalkyl. In some embodiments, R^{19} is optionally substituted 4 to 7-membered heteroaryl.

[00154] In some embodiments, the second terminus comprises the structure of Formula (4-D):



Formula (4-D),

or a pharmaceutically acceptable salt thereof, wherein;

L^2 is an optionally substituted alkylene, -O-, or -NR^D-, wherein

R^D is hydrogen, deuterium, or optionally substituted C_1-C_3 alkyl;

R^{18} is an optionally substituted 5 to 6-membered heteroaryl;

R^{20} is hydrogen, deuterium, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

x_3 is an integer from 1-3; and

y_3 is an integer from 1-4.

[00155] In some embodiments, L^2 is an optionally substituted alkylene. In some embodiments, L^2 is C_2-C_4 alkylene, optionally substituted with one or more C_1-C_3 alkyl. In some embodiments, L^2 is absent.

[00156] In some embodiments, L^2 is -NR^D-. In some embodiments, L^2 is -NH-.

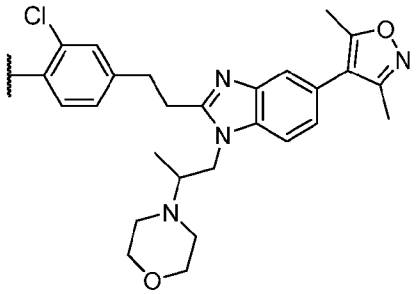
[00157] In some embodiments, R^{18} is an optionally substituted 5-membered heteroaryl. In some embodiments, R^{18} is optionally substituted oxazole, oxadiazole, thiazole, thiadiazole, pyrrole, or pyrazole. In some embodiments, R^{18} is optionally substituted oxazole.

[00158] In some embodiments, R^{20} is halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

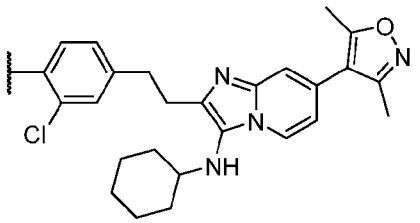
[00159] In some embodiments, x_3 is 1. In some embodiments, x_3 is 2. In some embodiments, x_3 is 3.

[00160] In some embodiments, y_4 is 1 or 2. In some embodiments, y_4 is 1. In some embodiments, y_4 is 2. In some embodiments, y_4 is 3. In some embodiments, y_4 is 4.

[00161] In some embodiments, the second terminus comprises the structure of Formula (4-E) or Formula (4-F), or a pharmaceutically acceptable salt thereof:

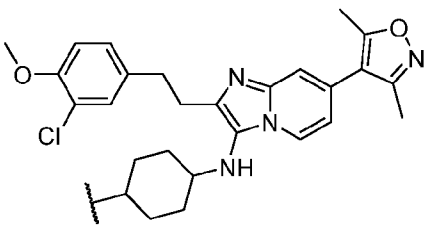


Formula (4-E) , or



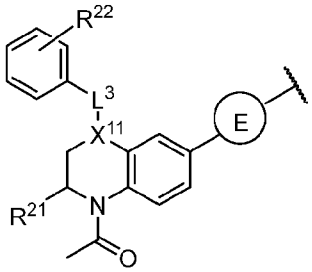
Formula (4F).

[00162] In some embodiments, the second terminus comprises the structure of Formula (4-G), or a pharmaceutically acceptable salt thereof:



Formula (4-G).

[00163] In some embodiments, the second terminus comprises the structure of Formula (5-A):



Formula (5-A),

or a pharmaceutically acceptable salt thereof, wherein;

Ring E is absent or an optionally substituted phenyl or optionally substituted 5 to 6-membered heteroaryl;

X^{11} is CH or N;

L^3 is $-NR^E-$ or $-CR^ER^E-$, wherein

each R^E is independently hydrogen, deuterium, or optionally substituted C_1-C_3 alkyl;

R^{21} is C_1-C_6 alkyl or C_3-C_6 cycloalkyl; and

R^{22} is halogen, CN, NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

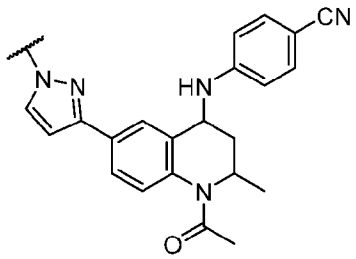
[00164] In some embodiments, Ring E is absent. In some embodiments, Ring E is an optionally substituted phenyl. In some embodiments, Ring E is an optionally substituted 5 to 6-membered heteroaryl. In some embodiments, Ring E is a 5-membered heteroaryl. In some embodiments, Ring E is a 6-membered heteroaryl.

[00165] In some embodiments, X^{11} is CH and L^3 is $-NR^E-$. In some embodiments, X^{11} is N and L^3 is $-CR^ER^E-$.

[00166] In some embodiments, R^{21} is C_1-C_6 alkyl. In some embodiments, R^{21} is methyl.

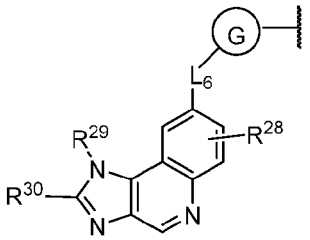
[00167] In some embodiments, R^{22} is halogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl. In some embodiments, R^{22} is CN, F, Cl, Br, or methyl.

[00168] In some embodiments, the second terminus comprises the structure of Formula (5-B), or a pharmaceutically acceptable salt thereof:



Formula (5-B).

[00169] In some embodiments, the second terminus comprises the structure of Formula (6-A):



Formula (6-A),

or a pharmaceutically acceptable salt thereof, wherein;

Ring G is an optionally substituted C_3-C_6 cycloalkyl or optionally substituted 4 to 6-membered heterocycloalkyl;

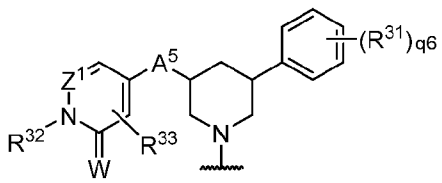
L^6 is -O-(optionally substituted alkylene);

R^{28} is an optionally substituted 5 to 6-membered heteroaryl;

R^{29} is optionally substituted C_1-C_6 alkyl(C_6-C_{10} aryl) or optionally substituted C_1-C_6 alkyl(6 to 10-membered heteroaryl); and

R^{30} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

[00170] In some embodiments, the second terminus comprises the structure of Formula (7-A):



Formula (7-A),

or a pharmaceutically acceptable salt thereof, wherein,

A^5 is -O-, -NH-, or -CH_2-;

Z^1 is CH or N;

W is O or S;

each R^{31} is independently hydrogen, halogen, -OH, -CN, -NO_2, -NH_2, optionally substituted C_1-C_{10} alkyl, optionally substituted C_1-C_{10} haloalkyl, optionally substituted C_1-C_{10} hydroxyalkyl, optionally substituted C_2-C_{10} alkenyl, optionally substituted C_2-C_{10} alkynyl, optionally substituted C_3-C_8 cycloalkyl, or optionally substituted 3 to 8-membered heterocycloalkyl;

or two R^{31} together with the atoms to which they are attached join together to form an optionally substituted C_5-C_8 cycloalkyl or optionally substituted 5 to 8-membered heterocycloalkyl;

R^{32} is hydrogen or optionally substituted C_1-C_{10} alkyl;

R^{33} is hydrogen, halogen, -OH, -CN, -NO_2, -NH_2, optionally substituted C_1-C_{10} alkyl, optionally substituted C_1-C_{10} haloalkyl, or optionally substituted C_1-C_{10} hydroxyalkyl; and

q_6 is 0-4.

[00171] In some embodiments, Z^1 is CH. In some embodiments, Z^1 is N.

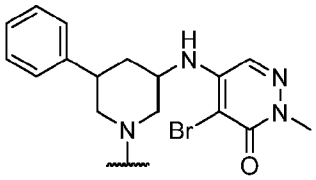
[00172] In some embodiments, W is O. In some embodiments, W is S.

[00173] In some embodiments, each R^{31} is independently an optionally substituted C_1-C_{10} alkyl, optionally substituted C_1-C_{10} haloalkyl, or optionally substituted C_1-C_{10} hydroxyalkyl. In some embodiments, each R^{31} is independently an optionally substituted C_3-C_8-cycloalkyl or optionally substituted 3 to 8-membered heterocycloalkyl. In some embodiments, each R^{31} is independently hydrogen, halogen, -OH, -CN, $-NO_2$, or $-NH_2$. In some embodiments, each R^{31} is hydrogen.

[00174] In some embodiments, R^{32} is an optionally substituted C_1-C_{10} alkyl. In some embodiments, R^{32} is methyl. In some embodiments, R^{32} is hydrogen.

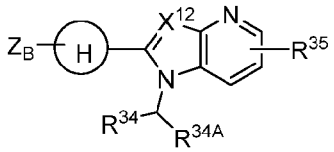
[00175] In some embodiments, R^{33} is hydrogen, halogen, -OH, -CN, $-NO_2$, or $-NH_2$. In some embodiments, R^{33} is an optionally substituted C_1-C_{10} alkyl, optionally substituted C_1-C_{10} haloalkyl, or optionally substituted C_1-C_{10} hydroxyalkyl.

[00176] In some embodiments, the second terminus comprises the structure of Formula (7-B), or a pharmaceutically acceptable salt thereof:



Formula (7-B).

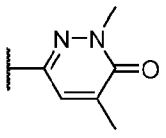
[00177] In some embodiments, the second terminus comprises the structure of Formula (8-A):



Formula (8-A),

or a pharmaceutically acceptable salt thereof, wherein,

Ring H is an optionally substituted phenyl or optionally substituted 6-membered heteroaryl;

or Ring H is  ;

Z_B is absent or an optionally substituted phenyl formamide;

X^{12} is CH or N;

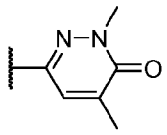
R^{34} is an optionally substituted phenyl or optionally substituted 6-membered heteroaryl;

R^{34A} is hydrogen, halogen, or optionally substituted C_1-C_3 alkyl;

R^{33} is independently halogen, optionally substituted C_1-C_{10} alkyl, or optionally substituted 5-membered heteroaryl; and

wherein the attachment to the linker is at R^{35}, Z_B, or Ring H.

[00178] In some embodiments, Ring H is an optionally substituted phenyl. In some embodiments, Ring H is an optionally substituted 6-membered heteroaryl.



[00179] In some embodiments, Ring H is .

[00180] In some embodiments, Z_B is absent. In some embodiments, Z_B is an optionally substituted phenyl formamide. In some embodiments, Z_B is -C(O)NH-phenyl.

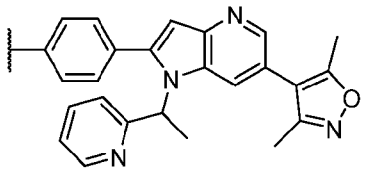
[00181] In some embodiments, X^{12} is CH. In some embodiments, X^{12} is N.

[00182] In some embodiments, R^{34} is an optionally substituted phenyl. In some embodiments, R^{34} is an optionally substituted 6-membered heteroaryl.

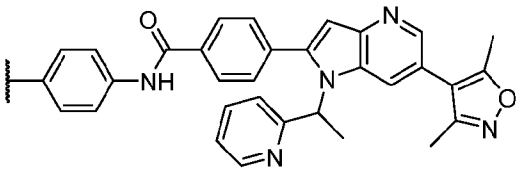
[00183] In some embodiments, R^{34A} is hydrogen or halogen. In some embodiments, R^{34A} is an optionally substituted C_1-C_3 alkyl. In some embodiments, R^{34A} is methyl.

[00184] In some embodiments, Formula (8-A) is attachment to the linker is at R^{35}. In some embodiments, Formula (8-A) is attached to the linker at Z_B. In some embodiments, Formula (8-A) is attached to the linker at Ring H.

[00185] In some embodiments, the second terminus comprises the structure of Formula (8-B) or Formula (8-C), or a pharmaceutically acceptable salt thereof:

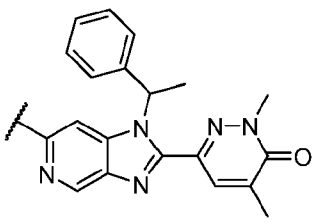


Formula (8-B) or



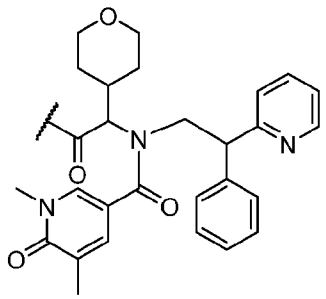
Formula (8-C).

[00186] In some embodiments, the second terminus comprises the structure of Formula (8-D), or a pharmaceutically acceptable salt thereof:



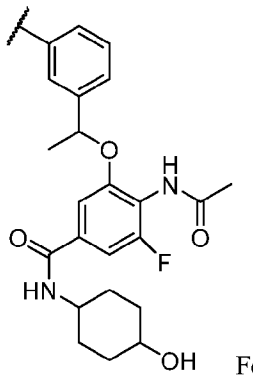
Formula (8-D).

[00187] In some embodiments, the second terminus comprises the structure of Formula (9-A), or a pharmaceutically acceptable salt thereof:

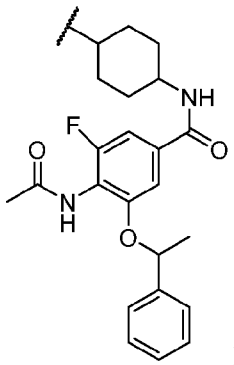


Formula (9-A).

[00188] In some embodiments, the second terminus comprises the structure of Formula (10-A) or Formula (10-B), or a pharmaceutically acceptable salt thereof:

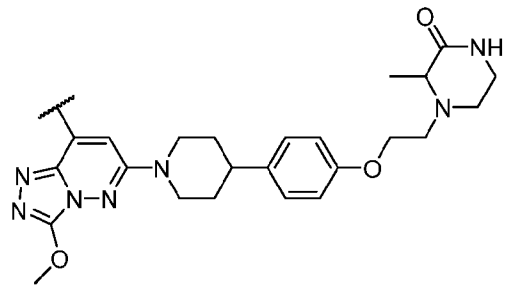


Formula (10-A) or



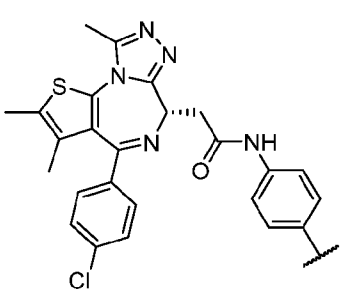
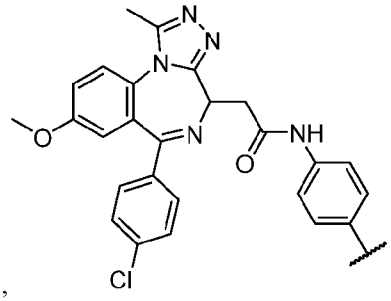
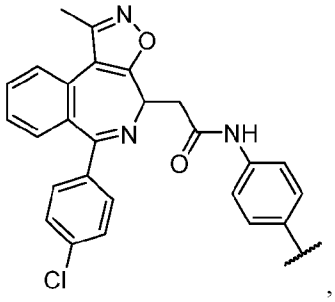
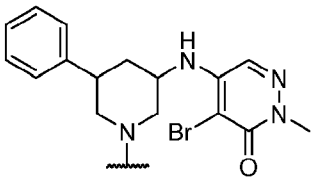
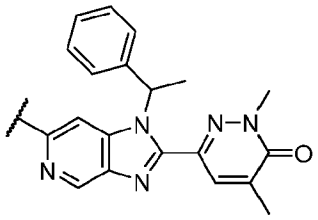
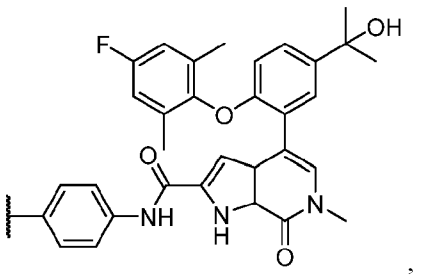
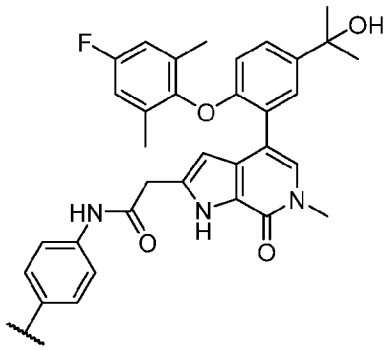
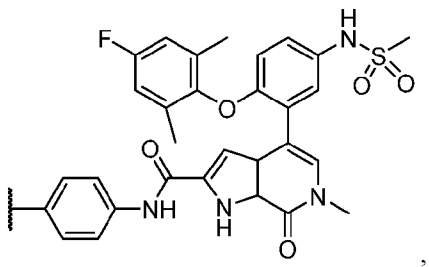
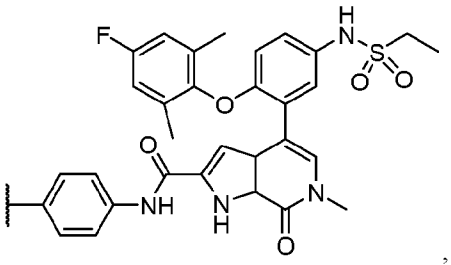
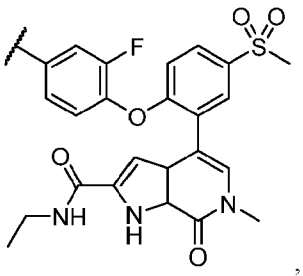
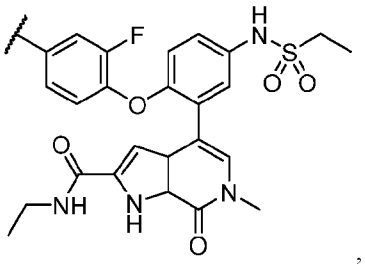
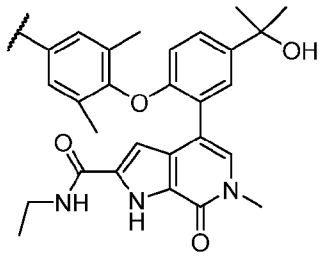
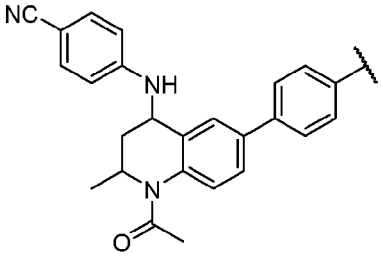
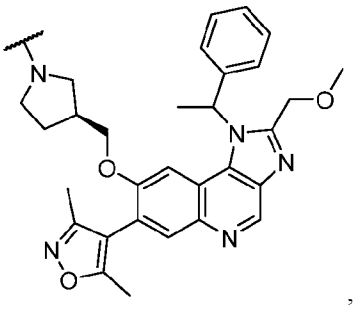
Formula (10-B).

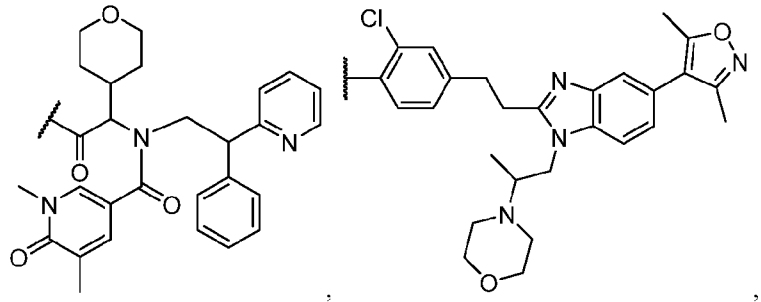
[00189] In some embodiments, the second terminus comprises the structure of Formula (11-A), or a pharmaceutically acceptable salt thereof:

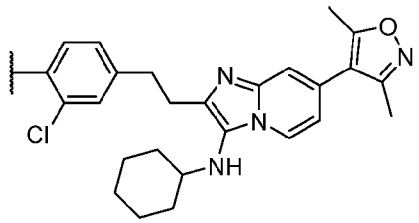
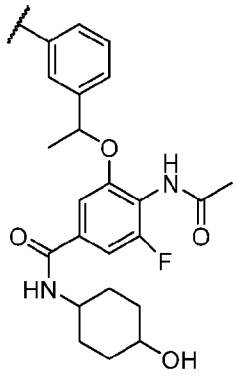
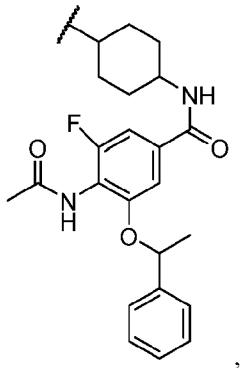


Formula (11-A).

[00190] In some embodiments, the second terminus is selected from the group consisting of:

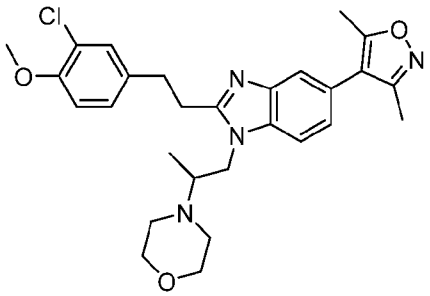
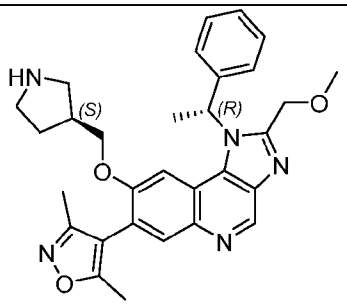

,

,

,

,

,

,

,

,

,

,

,

,

,

,

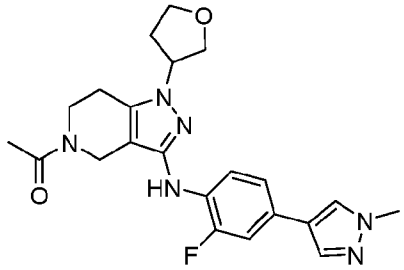
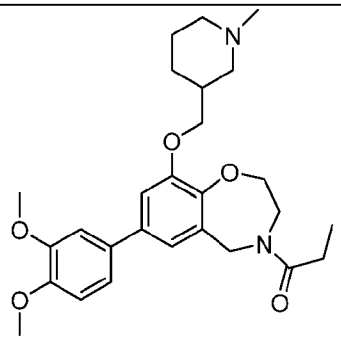
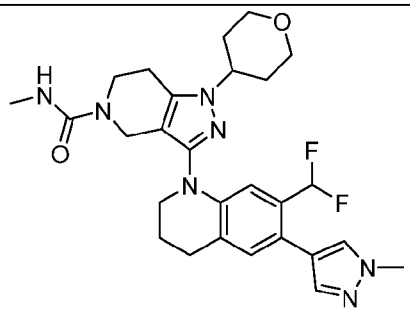
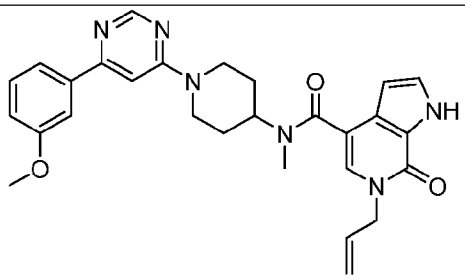
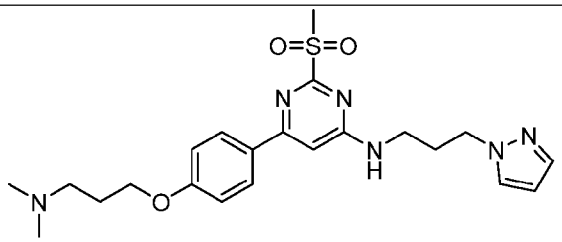
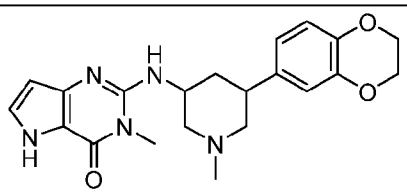

,

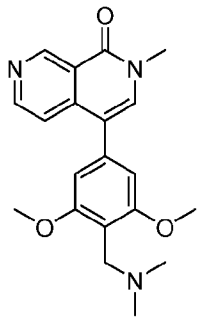
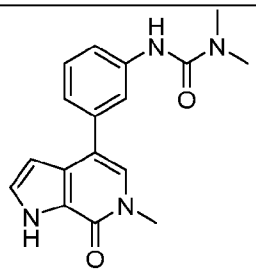
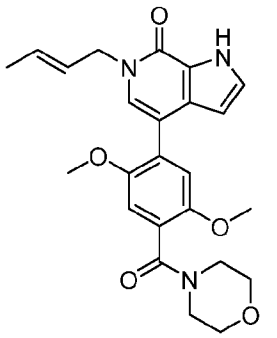
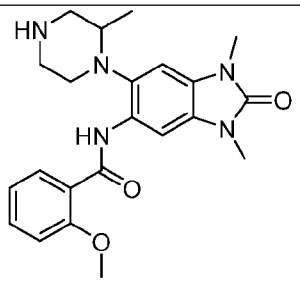
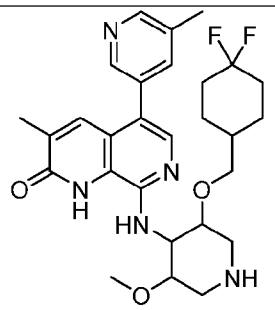
,

,

, and

, or a pharmaceutically acceptable salt thereof.

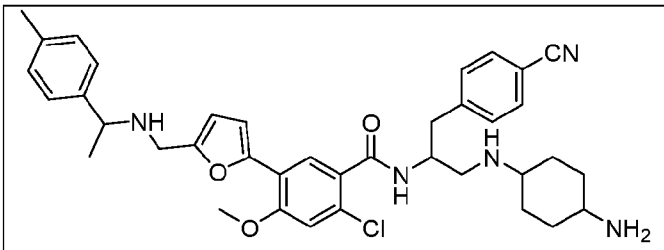
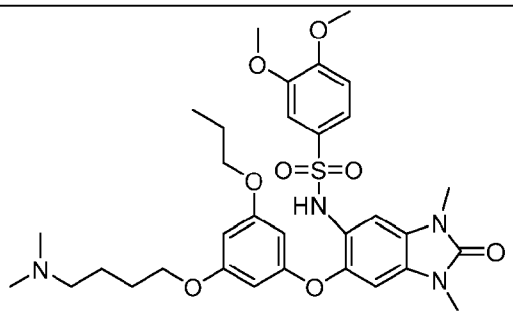
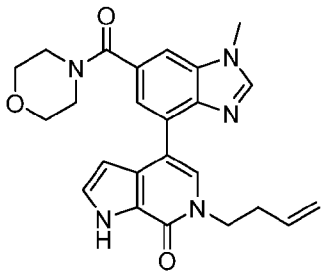
[00191] In some embodiments, the second terminus is selected from a moiety described in Table 2, or a pharmaceutically acceptable salt thereof.

Table 2. Exemplary bromodomain binding moieties.

Structure	Binder
	CBP/P300
	BET(BD1)
	PCAF

	
	CBP/P300
	CBP/P300
	CECR2
	BPTF
	PCAF

	BRD7/9
	TAF1
	BRD7/9
	BRPF
	ATAD2/ATAD2B

	ATAD2/ATAD2B
	TRIM24
	TAF1

Oligomeric Backbone – Linker Moiety

[00192] The oligomeric backbone is a linker that connects the first terminus and the second terminus and brings the regulatory molecule in proximity to the target gene to modulate gene expression.

[00193] The length of the linker depends on the type of regulatory protein and also the target gene. In some embodiments, the linker has a length of less than about 50 Angstroms. In some embodiments, the linker has a length of about 20 to 30 Angstroms.

[00194] In some embodiments, the oligomeric backbone comprises between 5 and 50 chain atoms.

[00195] In some embodiments, the oligomeric backbone comprises a multimer having 2 to 50 spacing moieties, wherein

each spacing moiety is independently selected from the group consisting of -$((CR^{3a}R^{3b})_x$-$O)_y$-, -$((CR^{3a}R^{3b})_x$-$NR^{4a})_y$-, -$((CR^{3a}R^{3b})_x$-CH=CH-$(CR^{3a}R^{3b})_x$-$O)_y$-, optionally substituted C_1-C_{12} alkyl, optionally substituted C_2-C_{10} alkenyl, optionally substituted C_2-C_{10} alkynyl, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5- to 10-membered heteroarylene, optionally substituted 4- to 10-membered heterocycloalkylene, amino acid residue, -O-, -C(O)NR^{4a}-, -NR^{4a}C(O)-, -C(O)-, -NR^{1a}-, -C(O)O-, -S-, -S(O)-, -$S(O)_2$-, -$S(O)_2NR^{4a}$-, -$NR^{4a}S(O)_2$-, and -P(O)OH-, and any combinations thereof; wherein

each x is independently 2-4;

each y is independently 1-10;

each R^{1a} is independently a hydrogen or optionally substituted C_1-C_6 alkyl;

each R^{3a} and R^{3b} is independently selected from hydrogen, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted amino, carboxyl, carboxyl ester, acyl, acyloxy, acyl amino, amino acyl, optionally substituted alkylamide, sulfonyl, optionally substituted thioalkoxy, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted cycloalkyl, and optionally substituted heterocyclyl; and

each R^{4a} is independently a hydrogen or an optionally substituted C_1-C_6 alkyl.

[00196] In some embodiments, the oligomeric backbone comprises a multimer having 2 to 50 spacing moieties, wherein each spacing moiety is independently selected from the group consisting of optionally substituted C_1-C_{12} alkyl, $-((CH_2)_x\text{-}O)_y-$, $-((CH_2)_x\text{-}NH)_y-$, -O-, -C(O)NH-, -NH-, and any combinations thereof.

[00197] In some embodiments, the oligomeric backbone comprises $-(T^1\text{-}V^1)_a\text{-}(T^2\text{-}V^2)_b\text{-}(T^3\text{-}V^3)_c\text{-}(T^4\text{-}V^4)_d\text{-}(T^5\text{-}V^5)_e-$, wherein

a, b, c, d and e are each independently 0 or 1, and where the sum of a, b, c, d and e is 1 to 5;

T^1, T^2, T^3, T^4 and T^5 are each independently selected from an optionally substituted C_1-C_{12} alkylene, optionally substituted alkenylene, optionally substituted alkynylene, $(EA)_w$, $(EDA)_m$, $(PEG)_n$, (modified $PEG)_n$, $(AA)_p$, $-(CR^{2a}OH)_h-$, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5- to 10-membered heteroarylene, optionally substituted 4- to 10-membered heterocycloalkylene, an acetal group, a disulfide, a hydrazine, a carbohydrate, a beta-lactam, and an ester,

(a) w is an integer from 1 to 20;

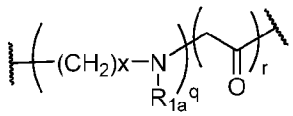
(b) m is an integer from 1 to 20;

(c) n is an integer from 1 to 30;

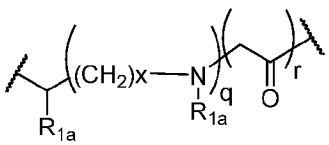
(d) p is an integer from 1 to 20;

(e) h is an integer from 1 to 12;

(f) EA has the following structure



(g) EDA has the following structure:



wherein each q is independently an integer from 1 to 6, each x is independently an integer from 1 to 4, and each r is independently 0 or 1;

(h) $(PEG)_n$ has the structure of $-(CR^{2a}R^{2b}\text{-}CR^{2a}R^{2b}\text{-}O)_n\text{-}CR^{2a}R^{2b}-$;

(i) (modified $PEG)_n$ has the structure of replacing at least one $-(CR^{2a}R^{2b}-CR^{2a}R^{2b}-O)-$ in $(PEG)_n$ with $-(CH_2-CR^{2a}=CR^{2a}-CH_2-O)-$ or $-(CR^{2a}R^{2b}-CR^{2a}R^{2b}-S)-$;

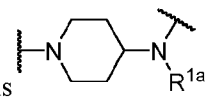
(j) AA is an amino acid residue;

(k) V^1, V^2, V^3, V^4 and V^5 are each independently selected from the group consisting of a bond, C(O)-, $-NR^{1a}-$, $-C(O)NR^{1a}-$, $-NR^{1a}C(O)-$, $-CONR^{1a}-C_1-C_4$ alkyl-, $-NR^{1a}C(O)-C_1-C_4$ alkyl-, -C(O)O-, -OC(O)-, -O-, -S-, -S(O)-, $-S(O)_2-$, $-S(O)_2NR^{1a}-$, $-NR^{1a}S(O)_2-$ and -P(O)OH-;

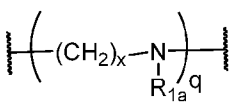
(l) each R^{1a} is independently hydrogen or and optionally substituted C_1-C_6 alkyl; and each R^{2a} and R^{2b} is independently selected from hydrogen, alkyl, substituted alkyl, alkenyl, substituted alkenyl, alkynyl, substituted alkynyl, halogen, alkoxy, substituted alkoxy, amino, substituted amino, carboxyl, carboxyl ester, acyl, acyloxy, acyl amino, amino acyl, alkylamide, substituted alkylamide, sulfonyl, thioalkoxy, substituted thioalkoxy, aryl, substituted aryl, heteroaryl, substituted heteroaryl, cycloalkyl, substituted cycloalkyl, heterocyclyl, and substituted heterocyclyl.

[00198] In some embodiments, the a, b, c, d and e are each independently 0 or 1, where the sum of a, b, c, d and e is 1. In some embodiments, the a, b, c, d and e are each independently 0 or 1, where the sum of a, b, c, d and e is 2. In some embodiments, the a, b, c, d and e are each independently 0 or 1, where the sum of a, b, c, d and e is 3. In some embodiments, the a, b, c, d and e are each independently 0 or 1, where the sum of a, b, c, d and e is 4. In some embodiments, the a, b, c, d and e are each independently 0 or 1, where the sum of a, b, c, d and e is 5.

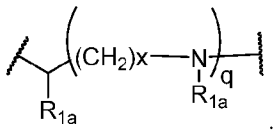
[00199] In some embodiments, n is 3-9. In some embodiments, n is 4-8. In some embodiments, n is 5 or 6.

[00200] In some embodiments, T^1, T^2, T^3, and T^4, and T^5 are each independently selected from C_1-C_{12} alkyl, substituted C_1-C_{12} alkyl, $(EA)_w$, $(EDA)_m$, $(PEG)_n$, (modified $PEG)_n$, $(AA)_p$, $-(CR^{2a}OH)_h-$, phenyl, substituted phenyl, piperidin-4-amino (P4A), para-amino-benzyloxycarbonyl (PABC), meta-amino-benzyloxycarbonyl (MABC), para-amino-benzyloxy (PABO), meta-amino-benzyloxy (MABO), para-aminobenzyl, an acetal group, a disulfide, a hydrazine, a carbohydrate, a beta-lactam, an ester, $(AA)_p$-MABC-$(AA)_p$, $(AA)_p$-MABO-$(AA)_p$, $(AA)_p$-PABO-$(AA)_p$ and $(AA)_p$-PABC-$(AA)_p$. In some embodiments, piperidin-4-amino (P4A) is , wherein R^{1a} is hydrogen or C_1-C_6 alkyl.

[00201] In some embodiments, T^1, T^2, T^3, T^4 and T^5 are each independently selected from (C_1-C_{12})alkyl, substituted C_1-C_{12} alkyl, $(EA)_w$, $(EDA)_m$, $(PEG)_n$, (modified $PEG)_n$, $(AA)_p$,$-(CR^{2a}OH)_h-$, optionally substituted C_6-C_{10} arylene, 4 to 10-membered heterocycloalkene, optionally substituted 5 to 10-membered heteroarylene. In some embodiments, EA has the following structure:

; and

EDA has the following structure:



[00202] In some embodiments, x is 2-3 and q is 1-3 for EA and EDA. In some embodiments, R^{1a} is hydrogen or C_1-C_6 alkyl.

[00203] In some embodiments, T^4 or T^5 is an optionally substituted C_6-C_{10} arylene.

[00204] In some embodiments, T^4 or T^5 is phenylene or substituted phenylene. In some embodiments, T^4 or T^5 is phenylene or phenylene substituted with 1-3 substituents selected from C_1-C_6 alkyl, halogen, OH or amine. In some embodiments, T^4 or T^5 is 5 to 10-membered heteroarylene or substituted heteroarylene. In some embodiments, T^4 or T^5 is 4 to 10-membered heterocyclene or substituted heterocylcylene. In some embodiments, T^4 or T^5 is heteroarylene or heterocylene optionally substituted with 1-3 substituents selected from C_1-C_6 alkyl, halogen, OH or amine.

[00205] In some embodiments, T^1, T^2, T^3, T^4 and T^5 and V^1, V^2, V^3, V^4 and V^5 are selected from the following Table 3.

Table 3. Exemplary linker units.

T^1	V^1	T^2	V^2	T^3	V^3	T^4	V^4	T^5	V^5
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	$NR^{1a}CO$	----	----	----	----
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	O	arylene	$NR^{1a}CO$	----	----
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	O	Subst. arylene	$NR^{1a}CO$	----	----
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	O	$NR^{1a}C$ O	C_1-C_{12} alkyl	Subst. arylene	$NR^{1a}CO$
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	C_1-C_{12} alkyl	$NR^{1a}CO$-C_1-C_4 alkyl	Subst. arylene	NR^{11}	----	----
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	O	Subst. arylene	---	----	----
$(PEG)_n$	C(O)NR^{1a}-C_1-C_4 alkyl	----	----	----	----	----	----	----	---
$(EA)_w$	CO	C_1-C_{12} alkyl	C(O)NR^{1a}-C_1-C_4 alkyl	----	----	----	----	----	---
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	$(PEG)_n$	$NR^{1a}CO$-C_1-C_4 alkyl	----	----	----	----
$(EA)_w$	CO	$(PEG)_n$	O	phenyl	$NR^{1a}CO$-C_1-C_4 alkyl	----	----	----	----
C_1-C_{12} alkylene	C(O)NR^{1a}	$(PEG)_n$	CO	----	----	----	----	----	---
C_1-C_{12} alkylene	C(O)NR^{1a}	$(EA)_w$	CO	modifd. $(PEG)_n$	O	arylene	$NR^{1a}CO$	----	----

[00206] In some embodiments, the oligomeric backbone comprises $N(R^{1a})(CH_2)_xN(R^{1b})(CH_2)_xN–$, wherein R^{1a} and R^{1b} are each independently selected from hydrogen or optionally substituted C_1-C_6 alkyl; and each x is independently an integer in the range of 1-6.

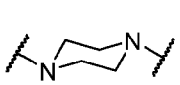
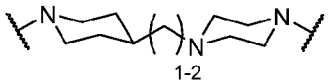
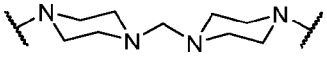
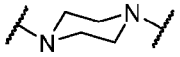
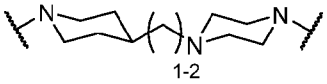
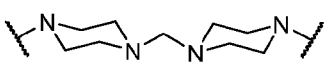
[00207] In some embodiments, the oligomeric backbone comprises $-(CH_2-C(O)N(R^{4a})-(CH_2)_q-N(R^{4a})-(CH_2)_q-N(R^{4a})C(O)-(CH_2)_x-C(O)N(R^{4a})-A-$, $-(CH_2)_x-C(O)N(R^{4a})-(CH_2CH_2O)_y(CH_2)_x-C(O)N(R^{4a})-A-$, or $-C(O)N(R^{4a})-(CH_2)_q-N(R^{4a})-(CH_2)_q-N(R^{4a})C(O)-(CH_2)_x-A-$; wherein each q is independently an integer from 2 to 10; each x is independently an integer from 1-6; and each A is independently selected from a bond, an optionally substituted C_1-C_{12} alkyl, an optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5 to 10-membered heteroarylene, and optionally substituted 4 to 10-membered heterocycloalkylene.

[00208] In some embodiments, the oligomeric backbone comprises $–(CH_2CH_2-O)_{x1}-$ or $–(CH_2CH_2-O)_{x2}-A-(CH_2CH_2-O)_{x3}-$, wherein A is an optionally substituted 4 to 10-membered heterocycloalkylene or spirocyclene, and each x_1, x_2, and x_3 is independently an integer from 1-15.

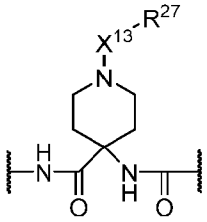
[00209] In some embodiments, the oligomeric backbone comprises $-NR^{4a}-(CH_2CH_2O)_y(CH_2)_x-$ or $-NR^{4a}-(CH_2)_q-C(O)NR^{4a}(CH_2CH_2O)_y(CH_2)_x-$, wherein q is 2-10, x is 1-4, y is 1-50, and each R^{4a} is independently hydrogen or an optionally substituted C_1-C_6 alkyl. In some embodiments, the oligomeric backbone comprises $-NR^{4a}-(CH_2CH_2O)_y(CH_2)_x$,. In some embodiments, the oligomeric backbone comprises $-NR^{4a}-(CH_2)_q-C(O)NR^{4a}(CH_2CH_2O)_y(CH_2)_x-$.

[00210] In some embodiments, the oligomeric backbone comprises $-(CH_2CH_2-O)_{x1}-$, $-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NH-$, $-NH-(CH_2CH_2-O)_{x1}-$, $-NH-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NH-$, $-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NHC(O)-$, or $-NH-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NHC(O)-$. In some embodiments, the oligomeric backbone comprises $-NH-(CH_2CH_2-O)_{x1}-$ or $-NH-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NH-$. In some embodiments, the oligomeric backbone comprises $-NH-(CH_2CH_2-O)_{x1}-$. In some embodiments, the oligomeric backbone comprises $-NH-(CH_2CH_2-O)_{x1}-(CH_2CH_2)-NH-$.

[00211] In some embodiments, the oligomeric backbone comprises polyethylene glycol (PEG). In some embodiments, the oligomeric backbone comprises 1-20 PEG units. In some embodiments, the oligomeric backbone comprises 1, 2, 3, 4,5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 PEG units.

[00212] In some embodiments, A is selected from , , or . In some embodiments, A is . In some embodiments, A is . In some embodiments, A is .

[00213] In some embodiments, A comprises a moiety having the structure:



or a pharmaceutically acceptable salt thereof, wherein:

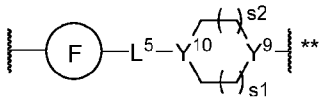
X^{13} is absent or -C(O)-; and

R^{27} is an optionally substituted C_1-C_{50} alkyl or optionally substituted C_1-C_{50} heteroalkyl.

[00214] In some embodiments, X^{13} is -C(O)-. In some embodiments, X^{13} is absent.

[00215] In some embodiments, R^{27} is C_1-C_{50} alkyl. In some embodiments, R^{27} is C_1-C_{40} alkyl. In some embodiments, R^{27} is C_1-C_{30} alkyl. In some embodiments, R^{27} is C_1-C_{20} alkyl. In some embodiments, R^{27} is C_1-C_{10} alkyl. In some embodiments, R^{27} is C_1-C_{50} heteroalkyl. In some embodiments, R^{26} is C_1-C_{40} heteroalkyl. In some embodiments, R^{27} is C_1-C_{30} heteroalkyl. In some embodiments, R^{27} is C_1-C_{20} heteroalkyl. In some embodiments, R^{27} is C_1-C_{10} heteroalkyl. In some embodiments, the heteroalkyl is polyethylene glycol (PEG).

[00216] In some embodiments, the oligomeric backbone comprises a moiety having a structure of Formula (C-1):



Formula (C-1),

or a pharmaceutically acceptable salt thereof, wherein:

Ring F is absent, arylene, or heterocycloalkylene;

L^5 is absent, optionally substituted alkylene, or optionally substituted alkynylene;

Y^9 and Y^{10} are each independently CH or N;

s_1 and s_2 are each independently 0-3; and

** denotes attachment to the second terminus.

[00217] In some embodiments, Ring F is absent. In some embodiments, Ring F is C_4-C_7 heterocycloalkylene.

[00218] In some embodiments, Y^9 is N. In some embodiments, Y^9 is CH.

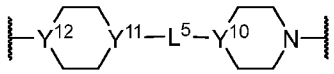
[00219] In some embodiments, Y^{10} is N. In some embodiments, Y^{10} is CH.

[00220] In some embodiments, L^5 is absent.

[00221] In some embodiments, L^5 is alkylene or alkynylene.

[00222] In some embodiments, L^5 is $-(C^{R1G}R^{1G})_x$-(alkylene)$_2$-$(CR^{1G}R^{1G})_y$-; wherein x and y are each independently 0 or 1; and each R^{1G} is hydrogen or C_1-C_3 alkyl.

[00223] In some embodiments, the oligomeric backbone comprises a moiety having a structure of Formula (C-2):



Formula (C-2),

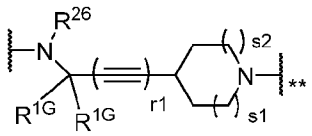
or a pharmaceutically acceptable salt thereof, wherein:

Y^{10}, Y^{11} and Y^{12} are each independently N or CH.

[00224] In some embodiments, each of Y^{11} and Y^{12} is independently N or CH; and Y^{10} is N.

[00225] In some embodiments, L^5 is C_1-C_3 alkylene or C_1-C_3 alkynylene. In some embodiments, L^5 is C_1-C_3 alkylene. In some embodiments, L^5 is C_1-C_3 alkynylene. In some embodiments, L^5 is -CH_2-, -CH_2CH_2-, —C≡C—, or —C≡C–C≡C—. In some embodiments, L^5 is -CH_2- or -CH_2CH_2-. In some embodiments, L^5 is —C≡C—. In some embodiments, L^5 is —C≡C–C≡C—.

[00226] In some embodiments, the oligomeric backbone comprises a moiety having the structure of Formula (C-3):



Formula (C-3),

or a pharmaceutically acceptable salt thereof, wherein:

s_1 and s_2 are each independently 0-3;

r_1 is an integer from 1-3;

R^{26} is an optionally substituted C_1-C_{20} alkylene or optionally substituted C_2-C_{20} heteroalkylene;

each R^{1G} is independently hydrogen or C_1-C_3 alkyl; and

** denotes attachment to the second terminus.

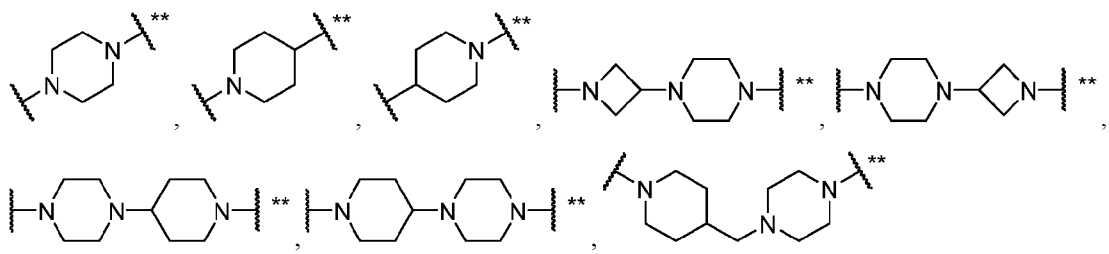
[00227] In some embodiments, R^{26} is an optionally substituted C_1-C_{20} heteroalkylene. In some embodiments, R^{26} is PEG.

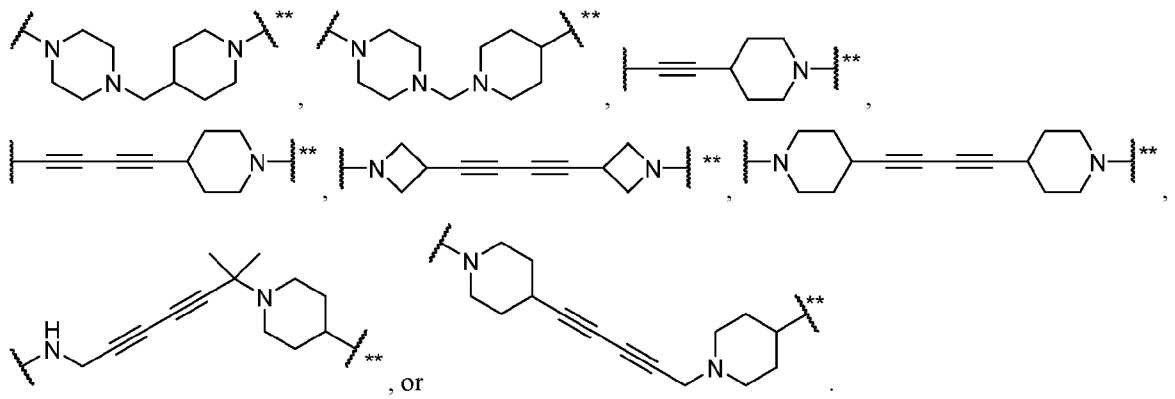
[00228] In some embodiments, each R^{1G} is independently hydrogen. In some embodiments, R^{1G} is independently C_1-C_3 alkyl. In some embodiments, the C_1-C_3 alkyl is methyl, ethyl or propyl. In some embodiments, each R^{1G} is independently methyl.

[00229] In some embodiments, s_1 and s_2 are each independently is 0, 1, or 2. In some embodiments, s_1 and s_2 are each independently 0. In some embodiments, s_1 and s_2 are each independently 1.

[00230] In some embodiments, r_1 is 1 or 2. In some embodiments, r_1 is 1. In some embodiments, r_1 is 2.

[00231] In some embodiments, the oligomeric backbone comprises:





[00232] In some embodiments, the oligomeric backbone is joined with the first terminus and/or with the second terminus with a group selected from -C(O)-, -NR^{1a}-, -C(O)NR^{1a}-, -NR^{1a}C(O)-, -C(O)$NR^{1a}$$C_1$-$C_4$alkyl-, -$NR^{1a}$C(O)-$C_1$-$C_4$alkyl-, -C(O)O-, -OC(O)-, -O-, -S-, -S(O)-, -$S(O)_2$-, -$S(O)_2NR^{1a}$-, -$NR^{1a}S(O)_2$-, -P(O)OH-, -$((CH_2)_x$-O)-, -$((CH_2)_y$-$NR^{1a})$-, optionally substituted C_1-C_{12} alkylene, optionally substituted C_2-C_{10} alkenylene, optionally substituted C_2-C_{10} alkynylene, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5 to 10-membered heteroarylene, and optionally substituted 4 to 10-membered heterocycloalkylene, wherein each x is independently 1-4, each y is independently 1-4, and each R^{1a} is independently a hydrogen or optionally substituted C_1-C_6 alkyl.

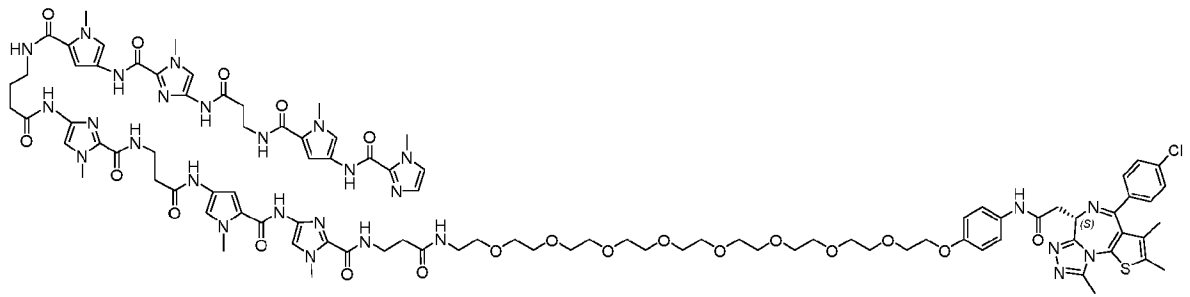
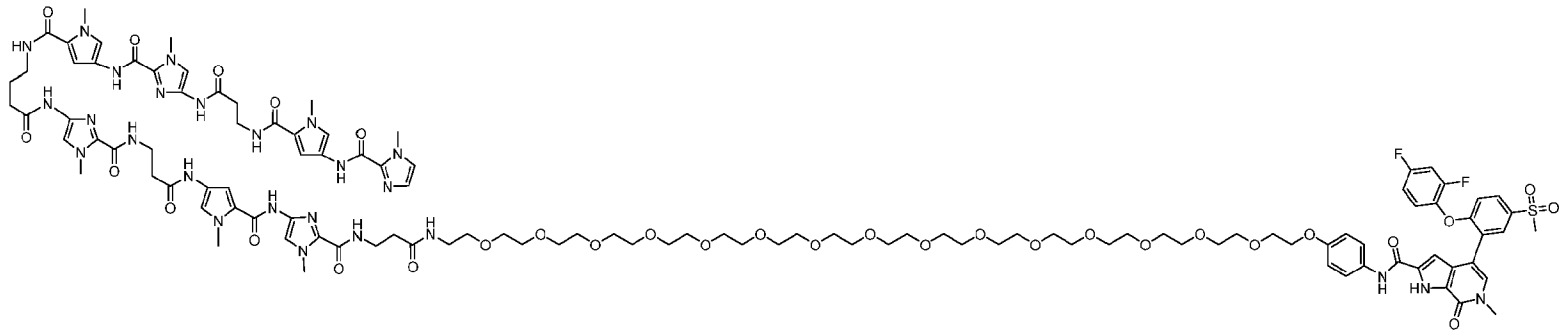
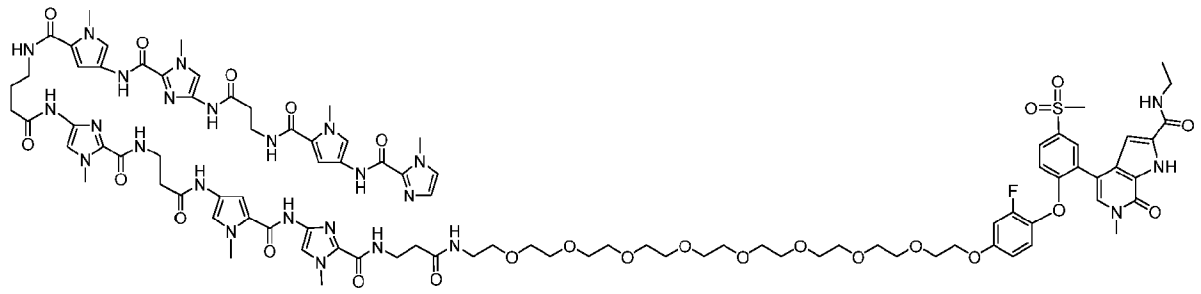
[00233] In some embodiments, the oligomeric backbone is joined with the first terminus with a group selected from -O-, -C(O)-, -NR^{1a}-, C_1-C_{12} alkyl, -C(O)NR^{1a}-, and -NR^{1a}C(O)-. In some embodiments, the oligomeric backbone is joined with the first terminus with a group selected from -O- or -NR^{1a}-.

[00234] In some embodiments, the oligomeric backbone is joined with the second terminus with a group selected from -C(O)-, -NR^{1a}-, -C(O)NR^{1a}-, -NR^{1a}C(O)-, -$((CH_2)_x$-O)-, -$((CH_2)_y$-$NR^{1a})$-, -O-, optionally substituted C_1-C_{12} alkyl, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5- to 10-membered heteroarylene, and optionally substituted 4- to 10-membered heterocycloalkylene, wherein each x is independently 1-4, each y is independently 1-4, and each R^{1a} is independently a hydrogen or optionally substituted C_1-C_6 alkyl.

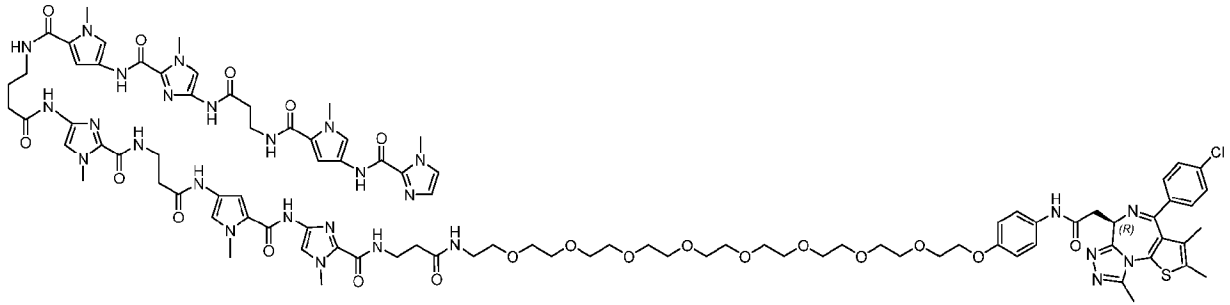
[00235] In some embodiments, the oligomeric backbone is joined with the second terminus with a group selected from -O-, -C(O)-, -NR^{1a}-, C_1-C_{12} alkyl, -C(O)NR^{1a}-, and -NR^{1a}C(O)-. In some embodiments, the oligomeric backbone is joined with the second terminus with a group selected from -O- or -NR^{1a}-. In some embodiments, the oligomeric backbone is joined with the second terminus with -O-. In some embodiments, the oligomeric backbone is joined with the second terminus with -NR^{1a}-. In some embodiments, the oligomeric backbone is joined with the second terminus with -NH-.

[00236] In some embodiments, non-limiting examples of the transcription modulator compounds described herein are presented below in Table 4 (next page).

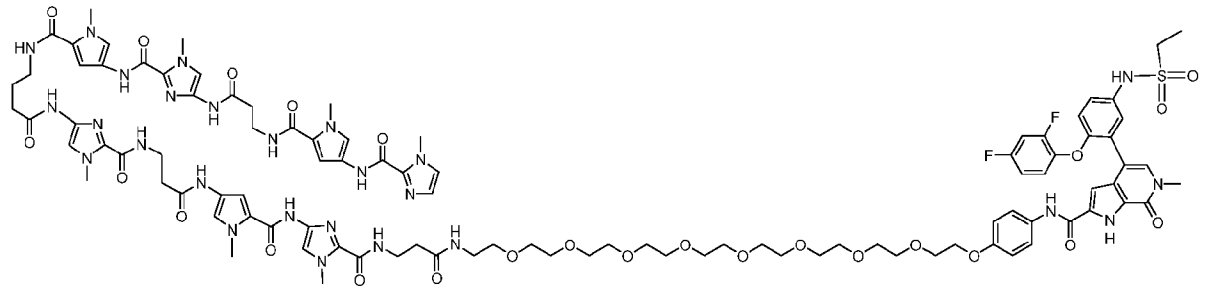
Table 4. Compounds of the disclosure.

Cmpd. No.	Structure
A-1	
A-2	
A-3	

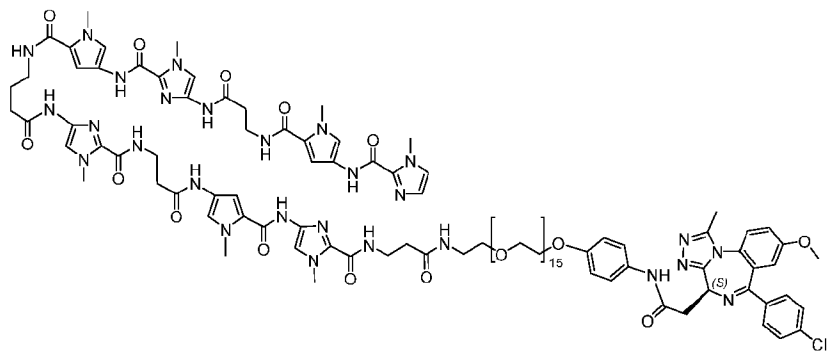
A-4



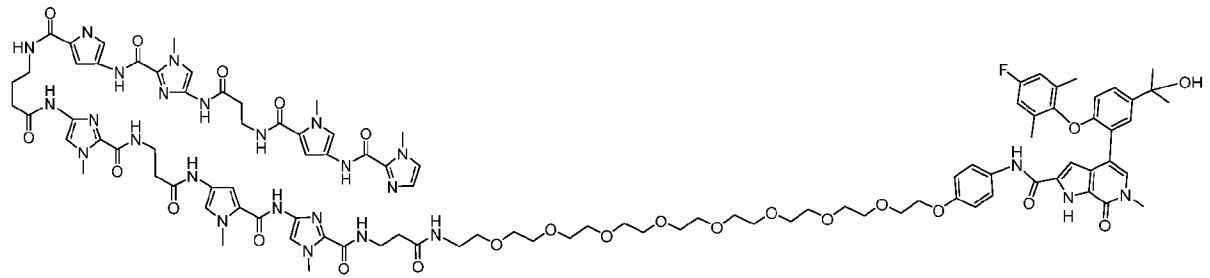
A-5



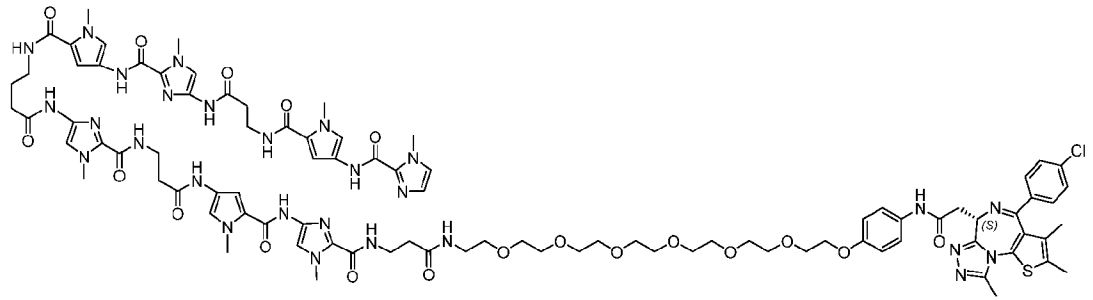
A-6



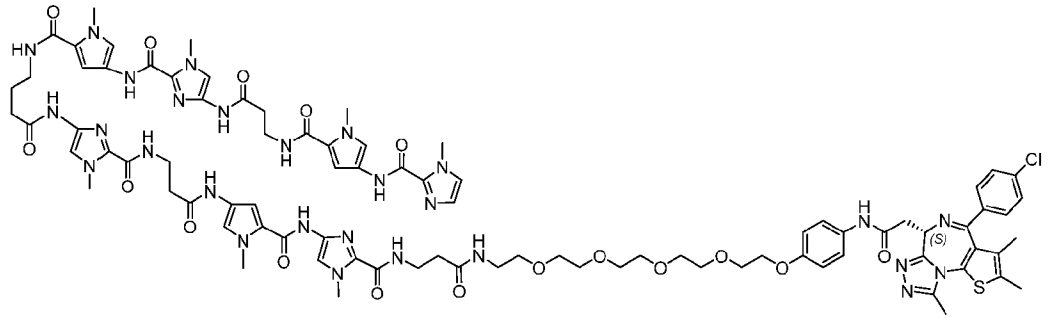
A-7



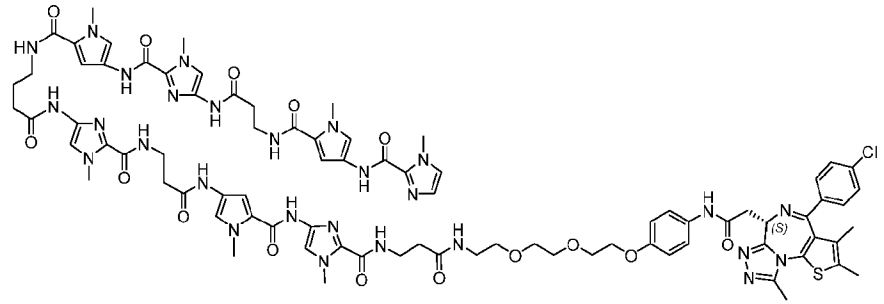
A-8



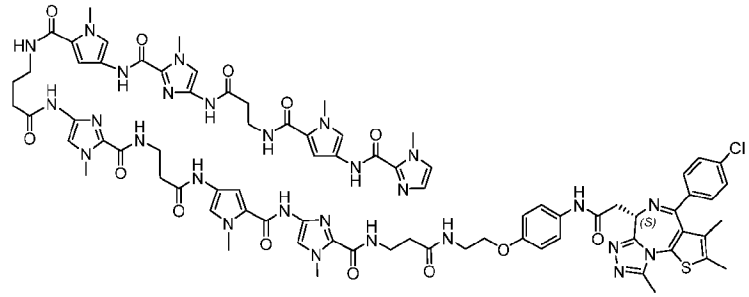
A-9



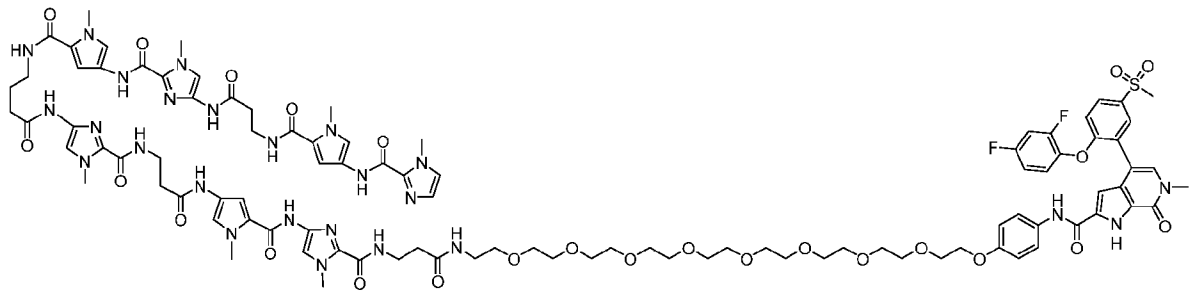
A-10



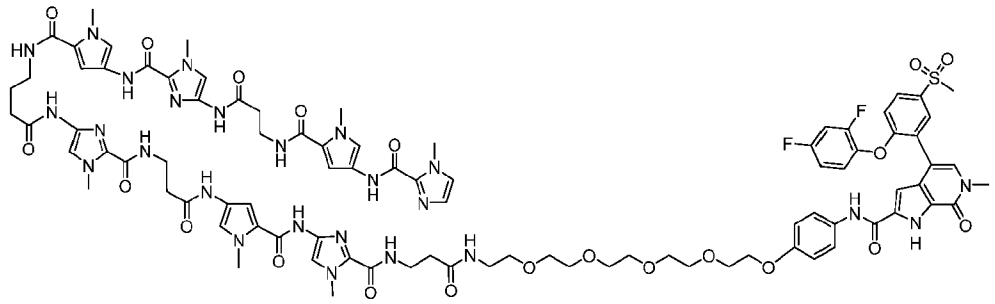
A-11



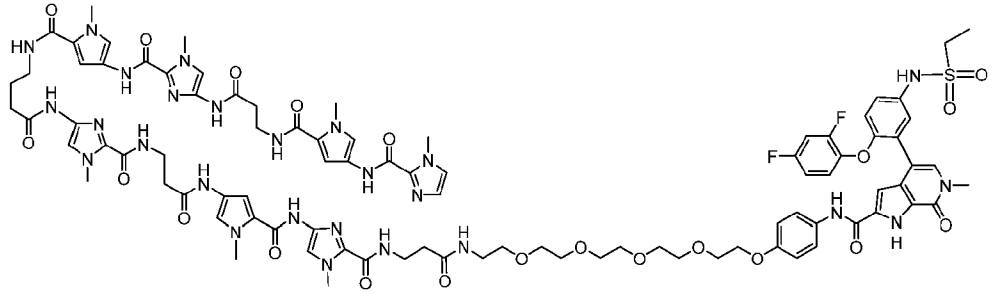
A-12



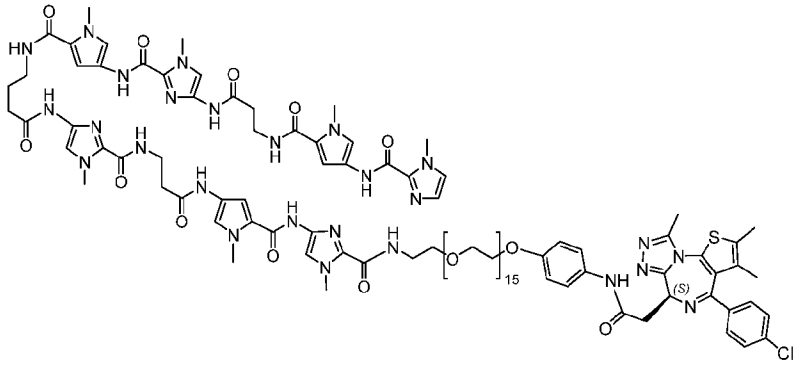
A-13



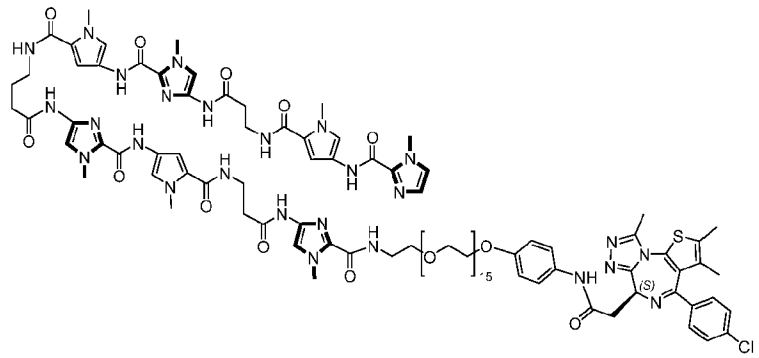
A-14



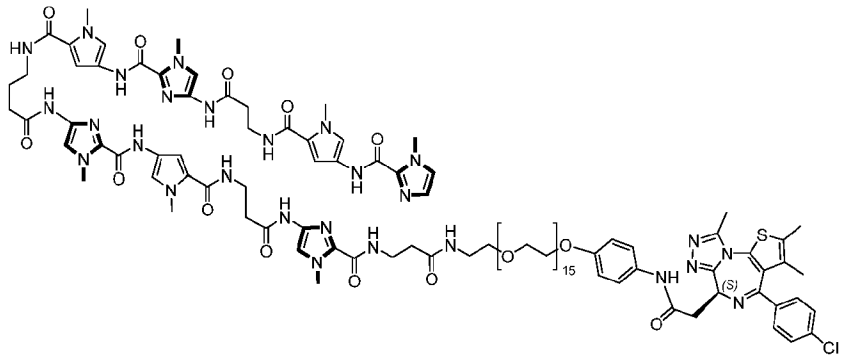
A-15



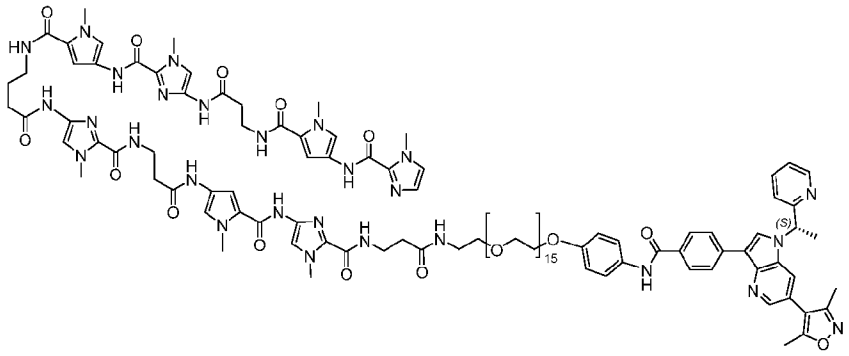
A-16



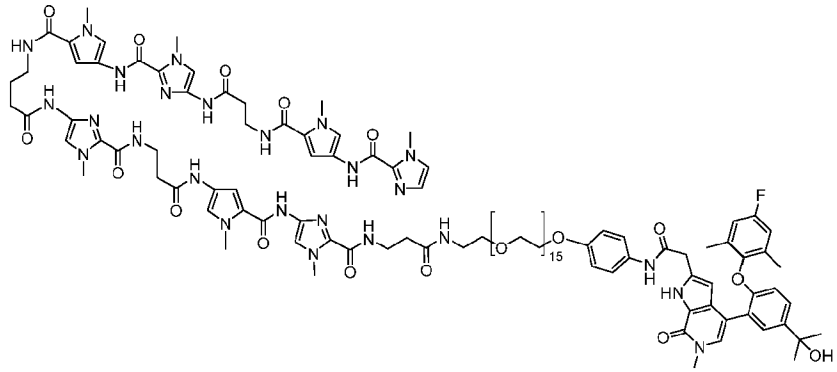
A-17



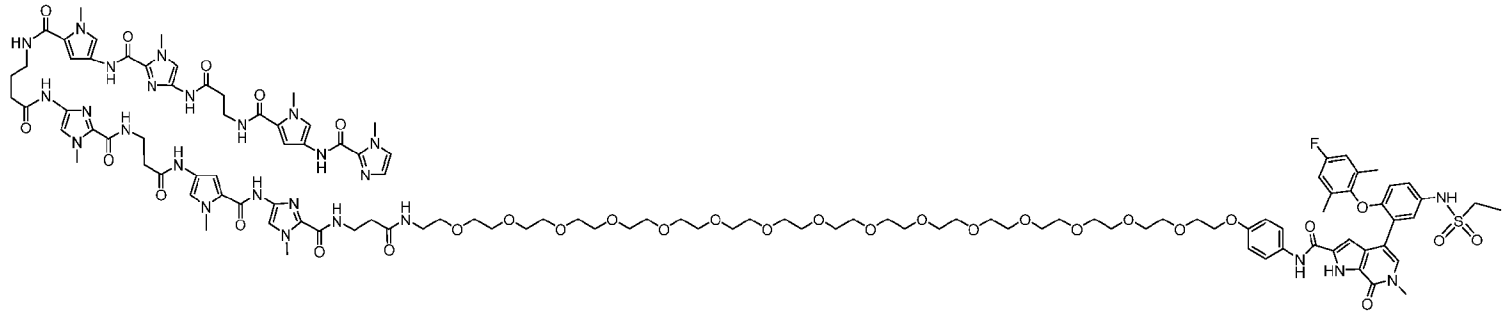
A-18



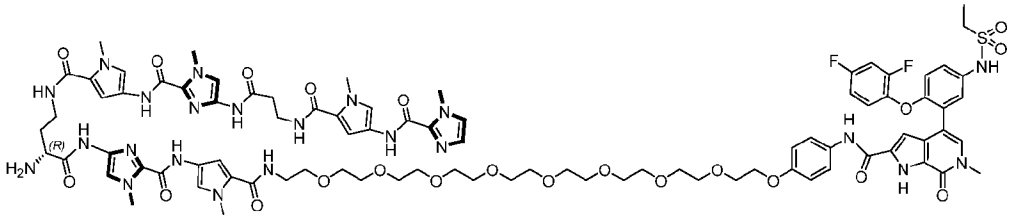
A-19



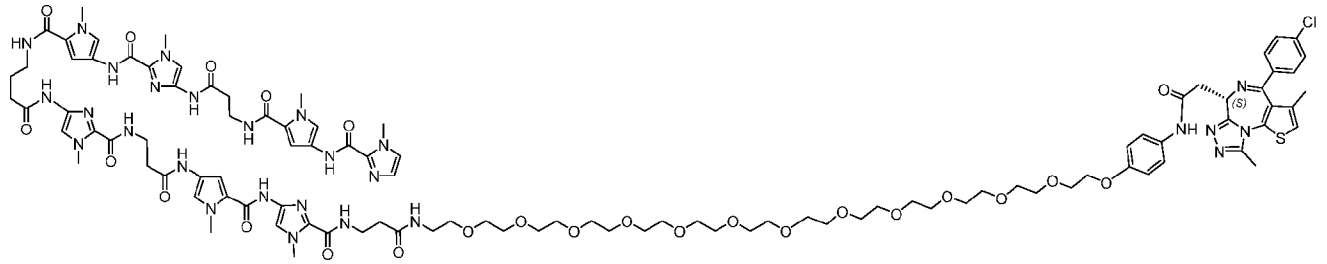
A-20



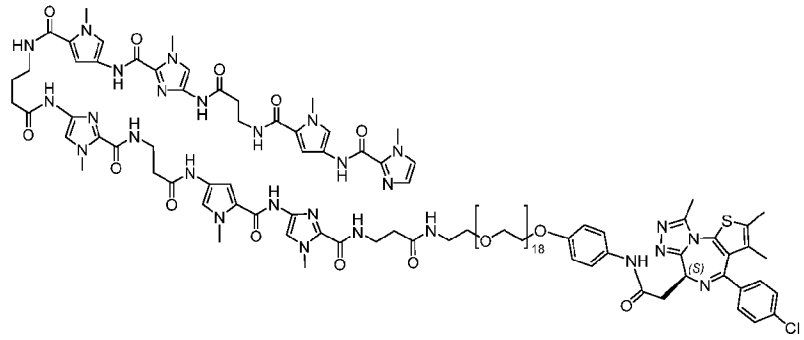
A-21



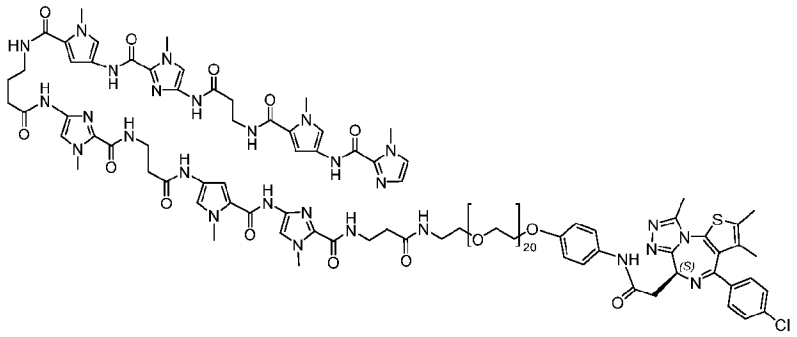
A-22



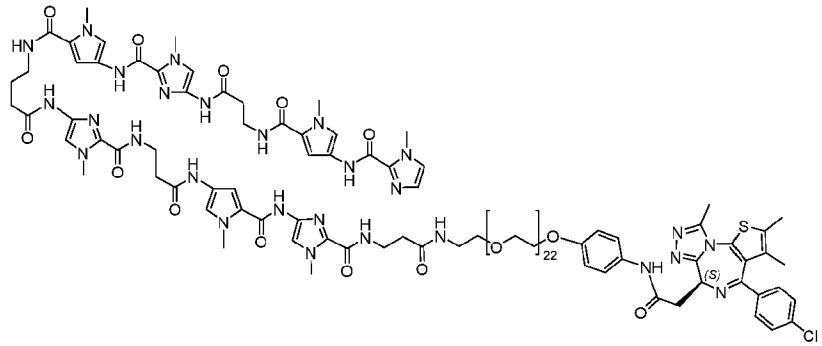
A-23



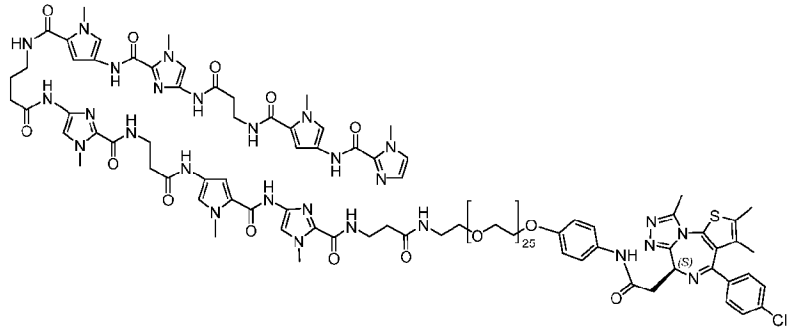
A-24



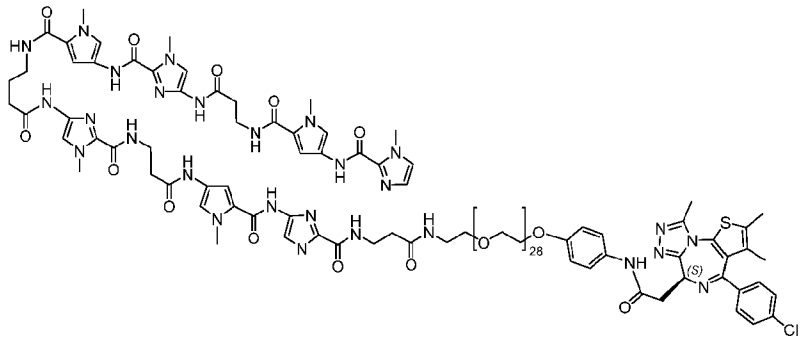
A-25



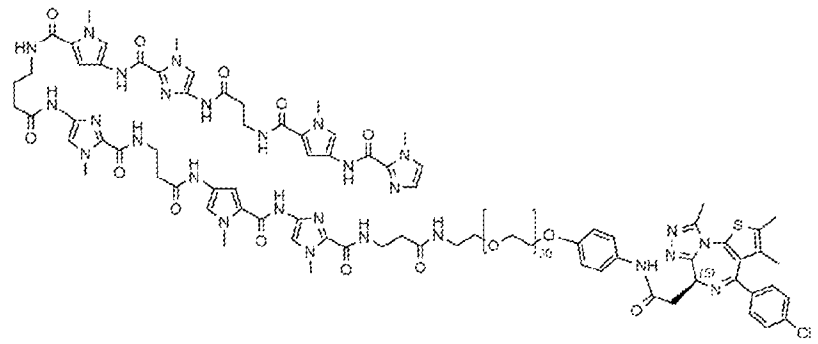
A-26



A-27



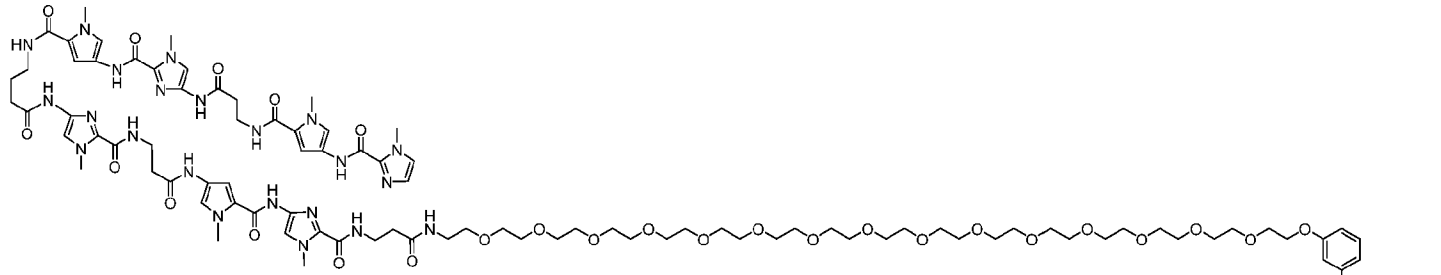
A-28



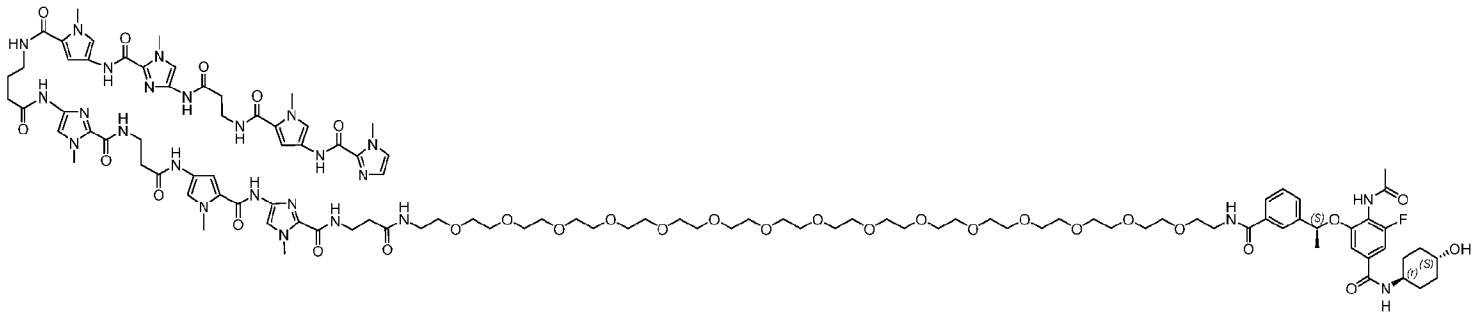
A-29

A-30

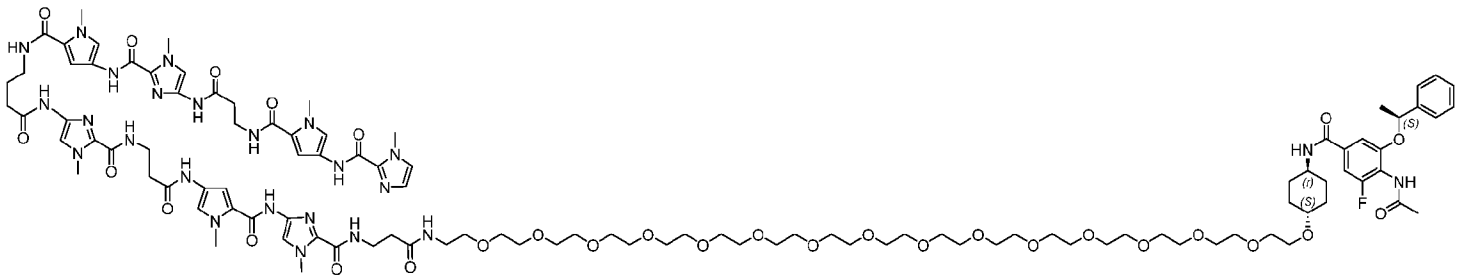
A-31

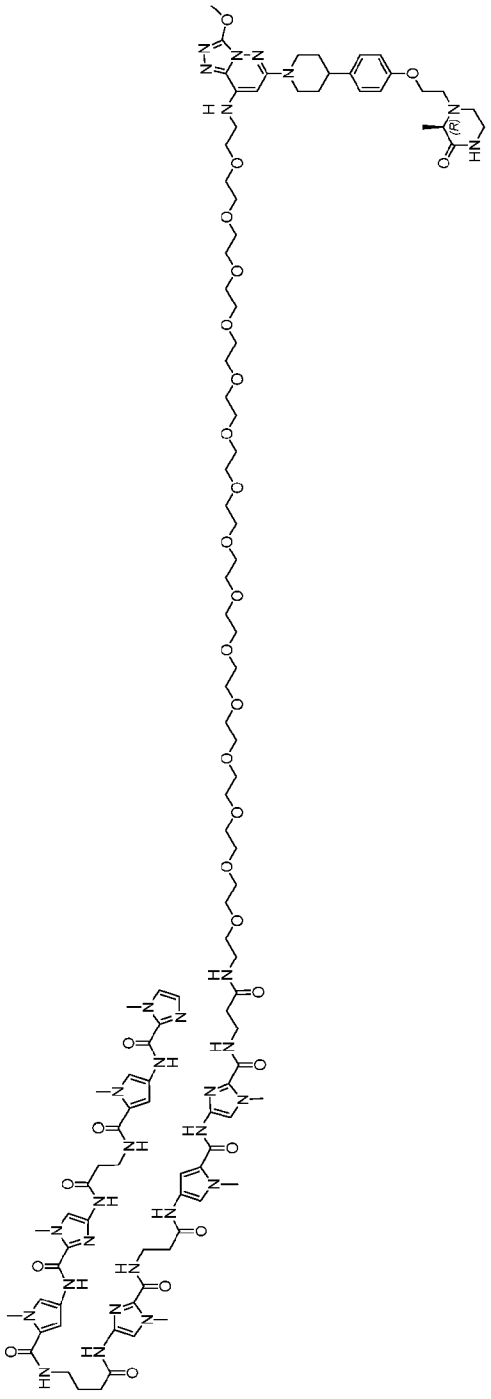


A-32

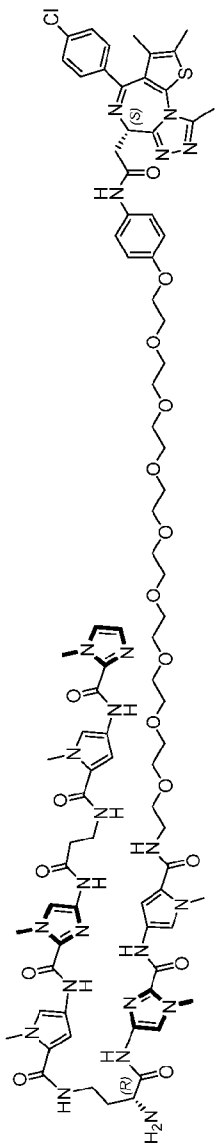


A-33

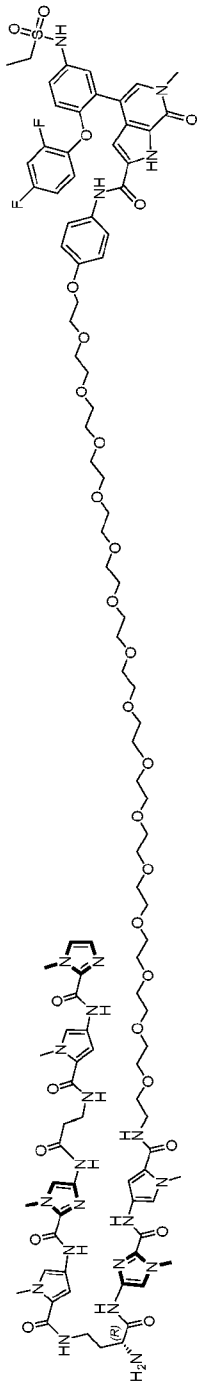




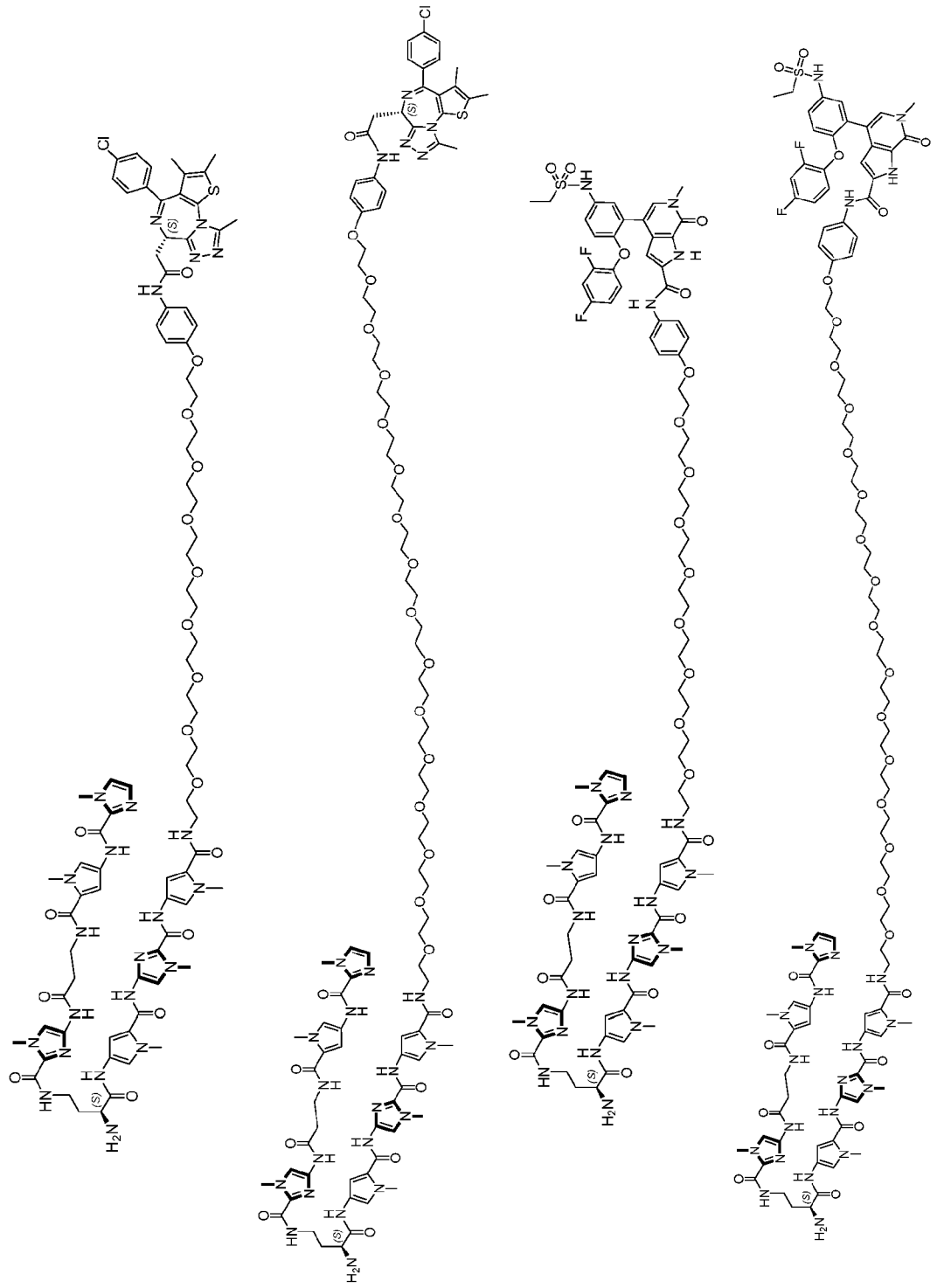
A-34



A-35

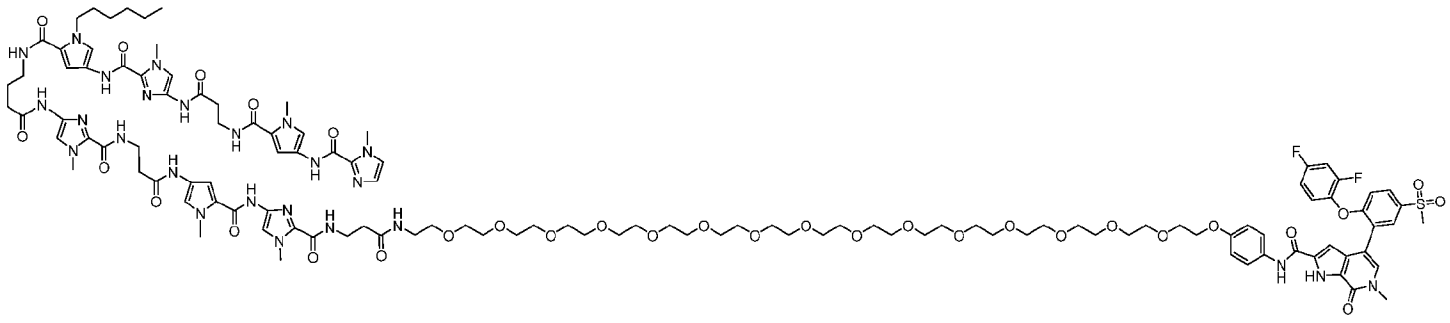


A-36

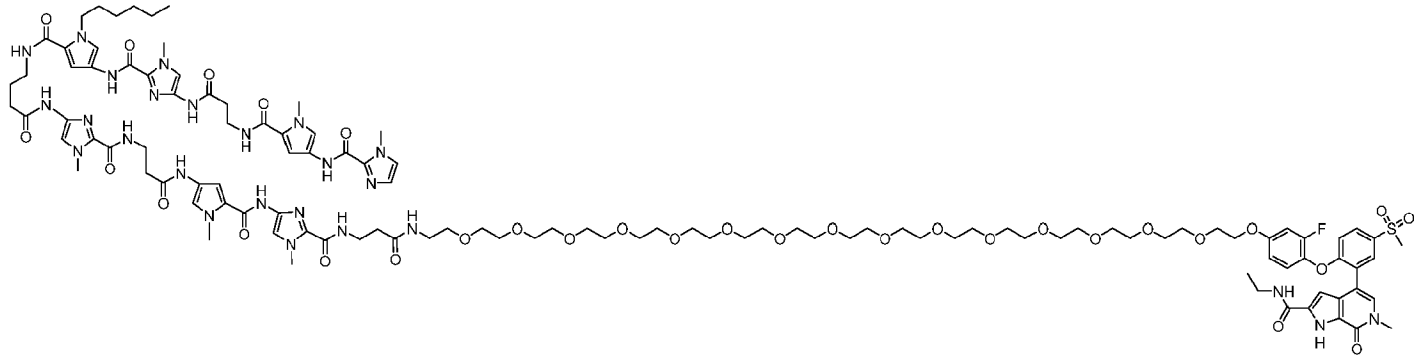


A-37 A-38 A-39 A-40

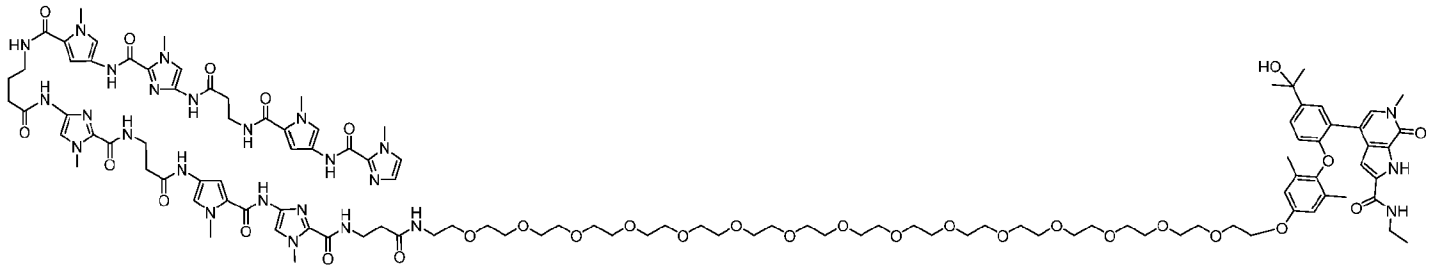
A-41



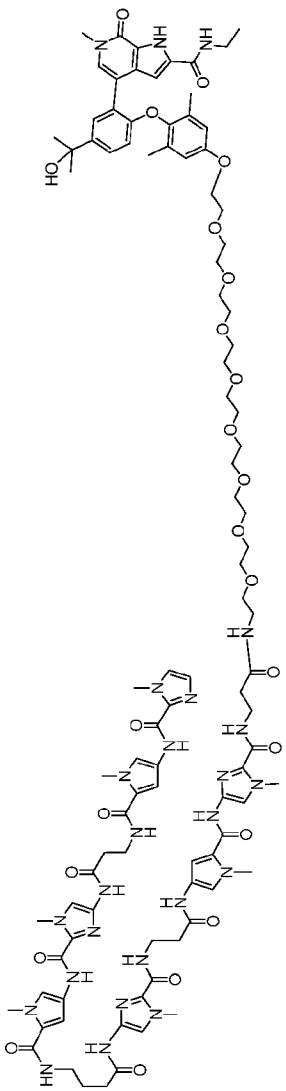
A-42



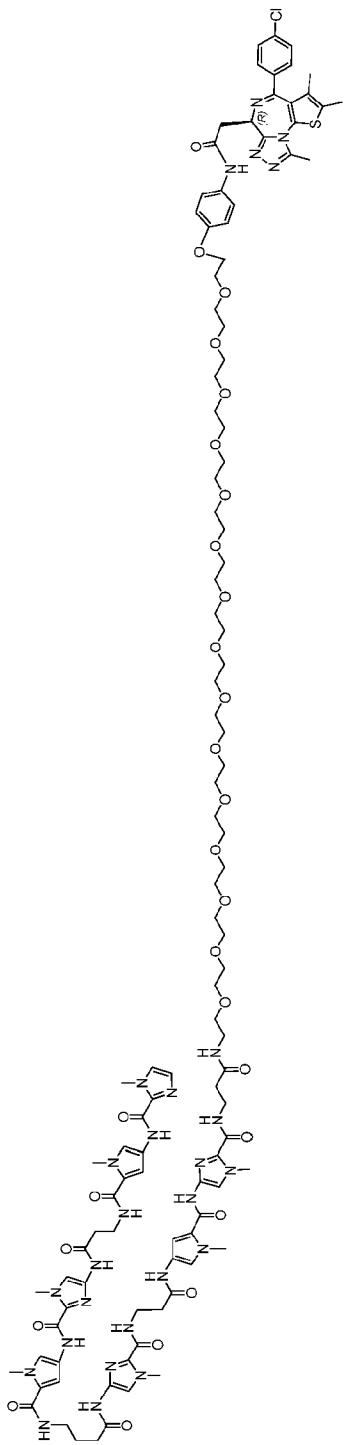
A-43



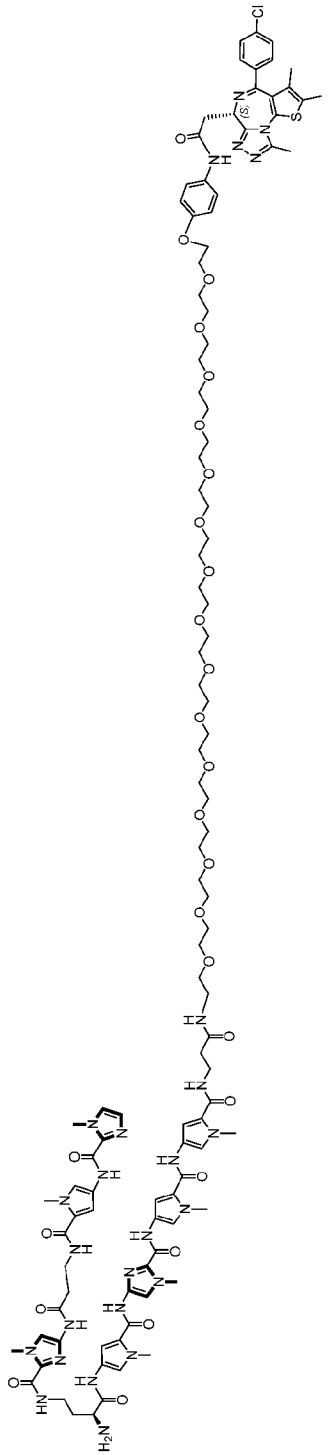
A-44



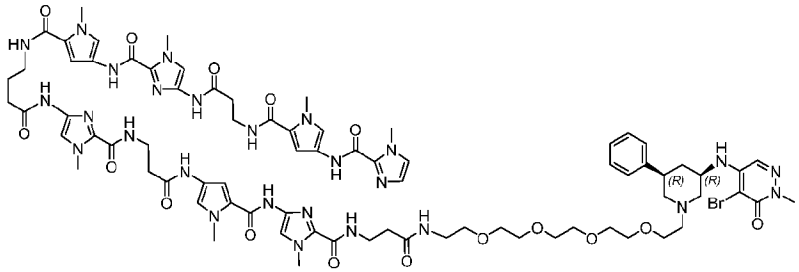
A-45



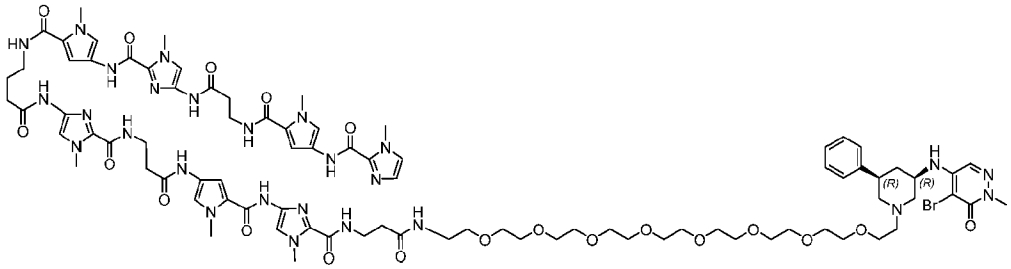
A-46



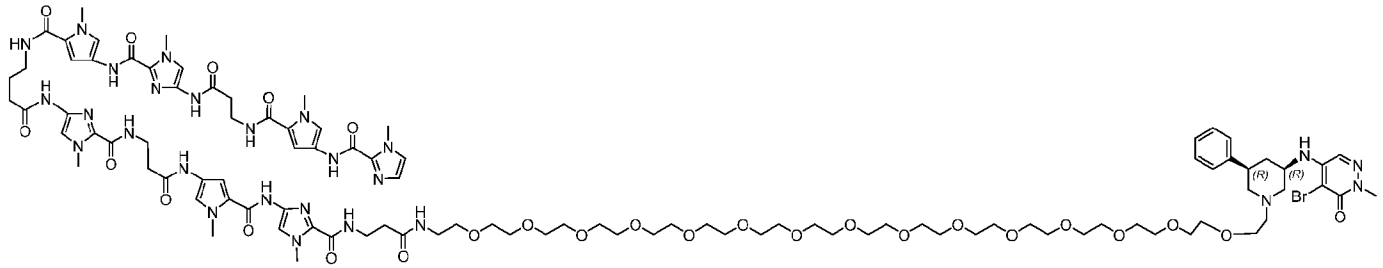
A-47



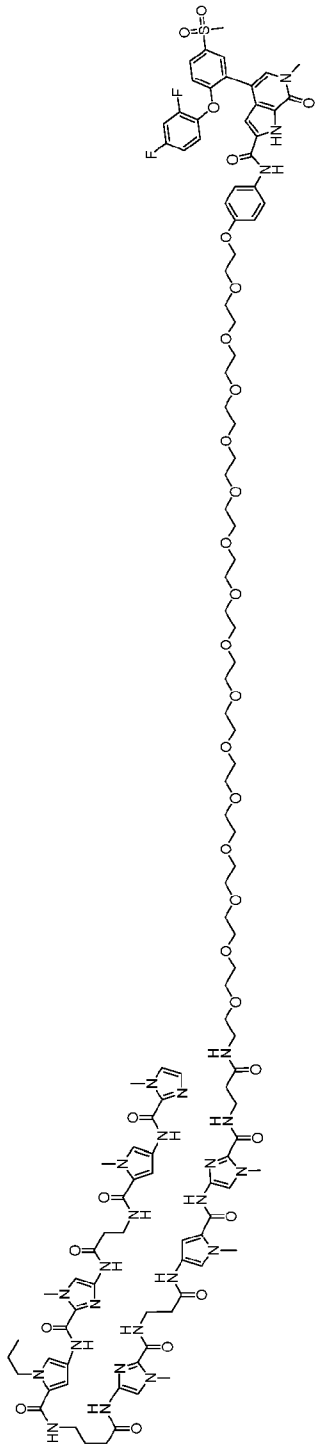
A-48



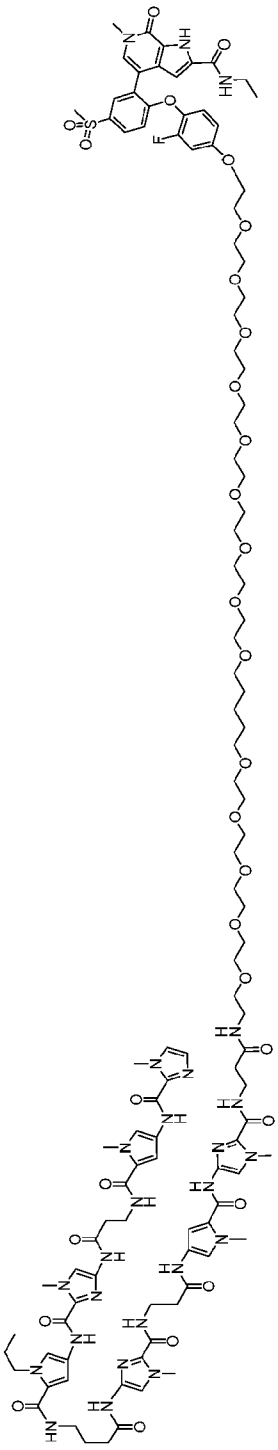
A-49



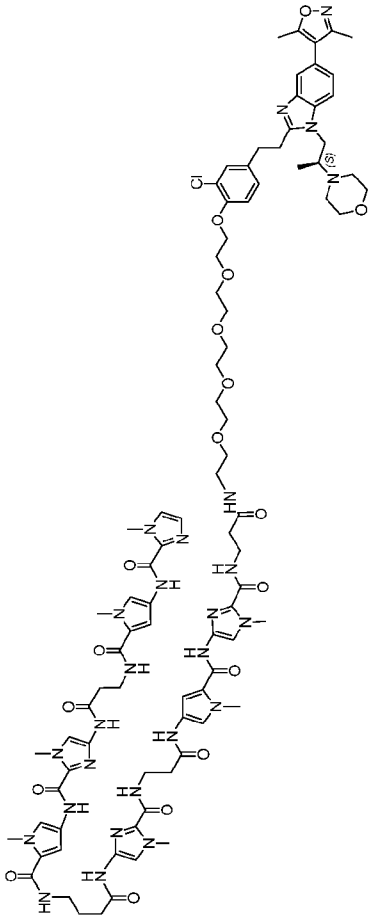
A-50



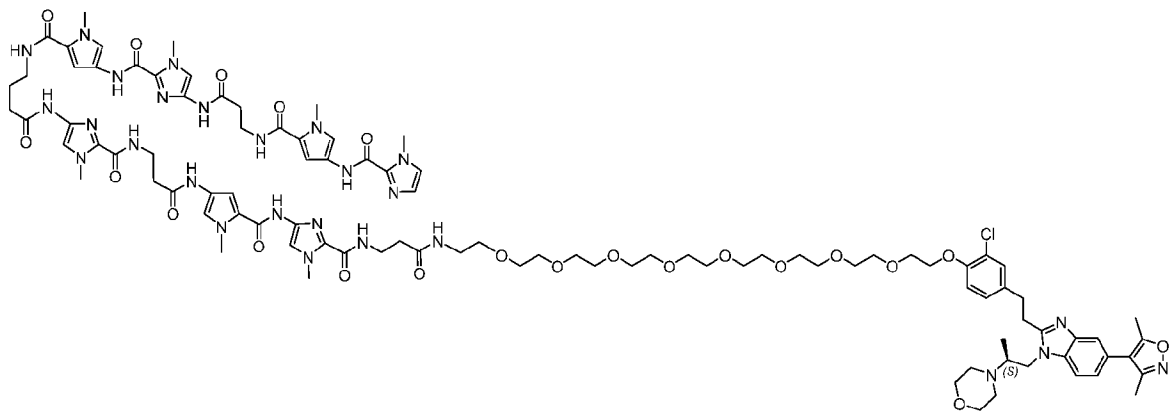
A-51



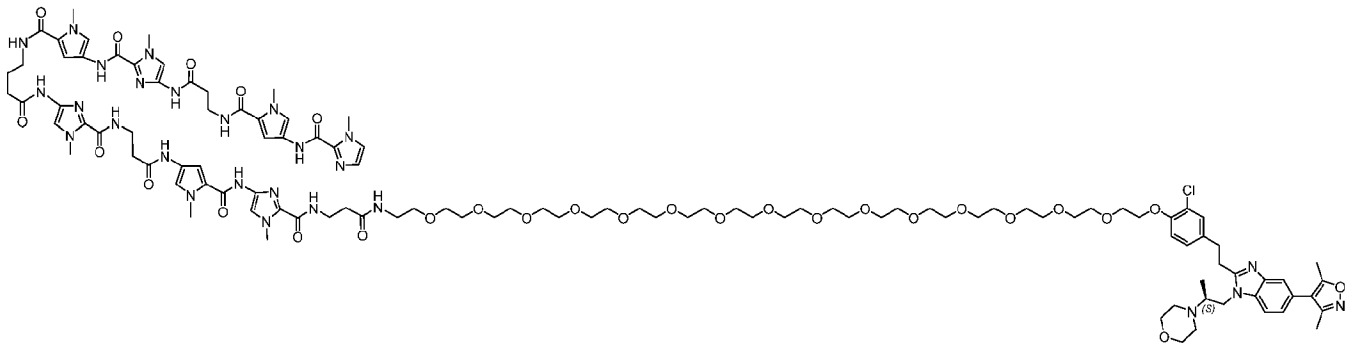
A-52



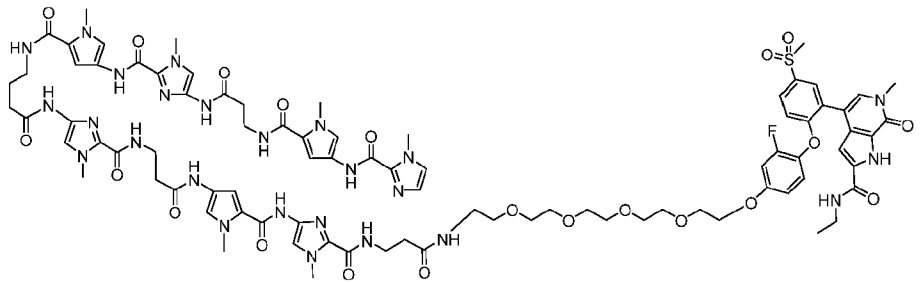
A-53



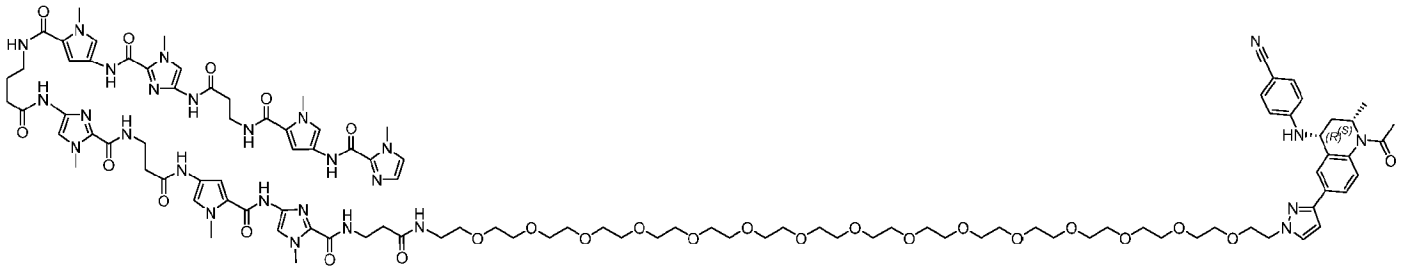
A-54



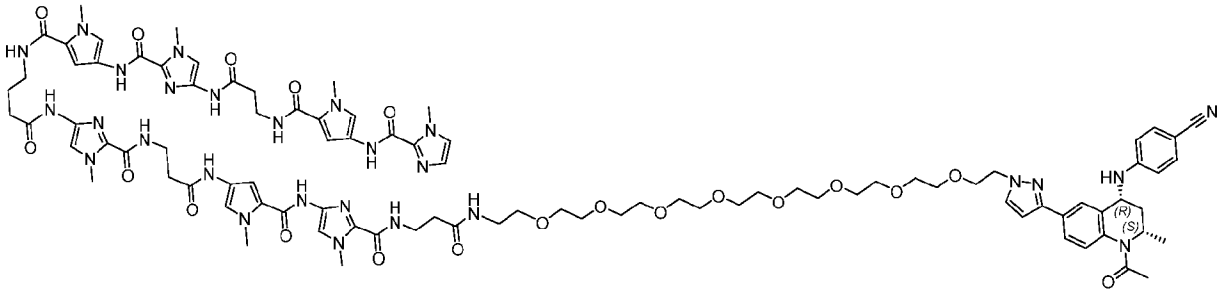
A-55



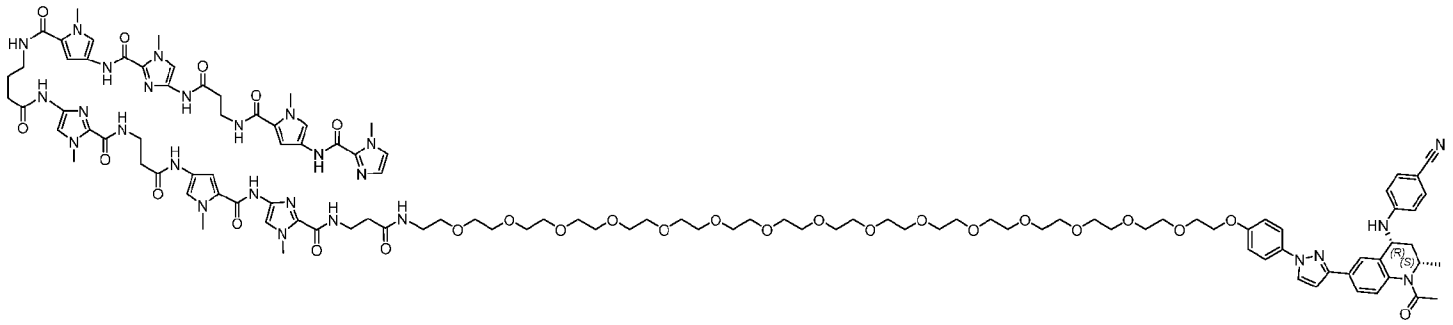
A-56



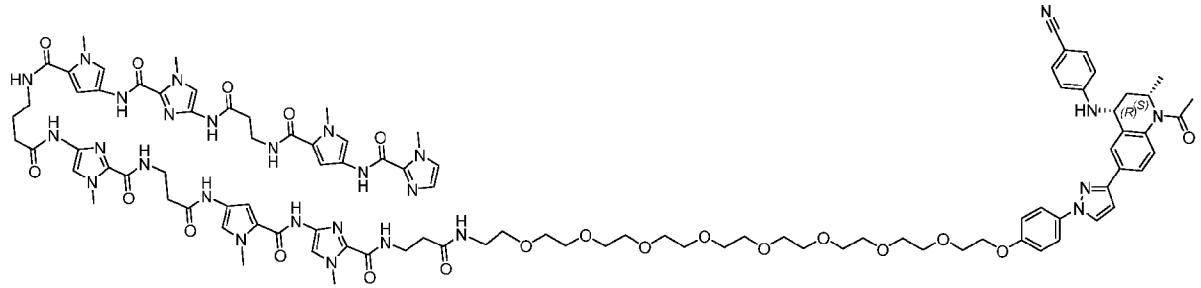
A-57



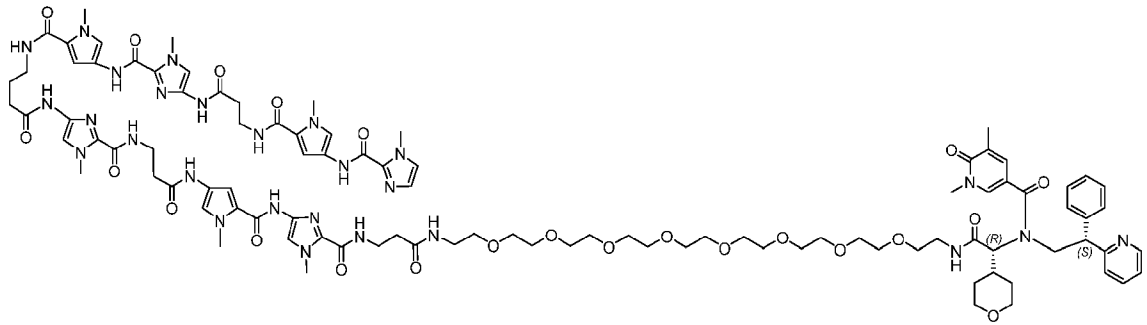
A-58



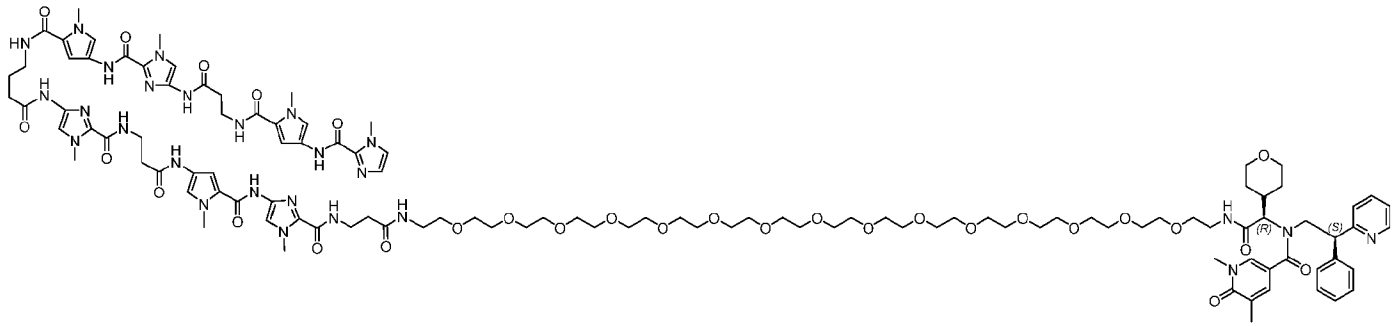
A-59



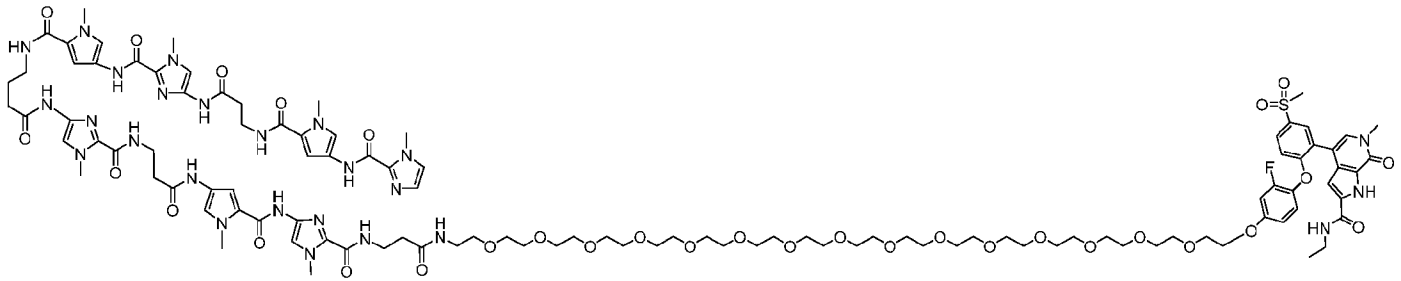
A-60



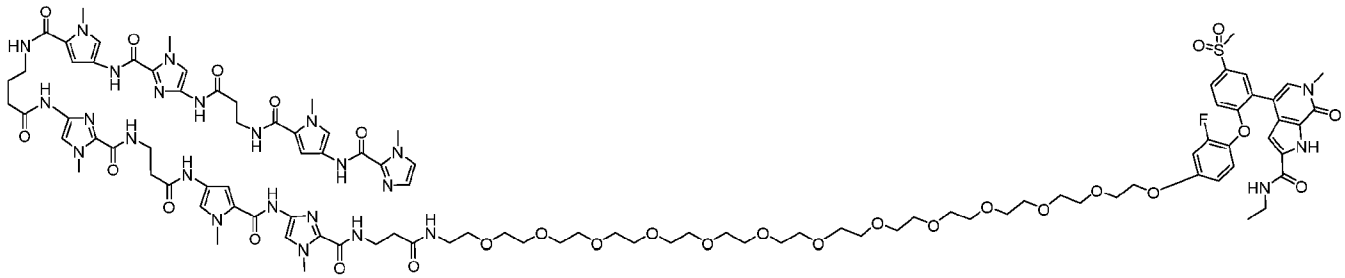
A-61



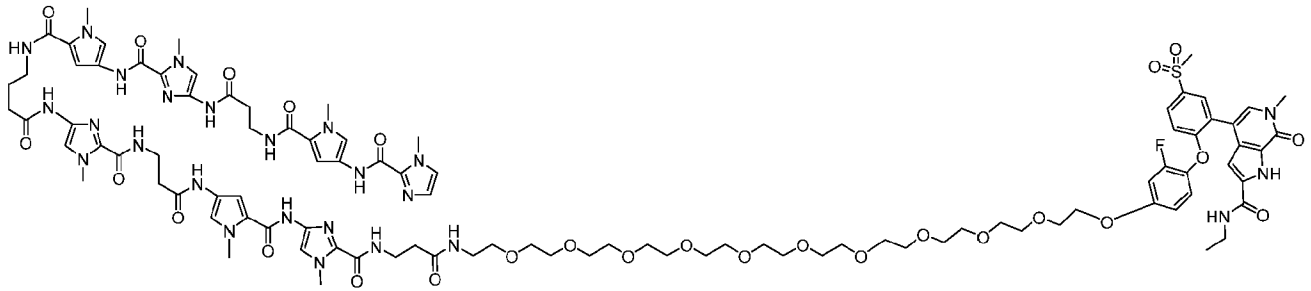
A-62



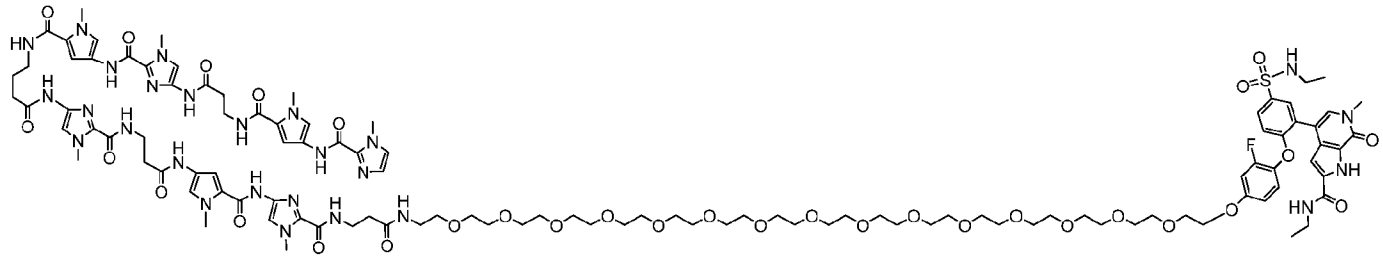
A-63



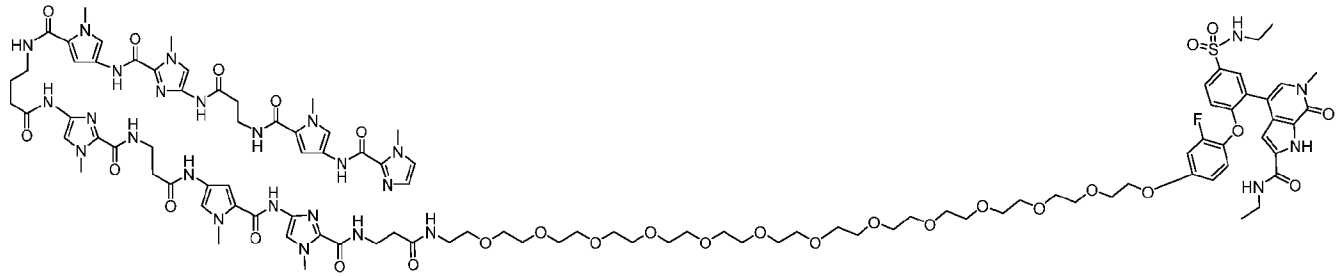
A-64



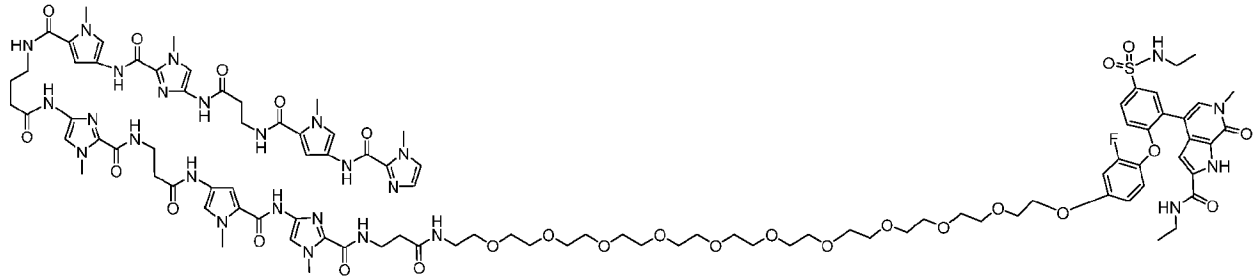
A-65



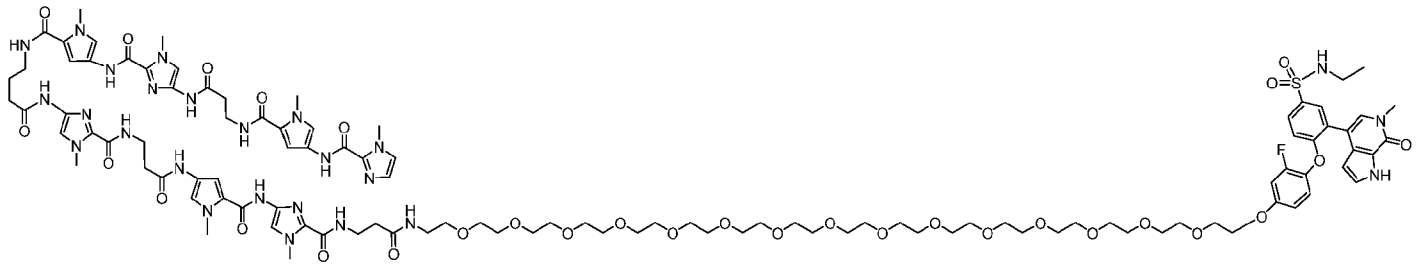
A-66



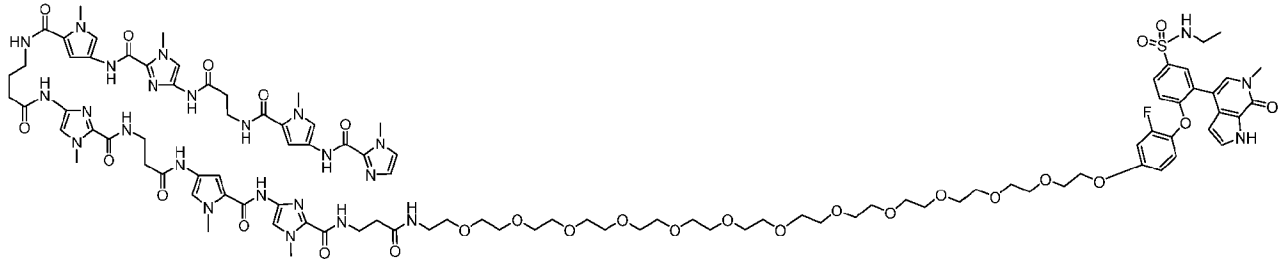
A-67



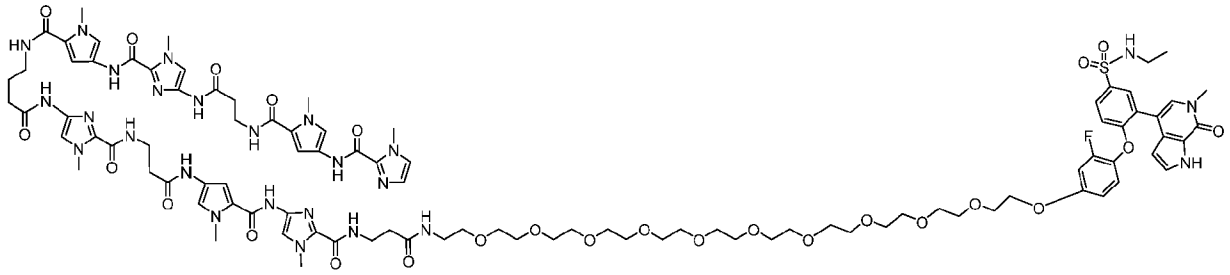
A-68



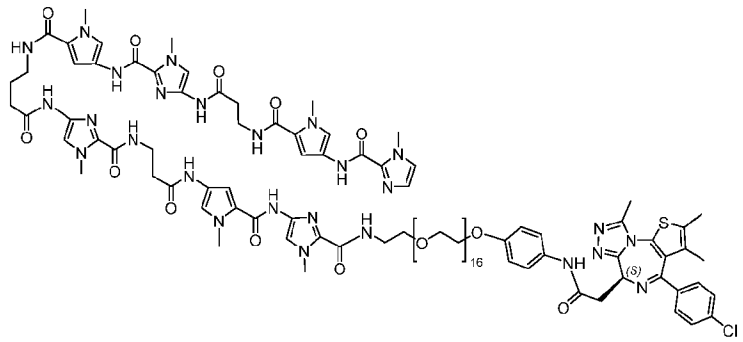
A-69



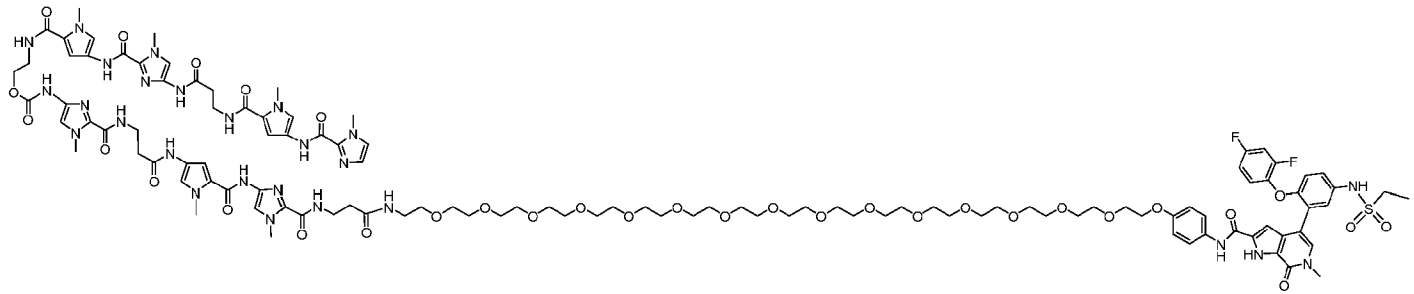
A-70



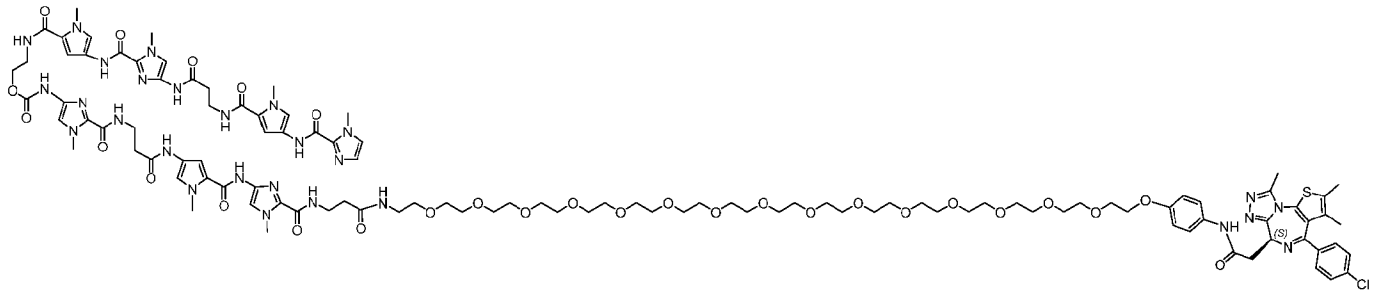
A-71



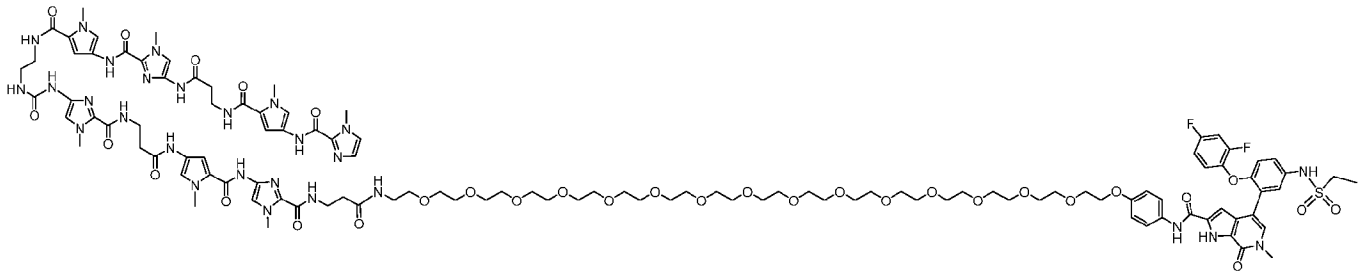
A-72



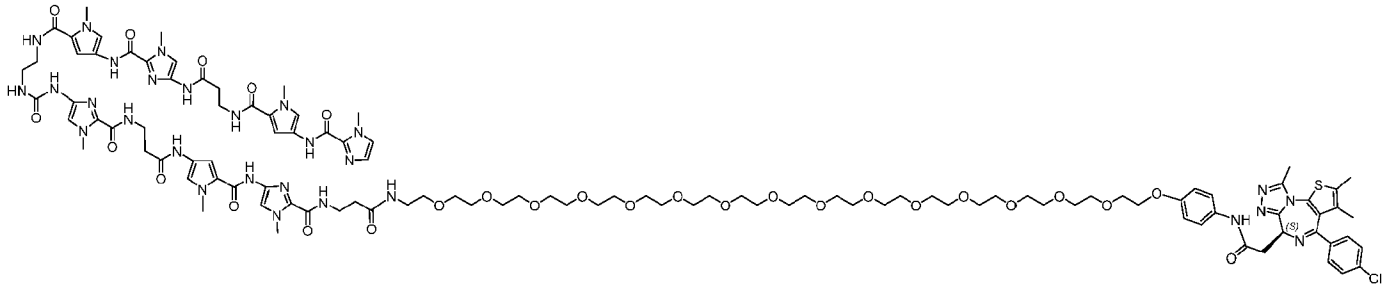
A-73



A-74



A-75



[00237] As used herein, two embodiments are "mutually exclusive" when one is defined to be something which is different than the other. For example, an embodiment wherein two groups combine to form a cycloalkyl are mutually exclusive with an embodiment in which one group is ethyl and the other group is hydrogen. Similarly, an embodiment wherein one group is CH_2 is mutually exclusive with an embodiment wherein the same group is NH.

Methods of Use

[00238] In another aspect, provided herein is a method of treating an individual having an expanded nucleotide repeat disorder, such nucleotide repeat comprising CAG, the method comprising administering a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus.

[00239] In another aspect, provided herein is a method of decreasing expression of a gene having an expanded nucleotide repeat, such as CAG, in a cell, the method comprising contacting the cell with a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus.

[00240] In some embodiments, the expanded nucleotide repeat disorder is an expanded CAG repeat disorder.

[00241] In some embodiments, the expanded nucleotide repeat disorder is Huntington's disease (HD). In some embodiments, the expanded nucleotide repeat disorder is a Huntington's disease-like syndrome. In some embodiments, the expanded nucleotide repeat disorder is Juvenile Huntington's disease.

[00242] In some embodiments, the expanded nucleotide repeat has at least about 36 repeats, at least about 40 repeats, at least about 50 repeats, at least about 60 repeats, at least about 70 repeats, at least about 80 repeats, at least about 90 repeats, at least about 100 repeats, at least about 110 repeats, at least about 120 repeats, or more.

[00243] In some embodiments, the expanded nucleotide repeat comprises CAG.

[00244] In some embodiments, the method results in reduction of expression of the gene having the expanded nucleotide repeat. In some embodiments, the reduction of expression is at least about 10%, at least

about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, or more, compared to an untreated individual.

[00245] In some embodiments, the gene is huntingtin (HTT).

[00246] In another aspect, provide herein is a method of treating Huntington's disease (HD) in a patient in need thereof, the method comprising administering to the patient a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

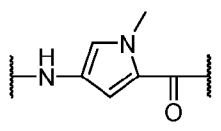
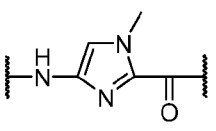
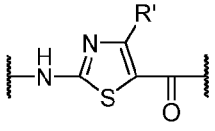
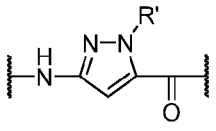
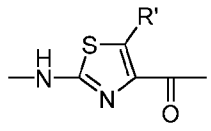
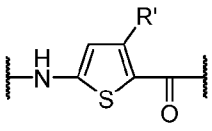
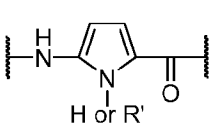
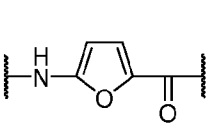
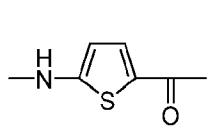
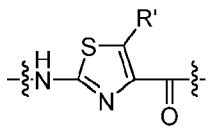
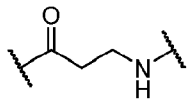
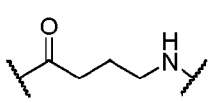
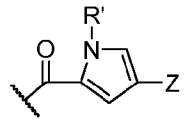
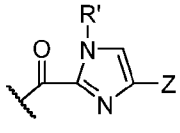
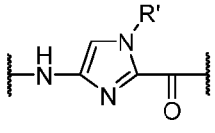
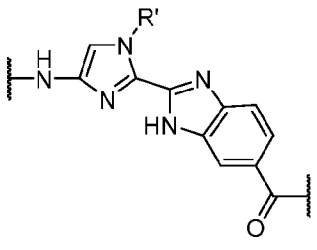
(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat; and

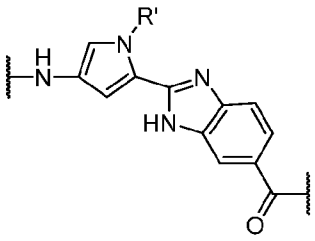
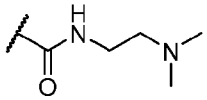
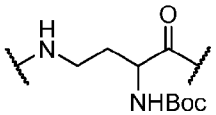
(c) the oligomeric backbone links the first terminus and the second terminus; and

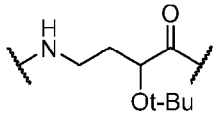
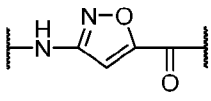
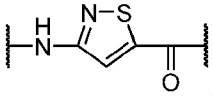
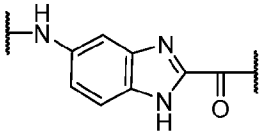
wherein the DNA-binding moiety comprises –NH-Q-C(O)-, wherein Q is an optionally substituted C_{6-10} arylene, optionally substituted 4 to 10-membered heterocyclene, optionally substituted 5 to 10-membered heteroarylene group, or an optionally substituted alkylene group.

[00247] In some embodiments, the DNA-binding moiety comprise a polyamide.

[00248] In some embodiments, the DNA-binding moiety comprises a polyamide of one or more of the following subunits selected from


(Py),

(Im),

(Th),

(Pz),

(Nt),

(Tn),

(Nh),

(Fr),

(Tp),

(iNt),

(beta or β),

(gAB),

(PyT),

(ImT),

(iIm),


(ImBi),  (PyBi),  (Dp), ,

, , , -NH-benzopyrazinylene-C(O)-, -NH-phenylene-C(O)-, -NH-pyridinylene-(C)O-, -NH-piperidinylene-C(O)-, -NH-pyrimidinylene-C(O)-, -NH-anthracenylene-C(O)-, -NH-quinolinylene-C(O)-, and , wherein each R' is independently hydrogen, optionally substituted C_1-C_{20} alkyl, C_1-C_{20} heteroalkyl, C_1-C_{20} haloalkyl, or C_1-C_{20} alkylamino; and Z is H, NH_2, C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 alkyl-NH_2.

[00249] In some embodiments, the DNA-binding moiety comprises a structure of any one of Formulas (A-1) – (A-13).

[00250] In some embodiments, the method reduces one or more symptoms of Huntington's disease.

[00251] In some embodiments, the one or more symptoms are selected from chorea, cognitive decline, abnormal libido, abnormal eye movement, abnormal sense of smell, aggression, agitation, anxiety, apathy, bradykinesia, bradyphrenia, clumsiness, delusions, depression, difficulty walking, disinhibition, dystonia, gait imbalance, muscle weakness, hallucinations, hostility, hypokinesia, irritability, memory impairment, myoclonus, obsessive-compulsive behavior, poor fine motor coordination, seizure, speech articulation difficulties, staring gaze, weight loss, abnormal cholesterol metabolism, abnormal cerebral white matter, alcoholism, Babinski sign, caudate atrophy, cerebral atrophy, choking, clonus, degeneration of the striatum, excessive daytime sleepiness, impaired visuospatial constructive cognition, inability to walk, insomnia, mutism, oral-pharyngeal dysphagia, rigidity, suicidal ideation, cerebellar atrophy, dementia, gate ataxia, gliosis, hyperreflexia, neuronal loss, or personality changes.

Pharmaceutical Compositions and Administration

[00252] Provided herein, in some embodiments, are compositions comprising a therapeutically effective amount of a transcription modulator molecule described herein (also referred to herein as "a pharmaceutical agent").

[00253] Pharmaceutical compositions may be formulated using one or more physiologically acceptable carriers including excipients and auxiliaries which facilitate processing of the pharmaceutical agent into preparations which are used pharmaceutically. Proper formulation is dependent upon the route of administration chosen. A summary of pharmaceutical compositions is found, for example, in Remington: The Science and Practice of Pharmacy, Nineteenth Ed (Easton, Pa., Mack Publishing Company, 1995);

Hoover, John E., Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pennsylvania 1975; Liberman, H.A. and Lachman, L., Eds., Pharmaceutical Dosage Forms, Marcel Decker, New York, N.Y., 1980; and Pharmaceutical Dosage Forms and Drug Delivery Systems, Seventh Ed. (Lippincott Williams & Wilkins, 1999).

[00254] The compositions and methods of the present disclosure may be utilized to treat an individual in need thereof. In certain embodiments, the individual is a mammal such as a human, or a non-human mammal. When administered to an animal, such as a human, the composition or the pharmaceutical agent, is preferably administered as a pharmaceutical composition comprising, for example, a pharmaceutical agent and a pharmaceutically acceptable carrier or excipient. Pharmaceutically acceptable carriers are well known in the art and include, for example, aqueous solutions such as water or physiologically buffered saline or other solvents or vehicles such as glycols, glycerol, oils such as olive oil, or injectable organic esters. In a preferred embodiment, when such pharmaceutical compositions are for human administration, particularly for invasive routes of administration, e.g., routes, such as injection or implantation, that circumvent transport or diffusion through an epithelial barrier, the aqueous solution is pyrogen-free, or substantially pyrogen-free. The excipients can be chosen, for example, to effect delayed release of an agent or to selectively target one or more cells, tissues or organs. The pharmaceutical composition can be in dosage unit form such as tablet, capsule, granule, lyophile for reconstitution, powder, solution, syrup, suppository, injection or the like. The composition can also be present in a transdermal delivery system, e.g., a skin patch. The composition can also be present in a solution suitable for topical administration, such as an eye drop.

[00255] A pharmaceutically acceptable excipient can contain physiologically acceptable agents that act, for example, to stabilize, increase solubility or to increase the absorption of a compound such as a pharmaceutical agent. Such physiologically acceptable agents include, for example, carbohydrates, such as glucose, sucrose or dextrans, antioxidants, such as ascorbic acid or glutathione, chelating agents, low molecular weight proteins or other stabilizers or excipients. The choice of a pharmaceutically acceptable excipient, including a physiologically acceptable agent, depends, for example, on the route of administration of the composition. The preparation or pharmaceutical composition can be a self-emulsifying drug delivery system or a self microemulsifying drug delivery system. The pharmaceutical composition (preparation) also can be a liposome or other polymer matrix, which can have incorporated therein, for example, a compound of the invention. Liposomes, for example, which comprise phospholipids or other lipids, are nontoxic, physiologically acceptable and metabolizable carriers that are relatively simple to make and administer.

[00256] A pharmaceutical composition (preparation) can be administered to a subject by any of a number of routes of administration including, for example, orally, for example, drenches as in aqueous or non-aqueous solutions or suspensions, tablets, capsules, including sprinkle capsules and gelatin capsules, boluses, powders, granules, pastes for application to the tongue; absorption through the oral mucosa, e.g., sublingually; anally, rectally or vaginally, for example, as a pessary, cream or foam; parenterally, including intramuscularly, intravenously, subcutaneously or intrathecally as, for example, a sterile solution or suspension; nasally; intraperitoneally; subcutaneously; transdermally, for example, as a patch applied to the

skin; and topically, for example, as a cream, ointment or spray applied to the skin, or as an eye drop. The compound may also be formulated for inhalation. In certain embodiments, a compound may be simply dissolved or suspended in sterile water.

[00257] A pharmaceutical composition may be a sterile aqueous or non-aqueous solution, suspension or emulsion, *e.g.*, a microemulsion. The excipients described herein are examples and are in no way limiting. An effective amount or therapeutically effective amount refers to an amount of the one or more pharmaceutical agents administered to a subject, either as a single dose or as part of a series of doses, which is effective to produce a desired therapeutic effect.

[00258] Subjects may generally be monitored for therapeutic effectiveness using assays and methods suitable for the condition being treated, which assays will be familiar to those having ordinary skill in the art and are described herein. Pharmacokinetics of a pharmaceutical agent, or one or more metabolites thereof, that is administered to a subject may be monitored by determining the level of the pharmaceutical agent or metabolite in a biological fluid, for example, in the blood, blood fraction, *e.g.*, serum, and/or in the urine, and/or other biological sample or biological tissue from the subject. Any method practiced in the art and described herein to detect the agent may be used to measure the level of the pharmaceutical agent or metabolite during a treatment course.

[00259] The dose of a pharmaceutical agent described herein for treating a disease or disorder may depend upon the subject's condition, that is, stage of the disease, severity of symptoms caused by the disease, general health status, as well as age, gender, and weight, and other factors apparent to a person skilled in the medical art. Pharmaceutical compositions may be administered in a manner appropriate to the disease to be treated as determined by persons skilled in the medical arts. In addition to the factors described herein and above related to use of pharmaceutical agent for treating a disease or disorder, suitable duration and frequency of administration of the pharmaceutical agent may also be determined or adjusted by such factors as the condition of the patient, the type and severity of the patient's disease, the particular form of the active ingredient, and the method of administration. Optimal doses of an agent may generally be determined using experimental models and/or clinical trials. The optimal dose may depend upon the body mass, weight, or blood volume of the subject. The use of the minimum dose that is sufficient to provide effective therapy is usually preferred. Design and execution of pre-clinical and clinical studies for a pharmaceutical agent, including when administered for prophylactic benefit, described herein are well within the skill of a person skilled in the relevant art. When two or more pharmaceutical agents are administered to treat a disease or disorder, the optimal dose of each pharmaceutical agent may be different, such as less than when either agent is administered alone as a single agent therapy. In certain particular embodiments, two pharmaceutical agents in combination may act synergistically or additively, and either agent may be used in a lesser amount than if administered alone. An amount of a pharmaceutical agent that may be administered per day may be, for example, between about 0.01 mg/kg and 100 mg/kg, *e.g.*, between about 0.1 to 1 mg/kg, between about 1 to 10 mg/kg, between about 10-50 mg/kg, between about 50-100 mg/kg body weight. In other embodiments, the amount of a pharmaceutical agent that may be administered per day is between about 0.01 mg/kg and

1000 mg/kg, between about 100-500 mg/kg, or between about 500-1000 mg/kg body weight. The optimal dose, per day or per course of treatment, may be different for the disease or disorder to be treated and may also vary with the administrative route and therapeutic regimen.

[00260] Pharmaceutical compositions comprising a pharmaceutical agent can be formulated in a manner appropriate for the delivery method by using techniques routinely practiced in the art. The composition may be in the form of a solid, *e.g.,* tablet, capsule, semi-solid, *e.g.,* gel, liquid, or gas, e.g., aerosol. In other embodiments, the pharmaceutical composition is administered as a bolus infusion.

[00261] Pharmaceutical acceptable excipients are well known in the pharmaceutical art and described, for example, in Rowe et al., Handbook of Pharmaceutical Excipients: A Comprehensive Guide to Uses, Properties, and Safety, 5th Ed., 2006, and in *Remington: The Science and Practice of Pharmacy* (Gennaro, 21st Ed. Mack Pub. Co., Easton, PA (2005)). Exemplary pharmaceutically acceptable excipients include sterile saline and phosphate buffered saline at physiological pH. Preservatives, stabilizers, dyes, buffers, and the like may be provided in the pharmaceutical composition. In addition, antioxidants and suspending agents may also be used. In general, the type of excipient is selected based on the mode of administration, as well as the chemical composition of the active ingredient(s). Alternatively, compositions described herein may be formulated as a lyophilizate. A composition described herein may be lyophilized or otherwise formulated as a lyophilized product using one or more appropriate excipient solutions for solubilizing and/or diluting the pharmaceutical agent(s) of the composition upon administration. In other embodiments, the pharmaceutical agent may be encapsulated within liposomes using technology known and practiced in the art. In certain particular embodiments, a pharmaceutical agent is not formulated within liposomes for application to a stent that is used for treating highly, though not totally, occluded arteries. Pharmaceutical compositions may be formulated for any appropriate manner of administration described herein and in the art.

[00262] A pharmaceutical composition, *e.g.,* for oral administration or for injection, infusion, subcutaneous delivery, intramuscular delivery, intraperitoneal delivery or other method, may be in the form of a liquid. A liquid pharmaceutical composition may include, for example, one or more of the following: a sterile diluent such as water, saline solution, preferably physiological saline, Ringer's solution, isotonic sodium chloride, fixed oils that may serve as the solvent or suspending medium, polyethylene glycols, glycerin, propylene glycol or other solvents; antibacterial agents; antioxidants; chelating agents; buffers and agents for the adjustment of tonicity such as sodium chloride or dextrose. A parenteral composition can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic. The use of physiological saline is preferred, and an injectable pharmaceutical composition is preferably sterile. In another embodiment, for treatment of an ophthalmological condition or disease, a liquid pharmaceutical composition may be applied to the eye in the form of eye drops. A liquid pharmaceutical composition may be delivered orally.

[00263] For oral formulations, at least one of the pharmaceutical agents described herein can be used alone or in combination with appropriate additives to make tablets, powders, granules or capsules, and if desired, with diluents, buffering agents, moistening agents, preservatives, coloring agents, and flavoring agents. The

pharmaceutical agents may be formulated with a buffering agent to provide for protection of the compound from low pH of the gastric environment and/or an enteric coating. A pharmaceutical agent included in a pharmaceutical composition may be formulated for oral delivery with a flavoring agent, *e.g.*, in a liquid, solid or semi-solid formulation and/or with an enteric coating.

[00264] A pharmaceutical composition comprising any one of the pharmaceutical agents described herein may be formulated for sustained or slow release, also called timed release or controlled release. Such compositions may generally be prepared using well known technology and administered by, for example, oral, rectal, intradermal, or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain the compound dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane. Excipients for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. The amount of pharmaceutical agent contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release, and the nature of the condition, disease or disorder to be treated or prevented.

[00265] In certain embodiments, the pharmaceutical compositions comprising a pharmaceutical agent are formulated for transdermal, intradermal, or topical administration. The compositions can be administered using a syringe, bandage, transdermal patch, insert, or syringe-like applicator, as a powder/talc or other solid, liquid, spray, aerosol, ointment, foam, cream, gel, paste. This preferably is in the form of a controlled release formulation or sustained release formulation administered topically or injected directly into the skin adjacent to or within the area to be treated, e.g., intradermally or subcutaneously. The active compositions can also be delivered via iontophoresis. Preservatives can be used to prevent the growth of fungi and other microorganisms. Suitable preservatives include, but are not limited to, benzoic acid, butylparaben, ethyl paraben, methyl paraben, propylparaben, sodium benzoate, sodium propionate, benzalkonium chloride, benzethonium chloride, benzyl alcohol, cetypyridinium chloride, chlorobutanol, phenol, phenylethyl alcohol, thimerosal, and combinations thereof.

[00266] Pharmaceutical compositions comprising a pharmaceutical agent can be formulated as emulsions for topical application. An emulsion contains one liquid distributed in the body of a second liquid. The emulsion may be an oil-in-water emulsion or a water-in-oil emulsion. Either or both of the oil phase and the aqueous phase may contain one or more surfactants, emulsifiers, emulsion stabilizers, buffers, and other excipients. The oil phase may contain other oily pharmaceutically approved excipients. Suitable surfactants include, but are not limited to, anionic surfactants, non-ionic surfactants, cationic surfactants, and amphoteric surfactants. Compositions for topical application may also include at least one suitable suspending agent, antioxidant, chelating agent, emollient, or humectant.

[00267] Ointments and creams may, for example, be formulated with an aqueous or oily base with the addition of suitable thickening and/or gelling agents. Lotions may be formulated with an aqueous or oily base and will in general also contain one or more emulsifying agents, stabilizing agents, dispersing agents, suspending agents, thickening agents, or coloring agents. Liquid sprays may be delivered from pressurized

packs, for example, via a specially shaped closure. Oil-in-water emulsions can also be used in the compositions, patches, bandages and articles. These systems are semisolid emulsions, micro-emulsions, or foam emulsion systems.

Definitions

[00268] As used herein, the terms below have the meanings indicated.

[00269] It is to be understood that certain radical naming conventions can include either a mono-radical or a di-radical, depending on the context. For example, where a substituent requires two points of attachment to the rest of the molecule, it is understood that the substituent is a di-radical. For example, a substituent identified as alkyl that requires two points of attachment includes di-radicals such as $-CH_2-$, $-CH_2CH_2-$, $-CH_2CH(CH_3)CH_2-$, and the like. Other radical naming conventions clearly indicate that the radical is a di-radical such as "alkylene," "alkenylene," "arylene", "heteroarylene."

[00270] When ranges of values are disclosed, and the notation "from n_1 … to n_2" or "between n_1 … and n_2" is used, where n_1 and n_2 are the numbers, then unless otherwise specified, this notation is intended to include the numbers themselves and the range between them. This range may be integral or continuous between and including the end values. By way of example, the range "from 2 to 6 carbons" is intended to include two, three, four, five, and six carbons, since carbons come in integer units. Compare, by way of example, the range "from 1 to 3 μM (micromolar)," which is intended to include 1 μM, 3 μM, and everything in between to any number of significant figures (e.g., 1.255 μM, 2.1 μM, 2.9999 μM, etc.).

[00271] The term "about," as used herein, is intended to qualify the numerical values which it modifies, denoting such a value as variable within a margin of error. When no particular margin of error, such as a standard deviation to a mean value given in a chart or table of data, is recited, the term "about" should be understood to mean that range which would encompass the recited value and the range which would be included by rounding up or down to that figure as well, taking into account significant figures.

[00272] The term "polyamide" refers to polymers of linkable units chemically bound by amide (i.e., CONH) linkages; optionally, polyamides include chemical probes conjugated therewith. Polyamides may be synthesized by stepwise condensation of carboxylic acids (COOH) with amines (RR'NH) using methods known in the art. Alternatively, polyamides may be formed using enzymatic reactions in vitro, or by employing fermentation with microorganisms.

[00273] The term "linkable unit" refers to methylimidazoles, methylpyrroles, and straight and branched chain aliphatic functionalities (e.g., methylene, ethylene, propylene, butylene, and the like) which optionally contain nitrogen substituents, and chemical derivatives thereof. The aliphatic functionalities of linkable units can be provided, for example, by condensation of B-alanine or dimethylaminopropylamine during synthesis of the polyamide by methods well known in the art.

[00274] The term "linker" or "oligomeric backbone" refers to a chain of at least 10 contiguous atoms. In certain embodiments, the linker contains no more than 20 non-hydrogen atoms. The terms linker and oligomeric backbone can be used interchangeably. In some embodiments, the linker contains no more than 40 non-hydrogen atoms. In some embodiments, the linker contains no more than 60 non-hydrogen atoms. In

certain embodiments, the linker contains atoms chosen from C, H, N, O, and S. In some embodiments, every non-hydrogen atom is chemically bonded either to 2 neighboring atoms in the linker, or one neighboring atom in the linker and a terminus of the linker. In some embodiments, the linker forms an amide bond with at least one of the two other groups to which it is attached. In certain embodiments, the linker forms an ester or ether bond with at least one of the two other groups to which it is attached. In some embodiments, the linker forms a thioester or thioether bond with at least one of the two other groups to which it is attached. In some embodiments, the linker forms a direct carbon-carbon bond with at least one of the two other groups to which it is attached. In some embodiments, the linker forms an amine or amide bond with at least one of the two other groups to which it is attached. In some embodiments, the linker comprises –(CH_2OCH_2)- units. In some embodiments, the linker comprises –($CH(CH_3)OCH_2$)- units. In some embodiments, the linker comprises -($CH_2NR_NCH_2$) units, for $R_N = C_{1-4}$alkyl. In some embodiments, the linker comprises an arylene, cycloalkylene, or heterocycloalkylene moiety.

[00275] The term "spacer" refers to a chain of at least 5 contiguous atoms. In some embodiments, the spacer contains no more than 10 non-hydrogen atoms. In some embodiments, the spacer contains atoms chosen from C, H, N, O, and S. In some embodiments, the spacer forms amide bonds with the two other groups to which it is attached. In certain embodiments, the spacer comprises –(CH_2OCH_2)- units. In some embodiments, the spacer comprises -($CH_2NR_NCH_2$)- units, for $R_N = C_{1-4}$alkyl. In some embodiments, the spacer contains at least one positive charge at physiological pH.

[00276] The term "turn component" refers to a chain of about 4 to 10 contiguous atoms. In some embodiments, the turn component contains atoms chosen from C, H, N, O, and S. In some embodiments, the turn component forms amide bonds with the two other groups to which it is attached. In some embodiments, the turn component contains at least one positive charge at physiological pH.

[00277] The terms "nucleic acid and "nucleotide" refer to ribonucleotide and deoxyribonucleotide, and analogs thereof, well known in the art.

[00278] The term "oligonucleotide sequence" refers to a plurality of nucleic acids having a defined sequence and length (e.g., 2, 3, 4, 5, 6, or even more nucleotides). The term "oligonucleotide repeat sequence" refers to a contiguous expansion of oligonucleotide sequences.

[00279] The term "transcription," well known in the art, refers to the synthesis of RNA (i.e., ribonucleic acid) by DNA-directed RNA polymerase. The term "modulate transcription" refers to a change in transcriptional level which can be measured by methods well known in the art, for example, assay of mRNA, the product of transcription. In certain embodiments, modulation is an increase in transcription. In other embodiments, modulation is a decrease in transcription

[00280] The term "acyl," as used herein, alone or in combination, refers to a carbonyl attached to an alkenyl, alkyl, aryl, cycloalkyl, heteroaryl, heterocycle, or any other moiety were the atom attached to the carbonyl is carbon. An "acetyl" group refers to a –$C(O)CH_3$ group. An "alkylcarbonyl" or "alkanoyl" group refers to an alkyl group attached to the parent molecular moiety through a carbonyl group. Examples of such

groups include methylcarbonyl and ethylcarbonyl. Examples of acyl groups include formyl, alkanoyl and aroyl.

[00281] The term "alkenyl," as used herein, alone or in combination, refers to a straight-chain or branched-chain hydrocarbon radical having one or more double bonds and containing from 2 to 20 carbon atoms. In certain embodiments, said alkenyl will comprise from 2 to 6 carbon atoms. The term "alkenylene" refers to a carbon-carbon double bond system attached at two or more positions such as ethenylene [(-CH=CH-),(-C::C-)]. Examples of suitable alkenyl radicals include ethenyl, propenyl, 2-methylpropenyl, 1,4-butadienyl and the like. Unless otherwise specified, the term "alkenyl" may include "alkenylene" groups.

[00282] The term "alkoxy," as used herein, alone or in combination, refers to an alkyl ether radical, wherein the term alkyl is as defined below. Examples of suitable alkyl ether radicals include methoxy, ethoxy, n-propoxy, isopropoxy, n-butoxy, iso-butoxy, sec-butoxy, tert-butoxy, and the like.

[00283] The term "alkyl," as used herein, alone or in combination, refers to a straight-chain or branched-chain alkyl radical containing from 1 to 20 carbon atoms. In certain embodiments, said alkyl will comprise from 1 to 10 carbon atoms. In further embodiments, said alkyl will comprise from 1 to 8 carbon atoms. Alkyl groups may be optionally substituted as defined herein. Examples of alkyl radicals include methyl, ethyl, n-propyl, isopropyl, n-butyl, isobutyl, sec-butyl, tert-butyl, pentyl, iso-amyl, hexyl, octyl, noyl and the like. The term "alkylene," as used herein, alone or in combination, refers to a saturated aliphatic group derived from a straight or branched chain saturated hydrocarbon attached at two or more positions, such as methylene ($-CH_2-$). Unless otherwise specified, the term "alkyl" may include "alkylene" groups.

[00284] The term "alkylamino," as used herein, alone or in combination, refers to an alkyl group attached to the parent molecular moiety through an amino group. Suitable alkylamino groups may be mono- or dialkylated, forming groups such as, for example, N-methylamino, N-ethylamino, N,N-dimethylamino, N,N-ethylmethylamino and the like.

[00285] The term "alkylidene," as used herein, alone or in combination, refers to an alkenyl group in which one carbon atom of the carbon-carbon double bond belongs to the moiety to which the alkenyl group is attached.

[00286] The term "alkylthio," as used herein, alone or in combination, refers to an alkyl thioether (R–S–) radical wherein the term alkyl is as defined above and wherein the sulfur may be singly or doubly oxidized. Examples of suitable alkyl thioether radicals include methylthio, ethylthio, n-propylthio, isopropylthio, n-butylthio, iso-butylthio, sec-butylthio, tert-butylthio, methanesulfonyl, ethanesulfinyl, and the like.

[00287] The term "alkynyl," as used herein, alone or in combination, refers to a straight-chain or branched chain hydrocarbon radical having one or more triple bonds and containing from 2 to 20 carbon atoms. In certain embodiments, said alkynyl comprises from 2 to 6 carbon atoms. In further embodiments, said alkynyl comprises from 2 to 4 carbon atoms. The term "alkynylene" refers to a carbon-carbon triple bond attached at two positions such as ethynylene (-C:::C-, -C≡C-). Examples of alkynyl radicals include ethynyl, propynyl, hydroxypropynyl, butyn-1-yl, butyn-2-yl, pentyn-1-yl, 3-methylbutyn-1-yl, hexyn-2-yl, and the like. Unless otherwise specified, the term "alkynyl" may include "alkynylene" groups.

[00288] The terms “amido” and “carbamoyl,” as used herein, alone or in combination, refer to an amino group as described below attached to the parent molecular moiety through a carbonyl group, or vice versa. The term “C-amido” as used herein, alone or in combination, refers to a -C(O)N(RR’) group with R and R’ as defined herein or as defined by the specifically enumerated “R” groups designated. The term “N-amido” as used herein, alone or in combination, refers to a RC(O)N(R’)- group, with R and R’ as defined herein or as defined by the specifically enumerated “R” groups designated. The term "acylamino" as used herein, alone or in combination, embraces an acyl group attached to the parent moiety through an amino group. An example of an "acylamino" group is acetylamino ($CH_3C(O)NH$-).

[00289] The term “amide,” as used herein, alone in combination, refers to -C(O)NRR’, wherein R and R’ are independently chosen from hydrogen, alkyl, acyl, heteroalkyl, aryl, cycloalkyl, heteroaryl, and heterocycloalkyl, any of which may themselves be optionally substituted. Additionally, R and R’ may combine to form heterocycloalkyl, either of which may be optionally substituted. Amides may be formed by direct condensation of carboxylic acids with amines, or by using acid chlorides. In addition, coupling reagents are known in the art, including carbodiimide-based compounds such as DCC and EDCI.

[00290] The term “amino,” as used herein, alone or in combination, refers to -NRR’, wherein R and R’ are independently chosen from hydrogen, alkyl, acyl, heteroalkyl, aryl, cycloalkyl, heteroaryl, and heterocycloalkyl, any of which may themselves be optionally substituted. Additionally, R and R’ may combine to form heterocycloalkyl, either of which may be optionally substituted.

[00291] The term "aryl," as used herein, alone or in combination, means a carbocyclic aromatic system containing one, two or three rings wherein such polycyclic ring systems are fused together. The term "aryl" embraces aromatic groups such as phenyl, naphthyl, anthracenyl, and phenanthryl. The term "arylene" embraces aromatic groups such as phenylene, naphthylene, anthracenylene, and phenanthrylene.

[00292] The term “arylalkenyl” or “aralkenyl,” as used herein, alone or in combination, refers to an aryl group attached to the parent molecular moiety through an alkenyl group.

[00293] The term “arylalkoxy” or “aralkoxy,” as used herein, alone or in combination, refers to an aryl group attached to the parent molecular moiety through an alkoxy group.

[00294] The term “arylalkyl” or “aralkyl,” as used herein, alone or in combination, refers to an aryl group attached to the parent molecular moiety through an alkyl group.

[00295] The term “arylalkynyl” or “aralkynyl,” as used herein, alone or in combination, refers to an aryl group attached to the parent molecular moiety through an alkynyl group.

[00296] The term “arylalkanoyl” or “aralkanoyl” or “aroyl,” as used herein, alone or in combination, refers to an acyl radical derived from an aryl-substituted alkanecarboxylic acid such as benzoyl, napthoyl, phenylacetyl, 3-phenylpropionyl (hydrocinnamoyl), 4-phenylbutyryl, (2-naphthyl)acetyl, 4-chlorohydrocinnamoyl, and the like.

[00297] The term aryloxy as used herein, alone or in combination, refers to an aryl group attached to the parent molecular moiety through an oxy.

[00298] The terms "benzo" and "benz," as used herein, alone or in combination, refer to the divalent radical C_6H_4= derived from benzene. Examples include benzothiophene and benzimidazole.

[00299] The term "carbamate," as used herein, alone or in combination, refers to an ester of carbamic acid (-NHCOO-) which may be attached to the parent molecular moiety from either the nitrogen or acid end, and which may be optionally substituted as defined herein.

[00300] The term "O-carbamyl" as used herein, alone or in combination, refers to a -OC(O)NRR', group-with R and R' as defined herein.

[00301] The term "N-carbamyl" as used herein, alone or in combination, refers to a ROC(O)NR'- group, with R and R' as defined herein.

[00302] The term "carbonyl," as used herein, when alone includes formyl [-C(O)H] and in combination is a -C(O)- group.

[00303] The term "carboxyl" or "carboxy," as used herein, refers to -C(O)OH or the corresponding "carboxylate" anion, such as is in a carboxylic acid salt. An "O-carboxy" group refers to a RC(O)O- group, where R is as defined herein. A "C-carboxy" group refers to a -C(O)OR groups where R is as defined herein.

[00304] The term "cyano," as used herein, alone or in combination, refers to -CN.

[00305] The term "cycloalkyl," or, alternatively, "carbocycle," as used herein, alone or in combination, refers to a saturated or partially saturated monocyclic, bicyclic or tricyclic alkyl group wherein each cyclic moiety contains from 3 to 12 carbon atom ring members and which may optionally be a benzo fused ring system which is optionally substituted as defined herein. In certain embodiments, said cycloalkyl will comprise from 5 to 7 carbon atoms. Examples of such cycloalkyl groups include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, tetrahydronapthyl, indanyl, octahydronaphthyl, 2,3-dihydro-1H-indenyl, adamantyl and the like. "Bicyclic" and "tricyclic" as used herein are intended to include both fused ring systems, such as decahydronaphthalene, octahydronaphthalene as well as the multicyclic (multicentered) saturated or partially unsaturated type. The latter type of isomer is exemplified in general by, bicyclo[1,1,1]pentane, camphor, adamantane, and bicyclo[3,2,1]octane.

[00306] The term "ester," as used herein, alone or in combination, refers to a carboxy group bridging two moieties linked at carbon atoms.

[00307] The term "ether," as used herein, alone or in combination, refers to an oxy group bridging two moieties linked at carbon atoms.

[00308] The term "halo," or "halogen," as used herein, alone or in combination, refers to fluorine, chlorine, bromine, or iodine.

[00309] The term "haloalkoxy," as used herein, alone or in combination, refers to a haloalkyl group attached to the parent molecular moiety through an oxygen atom.

[00310] The term "haloalkyl," as used herein, alone or in combination, refers to an alkyl radical having the meaning as defined above wherein one or more hydrogens are replaced with a halogen. Specifically embraced are monohaloalkyl, dihaloalkyl and polyhaloalkyl radicals. A monohaloalkyl radical, for one example, may have an iodo, bromo, chloro or fluoro atom within the radical. dihalo and polyhaloalkyl

radicals may have two or more of the same halo atoms or a combination of different halo radicals. Examples of haloalkyl radicals include fluoromethyl, difluoromethyl, trifluoromethyl, chloromethyl, dichloromethyl, trichloromethyl, pentafluoroethyl, heptafluoropropyl, difluorochloromethyl, dichlorofluoromethyl, difluoroethyl, difluoropropyl, dichloroethyl and dichloropropyl. "Haloalkylene" refers to a haloalkyl group attached at two or more positions. Examples include fluoromethylene (-CFH-), difluoromethylene ($-CF_2-$), chloromethylene (-CHCl-) and the like.

[00311] The term "heteroalkyl," as used herein, alone or in combination, refers to a stable straight or branched chain, or combinations thereof, fully saturated or containing from 1 to 3 degrees of unsaturation, consisting of the stated number of carbon atoms and from one to three heteroatoms chosen from N, O, and S, and wherein the N and S atoms may optionally be oxidized and the N heteroatom may optionally be quaternized. The heteroatom(s) may be placed at any interior position of the heteroalkyl group. Up to two heteroatoms may be consecutive, such as, for example, $-CH_2-NH-OCH_3$.

[00312] The term "heteroaryl," as used herein, alone or in combination, refers to a 3 to 15 membered unsaturated heteromonocyclic ring, or a fused monocyclic, bicyclic, or tricyclic ring system in which at least one of the fused rings is aromatic, which contains at least one atom chosen from N, O, and S. In certain embodiments, said heteroaryl will comprise from 1 to 4 heteroatoms as ring members. In further embodiments, said heteroaryl will comprise from 1 to 2 heteroatoms as ring members. In certain embodiments, said heteroaryl will comprise from 5 to 7 atoms. The term also embraces fused polycyclic groups wherein heterocyclic rings are fused with aryl rings, wherein heteroaryl rings are fused with other heteroaryl rings, wherein heteroaryl rings are fused with heterocycloalkyl rings, or wherein heteroaryl rings are fused with cycloalkyl rings. Examples of heteroaryl groups include pyrrolyl, pyrrolinyl, imidazolyl, pyrazolyl, pyridyl, pyrimidinyl, pyrazinyl, pyridazinyl, triazolyl, pyranyl, furyl, thienyl, oxazolyl, isoxazolyl, oxadiazolyl, thiazolyl, thiadiazolyl, isothiazolyl, indolyl, isoindolyl, indolizinyl, benzimidazolyl, quinolyl, isoquinolyl, quinoxalinyl, quinazolinyl, indazolyl, benzotriazolyl, benzodioxolyl, benzopyranyl, benzoxazolyl, benzoxadiazolyl, benzothiazolyl, benzothiadiazolyl, benzofuryl, benzothienyl, chromonyl, coumarinyl, benzopyranyl, tetrahydroquinolinyl, tetrazolopyridazinyl, tetrahydroisoquinolinyl, thienopyridinyl, furopyridinyl, pyrrolopyridinyl and the like. Exemplary tricyclic heterocyclic groups include carbazolyl, benzidolyl, phenanthrolinyl, dibenzofuranyl, acridinyl, phenanthridinyl, xanthenyl and the like.

[00313] The terms "heterocycloalkyl" and, interchangeably, "heterocycle," as used herein, alone or in combination, each refer to a saturated, partially unsaturated, or fully unsaturated (but nonaromatic) monocyclic, bicyclic, or tricyclic heterocyclic group containing at least one heteroatom as a ring member, wherein each said heteroatom may be independently chosen from nitrogen, oxygen, and sulfur. In certain embodiments, said heterocycloalkyl will comprise from 1 to 4 heteroatoms as ring members. In further embodiments, said heterocycloalkyl will comprise from 1 to 2 heteroatoms as ring members. In certain embodiments, said heterocycloalkyl will comprise from 3 to 8 ring members in each ring. In further embodiments, said heterocycloalkyl will comprise from 3 to 7 ring members in each ring. In yet further

embodiments, said hetercycloalkyl will comprise from 5 to 6 ring members in each ring. "Heterocycloalkyl" and "heterocycle" are intended to include sulfones, sulfoxides, N-oxides of tertiary nitrogen ring members, and carbocyclic fused and benzo fused ring systems; additionally, both terms also include systems where a heterocycle ring is fused to an aryl group, as defined herein, or an additional heterocycle group. Examples of heterocycle groups include tetrhydroisoquinoline, aziridinyl, azetidinyl, 1,3-benzodioxolyl, dihydroisoindolyl, dihydroisoquinolinyl, dihydrocinnolinyl, dihydrobenzodioxinyl, dihydro[1,3]oxazolo[4,5-b]pyridinyl, benzothiazolyl, dihydroindolyl, dihy-dropyridinyl, 1,3-dioxanyl, 1,4-dioxanyl, 1,3-dioxolanyl, isoindolinyl, morpholinyl, piperazinyl, pyrrolidinyl, tetrahydropyridinyl, piperidinyl, thiomorpholinyl, and the like. The heterocycle groups may be optionally substituted unless specifically prohibited.

[00314] The term "hydrazinyl" as used herein, alone or in combination, refers to two amino groups joined by a single bond, i.e., -N-N-.

[00315] The term "hydroxy," as used herein, alone or in combination, refers to -OH.

[00316] The term "hydroxyalkyl," as used herein, alone or in combination, refers to a hydroxy group attached to the parent molecular moiety through an alkyl group.

[00317] The term "imino," as used herein, alone or in combination, refers to =N-.

[00318] The term "iminohydroxy," as used herein, alone or in combination, refers to =N(OH) and =N-O-.

[00319] The phrase "in the main chain" refers to the longest contiguous or adjacent chain of carbon atoms starting at the point of attachment of a group to the compounds or molecules of any one of the formulas disclosed herein.

[00320] The term "isocyanato" refers to a -NCO group.

[00321] The term "isothiocyanato" refers to a -NCS group.

[00322] The term "mercaptyl" as used herein, alone or in combination, refers to an RS- group, where R is as defined herein.

[00323] The term "nitro," as used herein, alone or in combination, refers to $-NO_2$.

[00324] The terms "oxy" or "oxa," as used herein, alone or in combination, refer to –O–.

[00325] The term "oxo," as used herein, alone or in combination, refers to =O.

[00326] The term "perhaloalkoxy" refers to an alkoxy group where all of the hydrogen atoms are replaced by halogen atoms.

[00327] The term "perhaloalkyl" as used herein, alone or in combination, refers to an alkyl group where all of the hydrogen atoms are replaced by halogen atoms.

[00328] The terms "sulfonate," "sulfonic acid," and "sulfonic," as used herein, alone or in combination, refer the $-SO_3H$ group and its anion as the sulfonic acid is used in salt formation.

[00329] The term "sulfanyl," as used herein, alone or in combination, refers to –S–.

[00330] The term "sulfinyl," as used herein, alone or in combination, refers to –S(O)–.

[00331] The term "sulfonyl," as used herein, alone or in combination, refers to $-S(O)_2-$.

[00332] The term "N-sulfonamido" refers to a $RS(=O)_2NR'$- group with R and R' as defined herein.

[00333] The term "S-sulfonamido" refers to a $-S(=O)_2NRR'$, group, with R and R' as defined herein.

[00334] The terms "thia" and "thio," as used herein, alone or in combination, refer to a –S– group or an ether wherein the oxygen is replaced with sulfur. The oxidized derivatives of the thio group, namely sulfinyl and sulfonyl, are included in the definition of thia and thio.

[00335] The term "thiol," as used herein, alone or in combination, refers to an –SH group.

[00336] The term "thiocarbonyl," as used herein, when alone includes thioformyl –C(S)H and in combination is a –C(S)– group.

[00337] The term "N-thiocarbamyl" refers to an ROC(S)NR'– group, with R and R' as defined herein.

[00338] The term "O-thiocarbamyl" refers to a –OC(S)NRR', group with R and R' as defined herein.

[00339] The term "thiocyanato" refers to a –CNS group.

[00340] The term "trihalomethanesulfonamido" refers to a $X_3CS(O)_2NR$– group with X is a halogen and R as defined herein.

[00341] The term "trihalomethanesulfonyl" refers to a $X_3CS(O)_2$– group where X is a halogen.

[00342] The term "trihalomethoxy" refers to a X_3CO– group where X is a halogen.

[00343] The term "trisubstituted silyl," as used herein, alone or in combination, refers to a silicone group substituted at its three free valences with groups as listed herein under the definition of substituted amino. Examples include trimethysilyl, tert-butyldimethylsilyl, triphenylsilyl and the like.

[00344] Any definition herein may be used in combination with any other definition to describe a composite structural group. By convention, the trailing element of any such definition is that which attaches to the parent moiety. For example, the composite group alkylamido would represent an alkyl group attached to the parent molecule through an amido group, and the term alkoxyalkyl would represent an alkoxy group attached to the parent molecule through an alkyl group.

[00345] When a group is defined to be "absent," what is meant is that said group is absent.

[00346] The term "optionally substituted" means the anteceding group may be substituted or unsubstituted. When substituted, the substituents of an "optionally substituted" group may include, without limitation, one or more substituents independently selected from the following groups or a particular designated set of groups, alone or in combination: alkyl, alkenyl, alkynyl, alkanoyl, heteroalkyl, heterocycloalkyl, haloalkyl, haloalkenyl, haloalkynyl, perhaloalkyl, perhaloalkoxy, cycloalkyl, phenyl, aryl, aryloxy, alkoxy, haloalkoxy, oxo, acyloxy, carbonyl, carboxyl, alkylcarbonyl, carboxyester, carboxamido, cyano, hydrogen, halogen, hydroxy, amino, lower alkylamino, arylamino, amido, nitro, thiol, alkylthio, haloalkylthio, perhaloalkylthio, arylthio, sulfonate, sulfonic acid, trisubstituted silyl, N_3, SH, SCH_3, $C(O)CH_3$, CO_2CH_3, CO_2H, pyridinyl, thiophene, furanyl, carbamate, and urea. Where structurally feasible, two substituents may be joined together to form a fused five-, six-, or seven-membered carbocyclic or heterocyclic ring consisting of zero to three heteroatoms, for example forming methylenedioxy or ethylenedioxy. An optionally substituted group may be unsubstituted (e.g., $-CH_2CH_3$), fully substituted (e.g., $-CF_2CF_3$), monosubstituted (e.g., $-CH_2CH_2F$) or substituted at a level anywhere in-between fully substituted and monosubstituted (e.g., $-CH_2CF_3$). Where

substituents are recited without qualification as to substitution, both substituted and unsubstituted forms are encompassed. Where a substituent is qualified as "substituted," the substituted form is specifically intended. Additionally, different sets of optional substituents to a particular moiety may be defined as needed; in these cases, the optional substitution will be as defined, often immediately following the phrase, "optionally substituted with".

[00347] As used herein, a substituted group is derived from the unsubstituted parent group in which there has been an exchange of one or more hydrogen atoms for another atom or group. Unless otherwise indicated, when a group is deemed to be "substituted," it is meant that the group is substituted with one or more substituents independently selected from C_1-C_6 alkyl, C_1-C_6 alkenyl, C_1-C_6 alkynyl, C_1-C_6 heteroalkyl, C_3-C_7 carbocyclyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), C_3-C_7-carbocyclyl-C_1-C_6-alkyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), 3-10 membered heterocyclyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), 3-10 membered heterocyclyl-C_1-C_6-alkyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), aryl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), aryl(C_1-C_6)alkyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), 5-10 membered heteroaryl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), 5-10 membered heteroaryl(C_1-C_6)alkyl (optionally substituted with halo, C_1-C_6 alkyl, C_1-C_6 alkoxy, C_1-C_6 haloalkyl, and C_1-C_6 haloalkoxy), halo, cyano, hydroxy, C_1-C_6 alkoxy, C_1-C_6 alkoxy(C_1-C_6)alkyl (i.e., ether), aryloxy, sulfhydryl (mercapto), halo(C_1-C_6)alkyl (e.g., $-CF_3$), halo(C_1-C_6)alkoxy (e.g., $-OCF_3$), C_1-C_6 alkylthio, arylthio, amino, amino(C_1-C_6)alkyl, nitro, O-carbamyl, N-carbamyl, O-thiocarbamyl, N-thiocarbamyl, C-amido, N-amido, S-sulfonamido, N-sulfonamido, C-carboxy, O-carboxy, acyl, cyanato, isocyanato, thiocyanato, isothiocyanato, sulfinyl, sulfonyl, and oxo (=O). Wherever a group is described as "optionally substituted" that group can be substituted with the above substituents.

[00348] The term R or the term R', appearing by itself and without a number designation, unless otherwise defined, refers to a moiety chosen from hydrogen, alkyl, cycloalkyl, heteroalkyl, aryl, heteroaryl and heterocycloalkyl, any of which may be optionally substituted. Such R and R' groups should be understood to be optionally substituted as defined herein. Whether an R group has a number designation or not, every R group, including R, R' and R^n where n = (1, 2, 3, ...n), every substituent, and every term should be understood to be independent of every other in terms of selection from a group. Should any variable, substituent, or term (e.g. aryl, heterocycle, R, etc.) occur more than one time in a formula or generic structure, its definition at each occurrence is independent of the definition at every other occurrence. Those of skill in the art will further recognize that certain groups may be attached to a parent molecule or may occupy a position in a chain of elements from either end as written. For example, an unsymmetrical group such as -C(O)N(R)- may be attached to the parent moiety at either the carbon or the nitrogen.

[00349] Asymmetric centers exist in the compounds or molecules disclosed herein. These centers are designated by the symbols "R" or "S," depending on the configuration of substituents around the chiral carbon atom. It should be understood that the disclosure encompasses all stereochemical isomeric forms, including diastereomeric, enantiomeric, and epimeric forms, as well as d-isomers and 1-isomers, and mixtures thereof. Individual stereoisomers of compounds or molecules can be prepared synthetically from commercially available starting materials which contain chiral centers or by preparation of mixtures of enantiomeric products followed by separation such as conversion to a mixture of diastereomers followed by separation or recrystallization, chromatographic techniques, direct separation of enantiomers on chiral chromatographic columns, or any other appropriate method known in the art. Starting compounds or molecules of particular stereochemistry are either commercially available or can be made and resolved by techniques known in the art. Additionally, the compounds or molecules disclosed herein may exist as geometric isomers. The present disclosure includes all cis, trans, syn, anti, entgegen (E), and zusammen (Z) isomers as well as the appropriate mixtures thereof. Additionally, compounds or molecules may exist as tautomers; all tautomeric isomers are provided by this disclosure. Additionally, the compounds or molecules disclosed herein can exist in unsolvated as well as solvated forms with pharmaceutically acceptable solvents such as water, ethanol, and the like. In general, the solvated forms are considered equivalent to the unsolvated forms.

[00350] The term "bond" refers to a covalent linkage between two atoms, or two moieties when the atoms joined by the bond are considered to be part of larger substructure. A bond may be single, double, or triple unless otherwise specified. A dashed line between two atoms in a drawing of a molecule indicates that an additional bond may be present or absent at that position.

[00351] The term "disease" as used herein is intended to be generally synonymous, and is used interchangeably with, the terms "disorder," "syndrome," and "condition" (as in medical condition), in that all reflect an abnormal condition of the human or animal body or of one of its parts that impairs normal functioning, is typically manifested by distinguishing signs and symptoms, and causes the human or animal to have a reduced duration or quality of life.

[00352] The term "combination therapy" means the administration of two or more therapeutic agents to treat a therapeutic condition or disorder described in the present disclosure. Such administration encompasses co-administration of these therapeutic agents in a substantially simultaneous manner, such as in a single capsule having a fixed ratio of active ingredients or in multiple, separate capsules for each active ingredient. In addition, such administration also encompasses use of each type of therapeutic agent in a sequential manner. In either case, the treatment regimen will provide beneficial effects of the drug combination in treating the conditions or disorders described herein.

[00353] The phrase "therapeutically effective" is intended to qualify the amount of active ingredients used in the treatment of a disease or disorder or on the effecting of a clinical endpoint.

[00354] The term "therapeutically acceptable" refers to those compounds or molecules (or salts, tautomers, zwitterionic forms, etc.) which are suitable for use in contact with the tissues of patients without undue

toxicity, irritation, and allergic response, are commensurate with a reasonable benefit/risk ratio, and are effective for their intended use.

[00355] As used herein, reference to "treatment" of a patient is intended to include prophylaxis. Treatment may also be preemptive in nature, i.e., it may include prevention of disease. Prevention of a disease may involve complete protection from disease, for example as in the case of prevention of infection with a pathogen, or may involve prevention of disease progression. For example, prevention of a disease may not mean complete foreclosure of any effect related to the diseases at any level, but instead may mean prevention of the symptoms of a disease to a clinically significant or detectable level. Prevention of diseases may also mean prevention of progression of a disease to a later stage of the disease.

[00356] The term "patient" is generally synonymous with the term "subject" and includes all mammals including humans. Examples of patients include humans, livestock such as cows, goats, sheep, pigs, and rabbits, and companion animals such as dogs, cats, rabbits, and horses. Preferably, the patient is a human.

[00357] The term "contacting" refers to bringing the compound (e.g. a transcription molecular molecule of the present disclosure) into proximity of the desired target gene. The contacting may result in the binding to or result in a conformational change of the target moiety.

[00358] The compounds or molecules disclosed herein can exist as therapeutically acceptable salts. The present disclosure includes compounds or molecules listed above in the form of salts, including acid addition salts. Suitable salts include those formed with both organic and inorganic acids. Such acid addition salts will normally be pharmaceutically acceptable. However, salts of non-pharmaceutically acceptable salts may be of utility in the preparation and purification of the compound or molecule in question. Basic addition salts may also be formed and be pharmaceutically acceptable. For a more complete discussion of the preparation and selection of salts, refer to *Pharmaceutical Salts: Properties, Selection, and Use* (Stahl, P. Heinrich. Wiley-VCHA, Zurich, Switzerland, 2002).

[00359] Basic addition salts can be prepared during the final isolation and purification of the compounds or molecules by reacting a carboxy group with a suitable base such as the hydroxide, carbonate, or bicarbonate of a metal cation or with ammonia or an organic primary, secondary, or tertiary amine. The cations of therapeutically acceptable salts include lithium, sodium, potassium, calcium, magnesium, and aluminum, as well as nontoxic quaternary amine cations such as ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, diethylamine, ethylamine, tributylamine, pyridine, *N,N*-dimethylaniline, *N*-methylpiperidine, *N*-methylmorpholine, dicyclohexylamine, procaine, dibenzylamine, *N,N*-dibenzylphenethylamine, 1-ephenamine, and *N,N'*-dibenzylethylenediamine. Other representative organic amines useful for the formation of base addition salts include ethylenediamine, ethanolamine, diethanolamine, piperidine, and piperazine.

[00360] Other carrier materials and modes of administration known in the pharmaceutical art may also be used. Pharmaceutical compositions of the disclosure may be prepared by any of the well-known techniques of pharmacy, such as effective formulation and administration procedures. Preferred unit dosage

formulations are those containing an effective dose, as herein below recited, or an appropriate fraction thereof, of the active ingredient.

[00361] It should be understood that in addition to the ingredients particularly mentioned above, the formulations described above may include other agents conventional in the art having regard to the type of formulation in question, for example those suitable for oral administration may include flavoring agents.

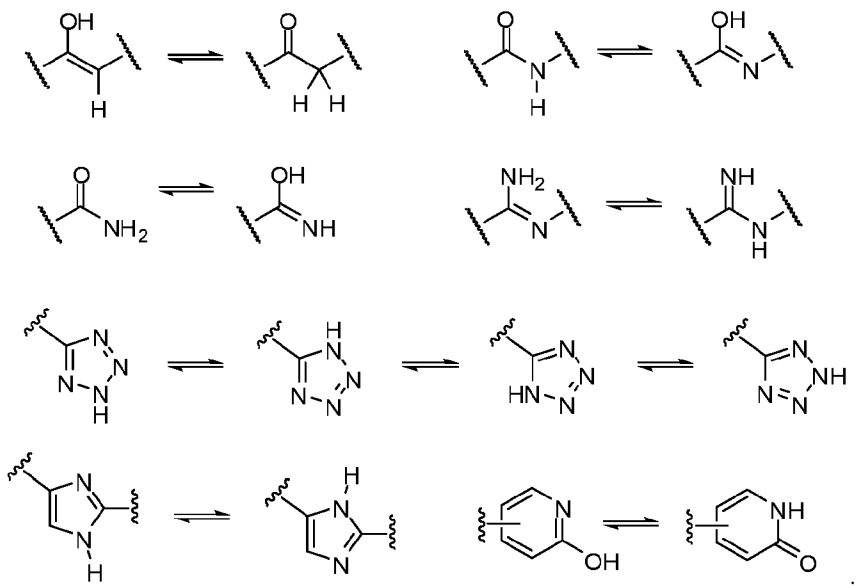
[00362] The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration.

[00363] The compounds or molecules can be administered in various modes, *e.g.* orally, topically, or by injection. The precise amount of compound administered to a patient will be the responsibility of the attendant physician. The specific dose level for any particular patient will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, sex, diets, time of administration, route of administration, rate of excretion, drug combination, the precise disorder being treated, and the severity of the indication or condition being treated. In addition, the route of administration may vary depending on the condition and its severity. The above considerations concerning effective formulations and administration procedures are well known in the art and are described in standard textbooks.

[00364] Included in the present disclosure are salts, particularly pharmaceutically acceptable salts, of the compounds described herein. The compounds of the present invention that possess a sufficiently acidic, a sufficiently basic, or both functional groups, can react with any of a number of inorganic bases, and inorganic and organic acids, to form a salt. Alternatively, compounds that are inherently charged, such as those with a quaternary nitrogen, can form a salt with an appropriate counterion, e.g., a halide such as bromide, chloride, or fluoride, particularly bromide.

[00365] Chemical entities having carbon-carbon double bonds or carbon-nitrogen double bonds may exist in *Z*- or *E*- form (or *cis*- or *trans*- form). Furthermore, some chemical entities may exist in various tautomeric forms. Unless otherwise specified, compounds described herein are intended to include all *Z*-, *E*- and tautomeric forms as well.

[00366] A "tautomer" refers to a molecule wherein a proton shift from one atom of a molecule to another atom of the same molecule is possible. The compounds presented herein, in certain embodiments, exist as tautomers. In circumstances where tautomerization is possible, a chemical equilibrium of the tautomers will exist. The exact ratio of the tautomers depends on several factors, including physical state, temperature, solvent, and pH. Some examples of tautomeric equilibrium include:



[00367] The compounds disclosed herein, in some embodiments, are used in different enriched isotopic forms, e.g., enriched in the content of ^{2}H, ^{3}H, ^{11}C, ^{13}C and/or ^{14}C. In one particular embodiment, the compound is deuterated in at least one position. Such deuterated forms can be made by the procedure described in U.S. Patent Nos. 5,846,514 and 6,334,997. As described in U.S. Patent Nos. 5,846,514 and 6,334,997, deuteration can improve the metabolic stability and or efficacy, thus increasing the duration of action of drugs.

[00368] Unless otherwise stated, compounds described herein are intended to include compounds which differ only in the presence of one or more isotopically enriched atoms. For example, compounds having the present structures except for the replacement of a hydrogen by a deuterium or tritium, or the replacement of a carbon by ^{13}C- or ^{14}C-enriched carbon are within the scope of the present disclosure.

[00369] The compounds of the present disclosure optionally contain unnatural proportions of atomic isotopes at one or more atoms that constitute such compounds. For example, the compounds may be labeled with isotopes, such as for example, deuterium (^{2}H), tritium (^{3}H), iodine-125 (^{125}I) or carbon-14 (^{14}C). Isotopic substitution with ^{2}H, ^{11}C, ^{13}C, ^{14}C, ^{15}C, ^{12}N, ^{13}N, ^{15}N, ^{16}N, ^{16}O, ^{17}O, ^{14}F, ^{15}F, ^{16}F, ^{17}F, ^{18}F, ^{33}S, ^{34}S, ^{35}S, ^{36}S, ^{35}Cl, ^{37}Cl, ^{79}Br, ^{81}Br, and ^{125}I are all contemplated. All isotopic variations of the compounds of the present invention, whether radioactive or not, are encompassed within the scope of the present invention. In some embodiments, where isotopic variations are illustrated, the remaining atoms of the compound may optionally contain unnatural portions of atomic isotopes.

[00370] In certain embodiments, the compounds disclosed herein have some or all of the ^{1}H atoms replaced with ^{2}H atoms. The methods of synthesis for deuterium-containing compounds are known in the art and include, by way of non-limiting example only, the following synthetic methods.

[00371] Deuterium substituted compounds are synthesized using various methods such as described in: Dean, Dennis C.; Editor. Recent Advances in the Synthesis and Applications of Radiolabeled Compounds for Drug Discovery and Development. [In: Curr., Pharm. Des., 2000; 6(10)] **2000**, 110 pp; George W.;

Varma, Rajender S. The Synthesis of Radiolabeled Compounds via Organometallic Intermediates, Tetrahedron, **1989**, 45(21), 6601-21; and Evans, E. Anthony. Synthesis of radiolabeled compounds, J. Radioanal. Chem., **1981**, 64(1-2), 9-32.

[00372] Deuterated starting materials are readily available and are subjected to the synthetic methods described herein to provide for the synthesis of deuterium-containing compounds. Large numbers of deuterium-containing reagents and building blocks are available commercially from chemical vendors, such as Aldrich Chemical Co.

[00373] The compounds described herein may in some cases exist as diastereomers, enantiomers, or other stereoisomeric forms. Where absolute stereochemistry is not specified, the compounds presented herein include all diastereomeric, enantiomeric, and epimeric forms as well as the appropriate mixtures thereof. Separation of stereoisomers may be performed by chromatography or by forming diastereomers and separating by recrystallization, or chromatography, or any combination thereof. (Jean Jacques, Andre Collet, Samuel H. Wilen, "Enantiomers, Racemates and Resolutions", John Wiley And Sons, Inc., 1981, herein incorporated by reference for this disclosure). Stereoisomers may also be obtained by stereoselective synthesis.

[00374] The methods and compositions described herein include the use of amorphous forms as well as crystalline forms (also known as polymorphs). The compounds described herein may be in the form of pharmaceutically acceptable salts. As well, in some embodiments, active metabolites of these compounds having the same type of activity are included in the scope of the present disclosure. In addition, the compounds described herein can exist in unsolvated as well as solvated forms with pharmaceutically acceptable solvents such as water, ethanol, and the like. The solvated forms of the compounds presented herein are also considered to be disclosed herein.

EXAMPLES

[00375] The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion. The present examples, along with the methods described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses which are encompassed within the spirit of the invention as defined by the scope of the claims will be known to those skilled in the art.

Compound Synthesis

[00376] Compounds of the present disclosure can be prepared using methods illustrated in general synthetic schemes and experimental procedures detailed below. General synthetic schemes and experimental procedures are presented for purposes of illustration and are not intended to be limiting. Starting materials used to prepare compounds of the present disclosure are commercially available or can be prepared using routine methods known in the art.

[00377] Synthetic chemistry transformations and methodologies useful in synthesizing the compounds described herein are known in the art and include, for example, those described in R. Larock, *Comprehensive Organic Transformations* (1989); T. W. Greene and P. G. M. Wuts, *Protective Groups in Organic Synthesis,* 2d. Ed. (1991); L. Fieser and M. Fieser, *Fieser and Fieser's Reagents for Organic Synthesis* (1994); and L. Paquette, ed., *Encyclopedia of Reagents for Organic Synthesis* (1995).

[00378] The following examples are intended to illustrate but not limit the disclosed embodiments. Scheme A describes the steps involved for preparing the polyamide, attaching the polyamide to the oligomeric backbone, and then attaching the ligand to the other end of the oligomeric backbone. The transcription modulator molecule such as those listed in Table 4 can be prepared using the synthesis.

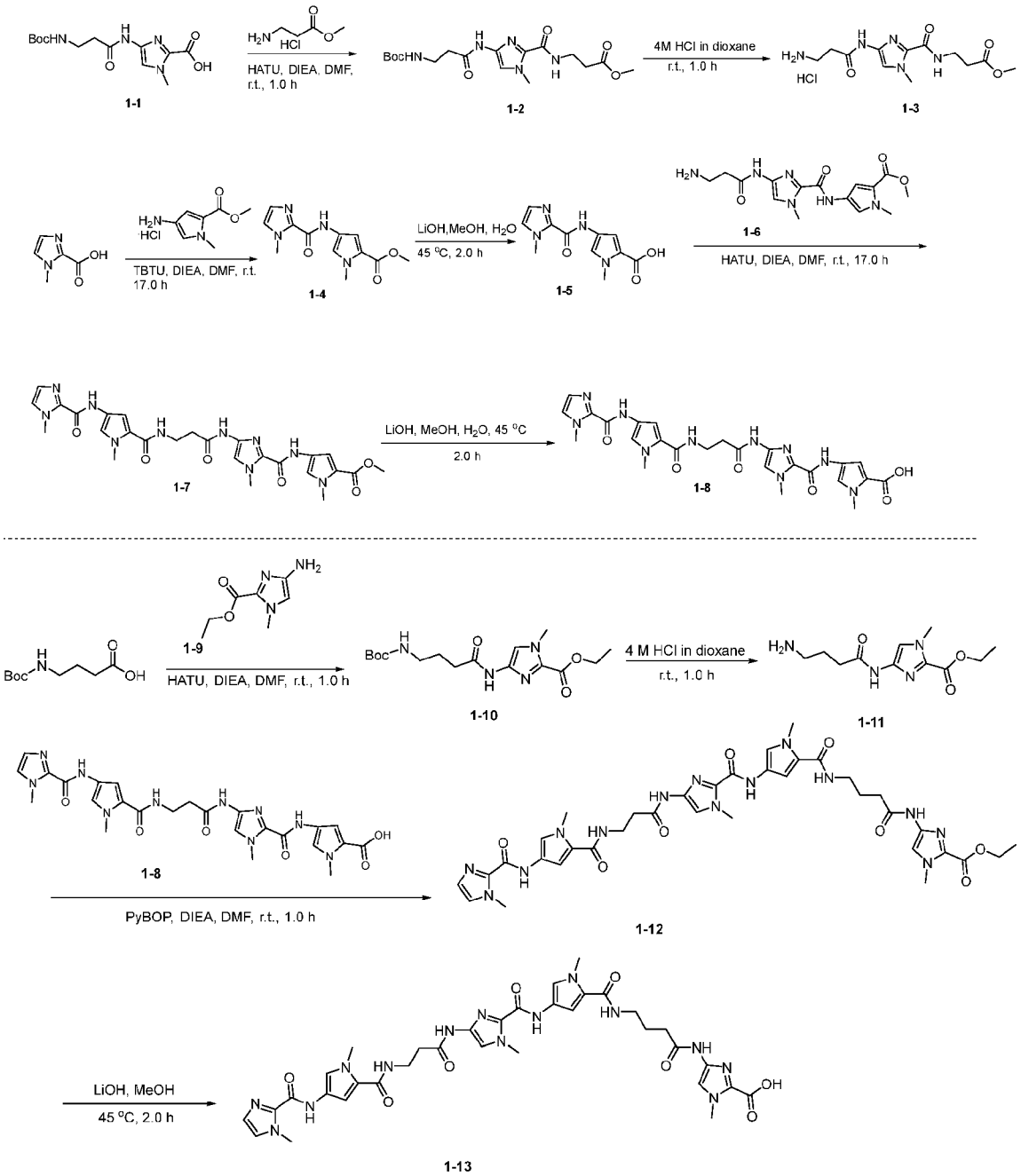
List of Abbreviation

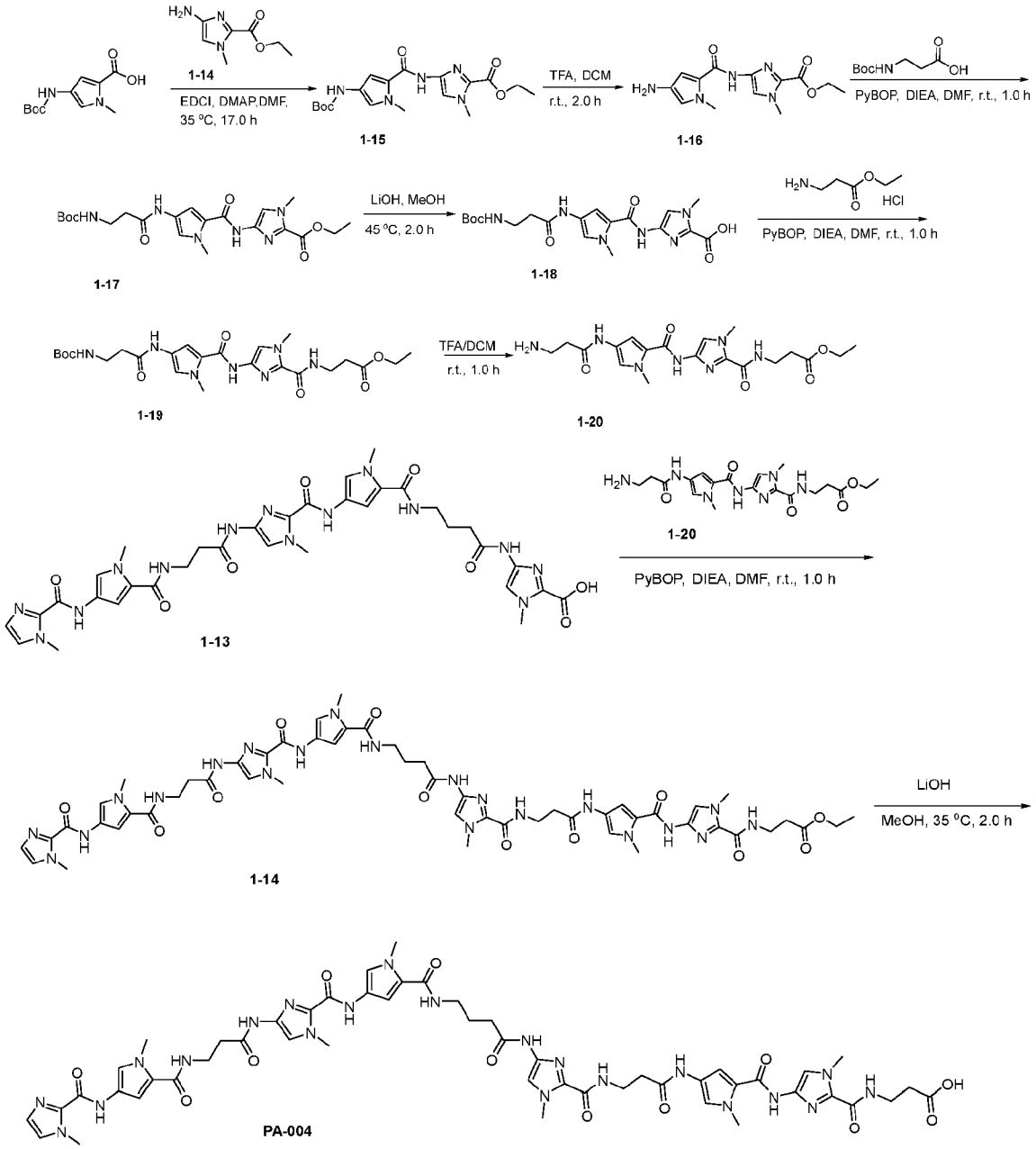
[00379] Ac_2O = acetic anhydride; AcCl = acetyl chloride; AcOH = acetic acid; AIBN = azobisisobutyronitrile; aq. = aqueous; Bu_3SnH = tributyltin hydride; CD_3OD = deuterated methanol; $CDCl_3$ = deuterated chloroform; CDI = 1,1′-Carbonyldiimidazole; DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene; DCM = dichloromethane; DEAD = diethyl azodicarboxylate; DIBAL-H = di-iso-butyl aluminium hydride; DIEA = DIPEA = N,N-diisopropylethylamine; DMAP = 4-dimethylaminopyridine; DMF = N,N-dimethylformamide; DMSO-d_6 = deuterated dimethyl sulfoxide; DMSO = dimethyl sulfoxide; DPPA = diphenylphosphoryl azide; EDC.HCl = EDCI.HCl = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; Et_2O = diethyl ether; EtOAc = ethyl acetate; EtOH = ethanol; h = hour; HATU=2-(1H-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium; HMDS = hexamethyldisilazane; HOBT = 1-hydroxybenzotriazole; i-PrOH = isopropanol; LAH = lithium aluminium hydride; LiHMDS = Lithium bis(trimethylsilyl)amide; MeCN = acetonitrile; MeOH = methanol; MP-carbonate resin = macroporous triethylammonium methylpolystyrene carbonate resin; MsCl = mesyl chloride; MTBE = methyl tertiary butyl ether; MW = microwave irradiation ; n-BuLi = n-butyllithium; NaHMDS = Sodium bis(trimethylsilyl)amide; NaOMe = sodium methoxide; NaOtBu = sodium t-butoxide; NBS = N-bromosuccinimide; NCS = N-chlorosuccinimide; NMP = N-Methyl-2-pyrrolidone; $Pd(Ph_3)_4$ = tetrakis(triphenylphosphine)palladium(0); $Pd_2(dba)_3$ = tris(dibenzylideneacetone)dipalladium(0); $PdCl_2(PPh_3)_2$ = bis(triphenylphosphine)palladium(II) dichloride; PG = protecting group; prep-HPLC = preparative high-performance liquid chromatography; PyBop = (benzotriazol-1-yloxy)-tripyrrolidinophosphonium hexafluorophosphate; Pyr = pyridine; RT = room temperature; RuPhos = 2-dicyclohexylphosphino-2′,6′-diisopropoxybiphenyl; sat. = saturated; ss = saturated solution; t-BuOH = tert-butanol; T3P = Propylphosphonic Anhydride; TBS = TBDMS = *tert*-butyldimethylsilyl; TBSCl = TBDMSCl = *tert*-butyldimethylchlorosilane; TEA = Et_3N = triethylamine; TFA = trifluoroacetic acid; TFAA = trifluoroacetic anhydride; THF = tetrahydrofuran; Tol = toluene; TsCl = tosyl chloride; XPhos = 2-dicyclohexylphosphino-2′,4′,6′-triisopropylbiphenyl.

[00380] SYNTHESIS OF REPRESENTATIVE POLYAMIDES

[00381] <u>Example 1. Synthesis of DNA-Binding Moiety PA-004</u>

[00382] Scheme 1.





[00383] ***Step 1: Synthesis of methyl 3-[(4-[3-[(tert-butoxycarbonyl)amino]propanamido]-1-methylimidazol-2-yl)formamido]propanoate***

[00384] Into a 1000 ml flask was added 4-[3-[(tert-butoxycarbonyl)amino] propanamido]-1-methylimidazole-2-carboxylic acid (11.00 g, 35.22 mmol, 1.00 equiv), DMF (300.00 mL), the mixture was cooled to 0 °C, then HATU (20.09 g, 52.83 mmol, 1.50 equiv), DIEA (18.21 g, 140.88 mmol, 4.00 equiv) was added dropwise, the mixture was stirred for 10 mins, methyl 3-aminopropanoate (3.63 g, 35.22 mmol, 1.00 equiv) was added in portions. The reaction was stirred at room temperature for 1.0 h. The reaction mixture was poured into water/ice (600 mL), the solid was filtered out and dried under vacuum. The aqueous phase was extracted by EA (3x200 mL), the organic phases were combined and washed by H_2O (1x200 mL)

and NaCl (1x200 mL), dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure. The residue was purified by silica gel column, eluted with pure EA. The fractions were combined and concentrated. Methyl 3-[(4-[3-[(tert-butoxycarbonyl)amino]propanamido]-1-methylimidazol-2-yl)formamido]propanoate (13.00 g, 87.95%) was obtained as a yellow solid. LC/MS: mass calcd. For $C_{17}H_{27}N_5O_6$: 397.20, found: 398.20 $[M+H]^+$.

[00385] ***Step 2: Synthesis of methyl 3-[[4-(3-aminopropanamido)-1-methylimidazol-2-yl] formamido]propanoate hydrochloride***

[00386] The procedure was the same as methyl 4-[4-(3-aminopropanamido)-1-methylimidazole-2-amido]-1-methylpyrrole-2-carboxylate hydrochloride, but the reaction time was 1.0 h. 11.00 g of methyl 3-[(4-[3-[(tert-butoxycarbonyl)amino]propanamido]-1-methylimidazol-2-yl)formamido]propanoate was used, 11.00 g crude of desired product was obtained as yellow oil. LC/MS: mass calcd. For $C_{12}H_{19}N_5O_4$: 297.14, found: 298.20 $[M+H]^+$.

[00387] ***Step 3: Synthesis of Methyl 1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2-carboxylate***

[00388] To a stirred solution of 1-methylimidazole-2-carboxylic acid (10.00 g, 79.29 mmol, 7.00 equiv) in DMF (150.00 mL) was added TBTU (38.19 g, 118.94 mmol, 1.50 equiv), methyl 4-amino-1-methylpyrrole-2-carboxylate hydrochloride (16.63 g, 87.24 mmol, 1.10 equiv) and DIEA (30.74 g, 237.88 mmol, 3.00 equiv) in portions at 0 °C. The resulting mixture was stirred for 17.0 h at room temperature. The reaction was poured into water/Ice (450 mL). The precipitated solids were collected by filtration and washed with H_2O (3x50 mL), dried under vacuum. Methyl 1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2-catboxylate (16.5 g, 78.37%) was obtained as a white solid. LC/MS: mass calcd. For $C_{12}H_{14}N_4O_3$: 262.11, found: 263.15 $[M+H]^+$.

[00389] ***Step 4: Synthesis of 1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2- carboxylic acid***

[00390] The procedure was the same as 4-[3-[(tert-butoxycarbonyl)amino] propanamido]-1-methylimidazole-2-carboxylic acid. 16.50 g of methyl 1-methyl-4-(1-methylimidazole-2-amido)pyrrole- 2-carboxylate was used, 12.00 g of 1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2-carboxylic acid (76.84% yield) was obtained as white solid. LC/MS: mass calcd. For $C_{11}H_{12}N_4O_3$: 248.09, found: 249.10 $[M+H]^+$.

[00391] ***Step 5: Synthesis of methyl 1-methyl-4-[1-methyl-4-(3-[[1-methyl-4-(1- methylimidazole-2-amido)pyrrol-2-yl]formamido]propanamido)imidazole-2- amido]pyrrole-2-carboxylate***

[00392] The procedure was the same as ethyl 3-[(4-[3-[(tert-butoxycarbonyl)amino] propanamido]-1-methylimidazol-2-yl)formamido]propanoate. 9.00 g of 1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2-carboxylic acid was used, 14.00 g of desired product (63.54% yield) was obtained yellow solid. LC/MS: mass calcd. For $C_{26}H_{30}N_{10}O_6$: 578.23, found: 579.10 $[M+H]^+$.

[00393] ***Step 6: Synthesis of 1-methyl-4-[1-methyl-4-(3-[[1-methyl-4-(1- methylimidazole-2-amido)pyrrol-2-yl]formamido]propanamido)imidazole-2-affordamido] pyrrole-2- carboxylic acid***

[00394] The procedure was the same as 4-[3-[(Tert-butoxycarbonyl)amino] propanamido]-1 -methylimidazole-2-carboxylic acid. 14.00 g of methyl 1-methyl-4-[1-methyl-4-(3-[[1-methyl-4-(1-

methylimidazole-2-amido) pyrrol-2-yl]formamido]propanamido)imidazole-2-amido]pyrrole-2-yl]formamidocarboxylate was used, 12.00 g of desired product (81.49% yield) was obtained as yellow solid. LC/MS: mass calcd. For $C_{25}H_{28}N_{10}O_6$: 564.22, found: 565.15$[M+H]^+$.

[00395] ***Step 7: Synthesis of ethyl 4-{4-[(tert-butoxycarbonyl)amino]butanamido} -1-methylimidazole-2-carboxylate***

[00396] The procedure was the same as ethyl 4-|3-|(tert-butoxycarbonyl)amino|propanamido|-1-methylimidazole-2-carboxylate. 7.80 g of 4-[(tert- butoxycarbonyl)amino]butanoic acid was obtained, 11.00 g of desired product was obtained as a pink solid (80.70% yield). LC/MS: mass calcd. For $C_{16}H_{26}N_4O_5$: 354.19, found: 355.15$[M+H]^+$.

[00397] ***Step 8: Synthesis of ethyl 4-(4-aminobutanamido)-1- methylimidazole-2-carboxylate***

[00398] The procedure was the same as methyl 4-[4-(3-aminopropanamido)-1- methylimidazole-2-amido]-1-methylpyrrole-2-carboxylate hydrochloride. 9.40 g of ethyl 4-{4-[(tert-butoxycarbonyl)amino]butanamido}- 1-methylimidazole-2-carboxylate was used, 6.20 g of desired product was obtained as a white solid (90.89% yield). LCMS: mass calcd. For $C_{11}H_{18}N_4O_3$: 254.14, found: 255.15$[M+H]^+$.

[00399] ***Step 9: Synthesis of ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl- 4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate***

[00400] To a stirred solution of 1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrole-2-carboxylic acid (18.20 g, 32.24 mmol, 1.00 equiv) in DMF (250.00 mL) was added DIEA (12.50 g, 96.71 mmol, 3.00 equiv), ethyl 4-(4-aminobutanamido)-1-methylimidazole-2-carboxylate (9.02 g, 35.46 mmol, 1.10 equiv) and PyBOP (20.13 g, 38.68 mmol, 1.20 equiv) at 0 °C. The resulting mixture was stirred for 1.0 h at room temperature. The reaction was poured into ice/water (800 mL). The precipitated solids were collected by filtration and washed with H_2O (3x200 mL), dried under vacuum. 24.70 g of ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate was obtained as a yellow solid (95.74% yield). LC/MS: mass calcd. For $C_{36}H_{44}N_{14}O_8$: 800.35, found: 801.30$[M+H]^+$.

[00401] ***Step 10: Synthesis of 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl- 4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylic acid***

[00402] The procedure was the same as 4-[3-[(tert-butoxycarbonyl)amino]propanamido]-1-methylimidazole-2-carboxylic acid. 24.00 g of ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate was used, 23.10 g of desired product was obtained as a yellow solid (99.36% yield). LC/MS: mass calcd. For $C_{34}H_{40}N_{14}O_8$: 772.32, found: 773.30$[M+H]^+$.

[00403] ***Step 11: Synthesis of ethyl 4-[4-[(tert-butoxycarbonyl)amino]-1-methylpyrrole- 2-amido]-1-methylimidazole-2-carboxylate***

[00404] To a stirred solution of 4-[(tert-butoxycarbonyl)amino]-1-methylpyrrole-2- carboxylic acid (11.50 g, 47.87 mmol, 1.00 equiv) in DMF (200.00 mL) was added EDCI (22.94 g, 119.66 mmol, 2.50 equiv), ethyl 4-amino-1- methylimidazole-2-carboxylate (8.10 g, 47.87 mmol, 1.00 equiv) and DMAP (14.62 g, 119.66 mmol, 2.50 equiv) at 0 °C. The resulting mixture was stirred for 17.0 h at 35 °C. After reaction, the reaction was poured into 500 mL ice/water. The precipitated solids were collected by filtration and washed with water (3x50 mL), dried under vacuum. This resulted in ethyl 4-{4-[(tert-butoxycarbonyl)amino]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylate (16.00 g, 85.48% yield) as a light yellow solid. LC/MS: mass calcd. For $C_{18}H_{25}N_5O_5$: 391.19, found: 392.30 $[M+H]^+$.

[00405] ***Step 12: Synthesis of ethyl 4-(4-amino-1-methylpyrrole-2-amido)-1- methylimidazole-2-carboxylate***

[00406] To a stirred solution of ethyl 4-{4-[(tert-butoxycarbonyl)amino]-1- methylpyrrole-2-amido}-1-methylimidazole-2-carboxylate (16.00 g, 40.88 mmol, 1.00 equiv) in DCM (135.00 mL) were added and TFA (45.00 mL) dropwise at room temperature. The resulting mixture was stirred for 2.0 h at room temperature. The resulting mixture was concentrated under vacuum. The residue brown oil was diluted with Et_2O (200 mL). The precipitated solids were collected by filtration and washed with Et_2O (2x100 mL). The resulting solid was dried under vacuum. This resulted in ethyl 4-(4-amino-1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate (16.00 g, crude) as a brown solid. LC/MS: mass calcd. For $C_{13}H_{17}N_5O_3$: 291.13, found: 292.15$[M+H]^+$.

[00407] ***Step 13: Synthesis of ethyl 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}- 1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate***

[00408] A solution of ethyl 4-(4-amino-1-methylpyrrole-2-amido)-1-methylimidazole- 2-carboxylate (12.00 g, 41.19 mmol, 1.00 equiv) and 3-[(tert-butoxycarbonyl)amino] propanoic acid (7.50 g, 39.64 mmol, 0.96 equiv), PyBOP (22.00 g, 42.28 mmol, 1.03 equiv), DIEA (45.00 g, 348.18 mmol, 8.45 equiv) in DMF (120.00 mL) was stirred for 1.0 h at room temperature. The reaction was poured into ice water (400 mL), and the mixture was stirred for 15 min. The precipitated solids were collected by filtration and washed with water (3x150 mL) and dried under vacuum. The aqueous phase was extracted by EA (3x150 mL), the combined organic phases were combined and washed by H_2O (200 mL), dried over anhydrous Na_2SO_4. The solid was filtered out and the filtrate was concentrated. The residue was purified by silica gel column chromatography, eluted with PE/EA (1:8). This resulted in 17.00 g of ethyl 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate was obtained as a yellow solid (89.28% yield). LC/MS: mass calcd. For $C_{21}H_{30}N_6O_6$: 462.22, found: 463.35$[M+H]^+$.

[00409] ***Step 14: Synthesis of 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1- methylpyrrole-2-amido)-1-methylimidazole-2-carboxylic acid***

[00410] The procedure was the same as 4-[3-[(tert-butoxycarbonyl)amino]propanamido] -1-methylimidazole-2-carboxylic acid. 12.00 g of ethyl 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate was used, 10.00 g of the desired product was obtained as a white solid (88.81% yield). LC/MS: mass calcd. For $C_{19}H_{26}N_6O_6$: 434.19, found: 435.25$[M+H]^+$.

[00411] ***Step 15: Synthesis of ethyl 3-{[4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazol-2-yl]formamido}propanoate***

[00412] A solution of 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1- methylpyrrole-2-amido)-1-methylimidazole-2-carboxylic acid (10.00 g, 23.02 mmol, 1.00 equiv) and β-alanine ethyl ester hydrochloride (4.90 g, 31.90 mmol, 1.39 equiv), PyBOP (12.50 g, 24.02 mmol, 1.04 equiv), DIEA (9.00 g, 69.64 mmol, 3.03 equiv) in DMF (120.00 mL) was stirred for 1.0 h at room temperature. The reaction was quenched by the addition of water (500 mL) at room temperature. The resulting mixture was extracted with EtOAc (3x400 mL). The combined organic layers were washed with brine (3x200 mL), dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure. The residue was purified by silica gel column chromatography, eluted with PE/EA (1:8) to afford ethyl 3-{[4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazol-2-yl]formamido}propanoate (12.00 g, 93.80%) as a yellow solid. LC/MS: mass calcd. For $C_{24}H_{35}N_7O_7$: 533.26, found: 534.30$[M+H]^+$.

[00413] ***Step 16: Synthesis of ethyl 3-({4-[4-(3-aminopropanamido)-1-methylpyrrole- 2-amido]-1-methylimidazol-2-yl}formamido)propanoate***

[00414] The procedure was the same as ethyl 4-(4-amino-1-methylpyrrole-2-amido)-1- methylimidazole-2-carboxylate. Ethyl 3-{[4-(4- {3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazol-2-yl]formamido}propanoate was used to obtain 12.00 g crude of the desired product was obtained as a white solid. LC/MS: mass calcd. For $C_{19}H_{27}N_7O_5$: 433.21, found: 434.25$[M+H]^+$.

[00415] ***Step 17: Synthesis of ethyl 3-[(1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-amido}imidazol-2-yl)formamido]propanoate***

[00416] The procedure was the same as ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate. 10.00 g of 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylic acid was used, 13.60 g of the desired product was obtained as a yellow solid (88.61% yield). Some pure product was obtained as a light yellow solid after purification by Prep-HPLC. HRMS: mass calcd. For $C_{53}H_{65}N_{21}O_{12}$: 1187.5122, found: 1188.5153$[M+H]^+$.

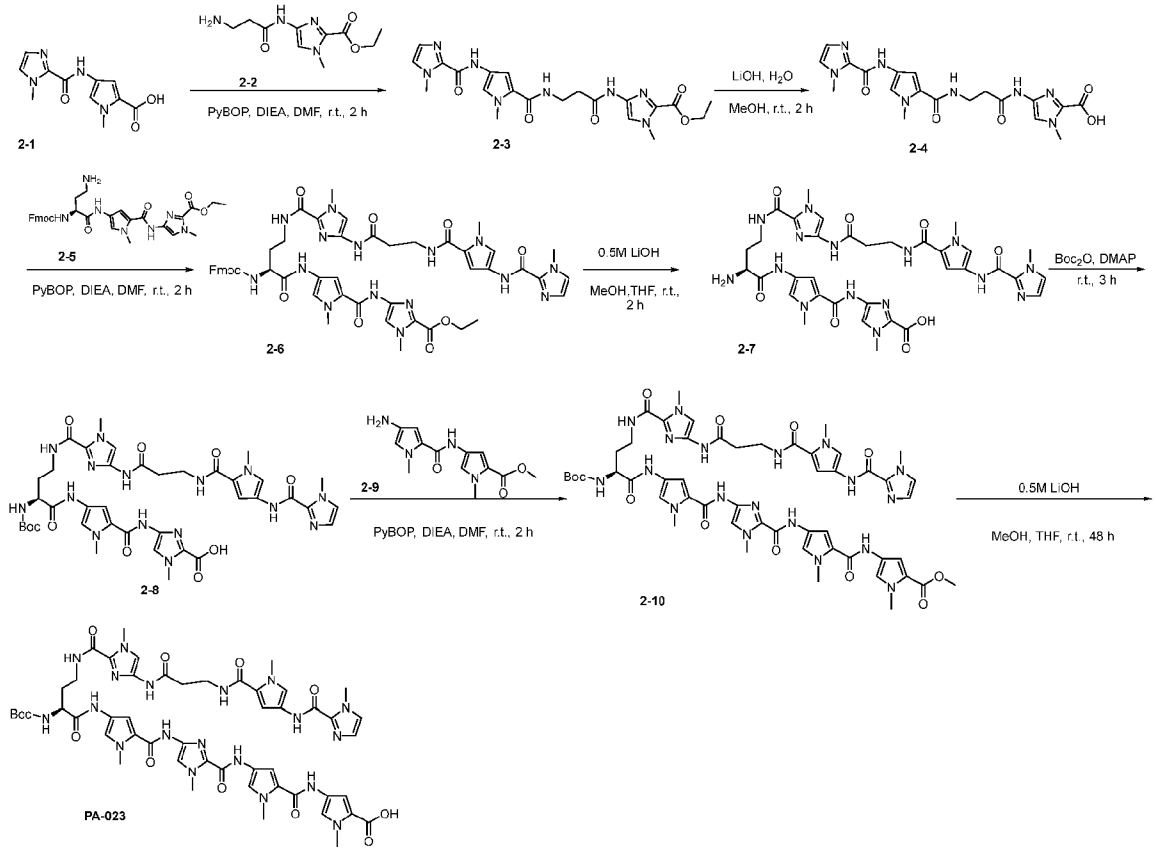
[00417] ***Step 18: Synthesis of 3-[(1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-***

amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-amido}imidazol-2-yl)formamido]propanoic acid (PA-004)

[00418] The procedure was the same as 4-[3-[(tert-butoxycarbonyl)amino]propanamido]-1-methylimidazole-2-carboxylic acid, but the reaction temperature was 35 °C. 10.60 g of ethyl 3-[(1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-amido}imidazol-2-yl)formamido]propanoate was used, 10.00 g crude of the desired product was obtained as a yellow solid. LC/MS: mass calcd. For $C_{51}H_{61}N_{21}O_{12}$: 1159.48, found: 581.25$[M/2+H]^+$.

[00419] **Example 2. Synthesis of DNA-Binding Moiety (PA-023)**

[00420] **Scheme 2.**



[00421] ***Step 1: Synthesis of ethyl 1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2- amido)pyrrol-2-yl][formamido}propanamido)imidazole-2-carboxylate***

[00422] The procedure was the same as ethyl 1-methyl-4-[4-({1-methyl-4-[1- methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate, but the reaction time was 2.0 h. 1.50 g of ethyl 4-(3-aminopropanamido)-1-methylimidazole-2-carboxylate was used, 2.00 g of desired product was obtained as an off-white solid (68.09% yield). LC/MS: mass calcd. For $C_{21}H_{26}N_8O_5$: 470.20, found: 471.40 $[M+H]^+$.

[00423] ***Step 2: Synthesis of 1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido) pyrrol-2-yl][formamido}propanamido)imidazole-2-carboxylic acid***

[00424] The procedure was the same as 4-[3-[(tert-butoxycarbonyl)amino]propanamido] -1-methylimidazole-2-carboxylic acid, but the reaction temperature was room temperature, the reaction time was 2.0 h. 2.00 g of ethyl 1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido} propanamido)imidazole-2-carboxylate was used, 1.80 g of 1-methyl-4-(3-{[1-methyl- 4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-carboxylic acid was obtained as an off-white solid (95.71% yield). LC/MS: mass calcd. For $C_{19}H_{22}N_8O_5$: 442.17, found: 443.10 $[M+H]^+$.

[00425] ***Step 3: Synthesis of ethyl 4-{4-[(2S)-2-{[(9H-fluoren-9-ylmethoxy)carbonyl] amino}-4-{[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl][formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylate***

[00426] The procedure was the same as ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl- 4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate, but the reaction time was 2.0 h. 1.60 g of ethyl 4-{4-[(2S)-4-amino-2-{[(9H-fluoren-9-ylmethoxy)carbonyl]amino}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylate was used, 1.90 g of desired product was obtained as a light yellow solid (70.20% yield). LC/MS: mass calcd. For $C_{51}H_{55}N_{15}O_{10}$: 1037.43, found: 1038.45 $[M+H]^+$.

[00427] ***Step 4: Synthesis of 4-[4-(4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4-{[1- methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl][formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-amido)-1-methylpyrrole-2-amido]-1-methylpyrrole-2-carboxylic acid***

[00428] A mixture of ethyl 4-{4-[(2S)-2-{[(9H-fluoren-9-ylmethoxy)carbonyl]amino}-4- {[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylate (1.90 g, 1.83 mmol, 1.00 equiv) and LiOH (0.22 g, 9.15 mmol, 5.00 equiv) in MeOH (5.00 mL), THF (15.00 mL) and H_2O (18.30 mL) was stirred for 2.0 h at room temperature. The resulting mixture was used in the next step directly without further purification. LC/MS: mass calcd. For $C_{34}H_{41}N_{15}O_8$: 787.33, found: 788.40 $[M+H]^+$.

[00429] ***Step 5: Synthesis of 4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4-{[1-methyl-4-(3- {[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl][formamido}propanamido)imidazol-2-yl][formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylic acid***

[00430] The mixture of 4-{4-[(2S)-2-amino-4-{[1-methyl-4-(3-{[1-methyl-4-(1- methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylic acid (1.40 g, 1.78 mmol, 1.00 equiv) in MeOH/THF/H_2O (5.00 mL/15.00 mL/18.30 mL) was added di-tert-butyl dicarbonate (0.78 g, 3.55 mmol, 2.00 equiv) and DMAP (0.02 g, 0.18 mmol, 0.10 equiv). The reaction was stirred at room temperature for 3.0 h. The mixture was added with H_2O (30 mL). The mixture was filtered through a Celite pad, and the solid was washed with ethyl acetate (3x30 mL) to afford 4-{4-[(2S)-2-[(tert- butoxycarbonyl)amino]-4-{[1-methyl-4-(3-{[1-methyl-

4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylic acid (1.20 g, 76.05% yield) as a yellow solid. LC/MS: mass calcd. For $C_{39}H_{49}N_{15}O_{10}$: 887.38, found: 888.45 $[M+H]^+$.

[00431] ***Step 6: Synthesis of methyl 4-[4-(4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4- {[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-amido)-1-methylpyrrole-2-amido]-1-methylpyrrole-2-carboxylate***

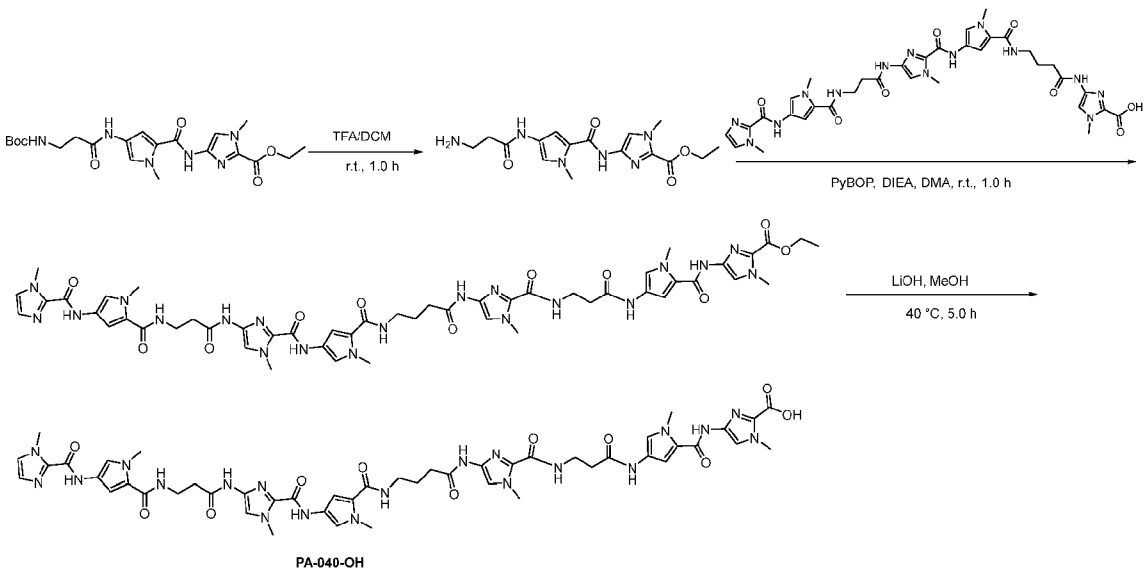
[00432] The procedure was the same as ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4- (3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido) imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate, but the reaction time was 2.0 h. 1.20 g of 4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4-{[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-carboxylic acid was used, 1.10 g of the desired product was obtained as a yellow solid (71.01% yield). LC/MS: mass calcd. For $C_{52}H_{63}N_{19}O_{12}$: 1145.49, found: 1146.50 $[M+H]^+$.

[00433] ***Step 7: Synthesis of 4-[4-(4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4-{[1- methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-amido)-1-methylpyrrole-2-amido]-1-methylpyrrole-2-carboxylic acid (PA-023)***

[00434] The procedure was the same as 4-[4-(4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]- 4-[(1-methyl-4-{1-methyl-4-[1-methyl-4-(1-methylimidazole-2-amido)pyrrole-2-amido]pyrrole-2-amido}imidazol-2-yl)formamido]butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-amido)-1-methylpyrrole-2-amido]-1-methylpyrrole-2-carboxylic acid. 1.00 g of methyl 4-[4-(4-{4-[(2S)-2-[(tert-butoxycarbonyl)amino]-4-{[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazol-2-yl]formamido}butanamido]-1-methylpyrrole-2-amido}-1-methylimidazole-2-amido)-1-methylpyrrole-2-amido]-1-methylpyrrole-2-carboxylate was used, 400.00 mg of the desired product was obtained as a white solid (39.16% yield). LC/MS: mass calcd. For $C_{51}H_{61}N_{19}O_{12}$: 1131.47, found: 1132.65 $[M+H]^+$.

[00435] **Example 3. Synthesis of DNA-Binding Moiety (PA-040)**

[00436] **Scheme 3.**



[00437] ***Step 1: Synthesis of ethyl 4-[4-(3-aminopropanamido)-1-methylpyrrole-2- amido]-1-methylimidazole-2-carboxylate***

[00438] The procedure was the same as ethyl 4-(4-amino-1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate (**Intermediate 1-16**, **Example 1**). 2.00 g of ethyl 4-(4-{3-[(tert-butoxycarbonyl)amino]propanamido}-1-methylpyrrole-2-amido)-1-methylimidazole-2-carboxylate was used, 2.00 g crude of the desired product was obtained as a white solid. LC/MS: mass calcd. For $C_{16}H_{22}N_6O_4$: 362.17, found: 363.25$[M+H]^+$.

[00439] ***Step 2: Synthesis of tert-butyl N-(3-[[3-(1,3-dioxoisoindol-2-yl)propyl](methyl) amino]propyl)-N-methylcarbamate***

[00440] The procedure was the same as ethyl 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylate (**Int.1-12**), but the solvent was DMA. 3.00 g of 1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazole-2-carboxylic acid was used, 4.30 g of the desired product was obtained as a yellow solid (96.84% yield). LC/MS: mass calcd. For $C_{50}H_{60}N_{20}O_{11}$: 1116.48, found:1117.60$[M+H]^+$.

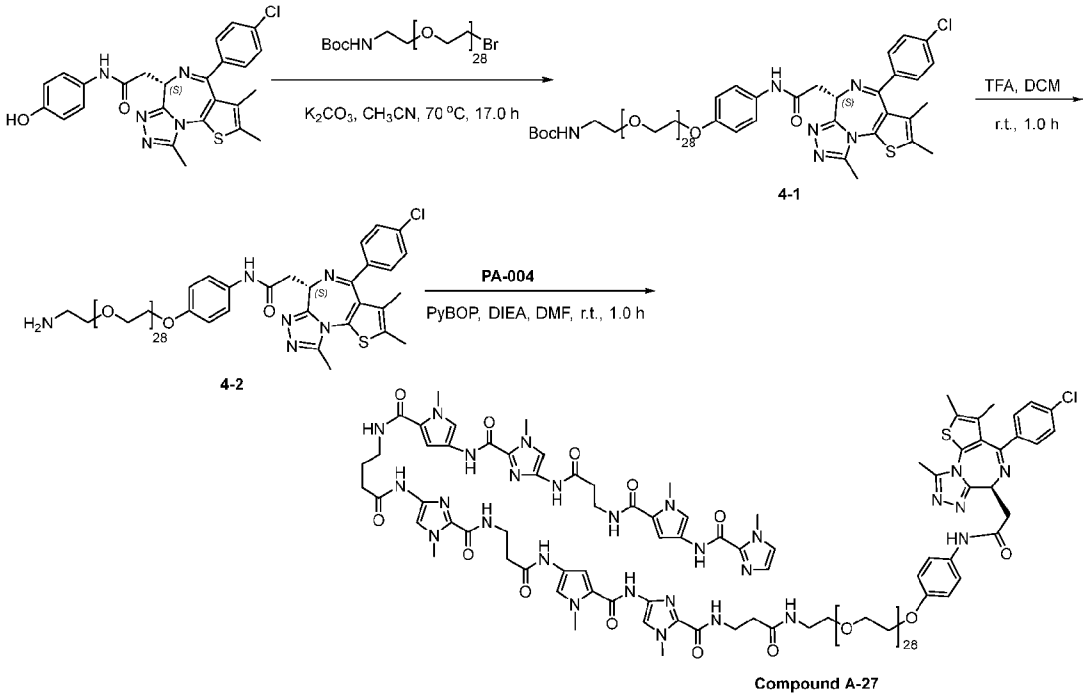
[00441] ***Step 3: Synthesis of 1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1- methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-amido}imidazole-2-carboxylic acid (PA-040-OH)***

[00442] The procedure was the same as **Example 1** (PA-004), but the reaction temperature was 40 °C, the reaction time was 5.0 h. 4.20 g of ethyl 1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-

amido}imidazole-2-carboxylate was used, 4.00 g of the desired product was obtained as a yellow solid (97.97% yield). LC/MS: mass calcd. For $C_{48}H_{56}N_{20}O_{11}$: 1088.44, found: 1089.55$[M+H]^+$.

[00443] **Example 4. Synthesis of A-27**

[00444] **Scheme 4.**



[00445] ***Step 1. Synthesis of Intermediate 4-1***

[00446] To a stirred solution of (S)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)-N-(4-hydroxyphenyl)acetamide (27.00 mg, 0.05 mmol, 1.00 equiv) in CH_3CN (1.50 mL) were added tert-butyl N-(86-bromo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaoxahexaoctacontan-1-yl)carbamate (80.00 mg, 0.05 mmol, 1.00 equiv) and K_2CO_3 (22.76 mg, 0.16 mmol, 3.00 equiv). The resulting mixture was stirred for 17.0 h at 70 °C. The resulting mixture was filtered, the filter cake was washed with EtOAc (3x8 mL). The filtrate was concentrated under reduced pressure and purified by TLC-plate (CH_2Cl_2/MeOH=8:1) to afford tert-butyl (S)-(86-(4-(2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)phenoxy)3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaoxahexaoctacontyl)carbamate (81.00 mg, 77.40% yield) as a brown solid.

[00447] LC/MS: mass calcd. For $C_{88}H_{147}ClN_6O_{32}S$: 1866.95, found: 623.90$[M/3+H]^+$.

[00448] ***Step 2. Synthesis of Intermediate 4-2***

[00449] A solution of tert-butyl (S)-(86-(4-(2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)phenoxy)-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-

octacosaoxahexaoctacontyl)carbamate (70.00 mg, 0.04 mmol, 1.00 equiv) and TFA (0.20 mL) in DCM (1.00 mL) was stirred for 1.0 h at room temperature. The resulting mixture was concentrated under reduced pressure. This resulted in (S)-N-(4-((86-amino-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaoxahexaoctacontyl)oxy)phenyl)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamide (70.00 mg, crude) as a yellow oil.

[00450] LC/MS: mass calcd. For $C_{83}H_{139}ClN_6O_{30}S$: 1766.89, found: 590.60[M/3+H]$^+$.

[00451] ***Step 3. Synthesis of A-26***

[00452] To a stirred solution of 3-[(1-methyl-4-{1-methyl-4-[3-({1-methyl-4-[4-({1-methyl-4-[1-methyl-4-(3-{[1-methyl-4-(1-methylimidazole-2-amido)pyrrol-2-yl]formamido}propanamido)imidazole-2-amido]pyrrol-2-yl}formamido)butanamido]imidazol-2-yl}formamido)propanamido]pyrrole-2-amido}imidazol-2-yl)formamido]propanoic acid (37.80 mg, 0.03 mmol, 1.00 equiv) in DMF (1.00 mL) was added DIEA (25.27 mg, 0.19 mmol, 6.00 equiv), (S)-N-(4-((86-amino-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaoxahexaoctacontyl)oxy)phenyl)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamide (57.62 mg, 0.03 mmol, 1.00 equiv) and PyBOP (25.43 mg, 0.05 mmol, 1.50 equiv) at 0 °C. The resulting mixture was stirred for 1.0 h at room temperature. The resulting mixture was filtered and purified by Perp-HPLC with the following condonation: Column: XBridge Prep C18 OBD Column, 19*150 mm, 5μm; Mobile Phase A: Water(10 mmol/L NH_4HCO_3), Mobile Phase B: ACN; Flow rate: 25 mL/min; Gradient: 45% B to 65% B in 11 min, 65% B; Wave Length: 254 nm; RT1(min): 3.05; Number Of Runs: 0. The fractions were combined and lyophilized to afford (S)-N-(3-((5-((2-((1-(4-(2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)phenoxy)-88-oxo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaoxa-87-azanonacontan-90-yl)carbamoyl)-1-methyl-1H-imidazol-4-yl)carbamoyl)-1-methyl-1H-pyrrol-3-yl)amino)-3-oxopropyl)-1-methyl-4-(4-(1-methyl-4-(1-methyl-4-(3-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)propanamido)-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1H-imidazole-2-carboxamide (13.80 mg, 14.46%) as a white solid.
HRMS: mass calcd. For $C_{134}H_{198}ClN_{27}O_{41}S$: 2908.3648, found: 2909.3729 [M+H]$^+$.

[00453] HPLC: 99.401% purity .

[00454] **Example 9. General synthesis and purification of the compounds of the disclosure**

[00455] Compounds of the disclosure were made by methods similar to **Example 4**. The compounds were subsequently purified by HRMS methods A or B.

[00456] Method A: Instrument: Waters Acquity I Class UPLC with Xevo G2-XSQ Tof HRMS; Column: ACQUITY UPLC BEH-C18, 2.1 x 50 mm, 2.7 μm; mobile phase A: H_2O (0.1% HCOOH), mobile B, CAN (0.1% HCOOH); Flow rate:0.4mL/min; Gradient: 10% B to 95% B in 1.5 min, hold 95% for another 0.5 min, then down to 10% B in 0.3 min, hold 10% B for another 0.7 min; detector: 254nm.

[00457] Method B: Instrument: Waters Acquity I Class UPLC with Xevo G2-XS Q Tof HRMS; Column: ACQUITY UPLC BEH-C18, 2.1 x 50 mm, 2.7μm; mobile phase A: H_2O (0.1% HCOOH), mobile B, CAN (0.1% HCOOH); Flow rate:0.4 mL/min; Gradient: 5% B to 40% B in 2.0min, to 95% in another 1.5 min, hold 95% for 1.5 min, then down to 5% B in 0.3 min, hold 5% B for another 0.7 min; detector: 254nm.

[00458] Experimental data for the compounds of the disclosure purified by Method A are provided in Table 5.

Table 5. LCMS analysis of compounds of the disclosure.

Comp. No.	Exact Mass	Free Base MW	Mass Found
A-1	2027.8405	2029.656	2028.8478
A-2	2410.0176	2411.5808	2411.0159
A-3	2035.8435	2037.1754	2036.8448
A-4	2027.8405	2029.656	2028.8377
A-5	2130.8607	2132.2518	2131.8592
A-6	2332.0468	2333.977	2333.0514
A-7	2091.9392	2093.2684	2092.9497
A-8	1939.7881	1941.55	1940.788
A-9	1851.7356	1853.444	1852.7413
A-10	1763.6832	1765.338	1764.6895
A-11	1675.6308	1677.232	1676.6254
A-12	2101.8341	2103.2098	2102.831
A-14	1954.7558	1956.0398	1955.7612
A-15	2264.9869	2266.948	2265.9864
A-16	2264.9869	2266.948	2265.9857
A-17	2336.024	2338.027	2337.0146
A-18	2374.1171	2375.633	2375.1068
A-19	2414.1383	2415.6664	2415.1287
A-20	2449.0849	2450.6864	2450.0921
A-21	1880.7541	1881.9938	1881.7648
A-22	2203.9453	2205.868	2204.9598
A-23	2468.1026	2470.186	2469.0977
A-24	2556.1551	2558.292	2557.1534
A-26	2776.2861	2778.557	2777.2906
A-27	2908.3648	2910.716	2909.3729
A-28	2996.4172	2998.822	2997.4188
A-29	2302.025	2303.947	2303.0312
A-30	2273.1382	2274.573	2274.1414
A-32	2302.107	2303.5354	2303.109
A-33	2272.0964	2273.5094	2273.1027
A-35	1777.7339	1779.398	1778.7456
A-36	2188.9376	2190.3648	2189.9359

Comp. No.	Exact Mass	Free Base MW	Mass Found
A-37	1777.7339	1779.398	1778.7411
A-38	2085.9174	2087.769	2086.9287
A-39	1880.7541	1881.9938	1881.7619
A-40	2188.9376	2190.3648	2189.9468
A-41	2480.0959	2481.7158	2481.0976
A-42	2414.1053	2415.6814	2415.1071
A-43	2334.1321	2335.605	2335.1469
A-44	2025.9486	2027.234	2026.955
A-45	2336.024	2338.027	2359.0190[M+Na]
A-46	2207.9654	2209.896	2208.9786
A-47	1722.6916	1724.709	1723.709
A-48	1898.7964	1900.921	1899.8109
A-49	2206.9799	2209.292	2207.9987
A-50	2438.0489	2439.6348	2439.0492
A-51	2370.0791	2371.6284	2373.0673
A-52	1854.8258	1856.47	1855.8324
A-53	2030.9307	2032.682	2031.9369
A-54	2339.1142	2341.053	2340.1233
A-55	1859.7387	1860.9634	1860.7429
A-56	2216.0803	2217.477	2217.0815
A-57	1907.8968	1909.106	1908.8999
A-58	2308.1065	2309.574	2309.1194
A-59	1999.923	2001.203	2000.9344
A-60	2024.9646	2026.25	1941.9203
A-61	2333.1481	2334.621	2250.0935
A-62	2344.0270	2345.5464	2345.0348
A-65	2373.0536	2374.5884	2374.0653
A-66	2240.9750	2242.4294	2241.9809
A-67	2152.9225	2154.3234	2153.9327
A-68	2302.0165	2303.5094	2303.0264
A-69	2169.9378	2171.3504	2170.9403
A-70	2081.8854	2083.2444	2082.8987
A-71	2309.0131	2311.0010	2310.0186
A-72	2441.0234	2442.5948	2442.0414
A-73	2338.0033	2339.9990	2339.0168
A-74	2440.0394	2441.6108	2441.0594
A-75	2337.0192	2339.0150	2338.0347

BIOLOGICAL EXAMPLES

[00459] **Example B1. Activity**

[00460] Fibroblast: a cell type derived from a skin biopsy of a patient. These cells are not altered genetically, so they serve as a primary cell culture model of disease

[00461] iPSC: induced pluripotent stem cell, a cell type that results as a reprogramming of another cell type (typically skin cells or blood cells) into a more embryonic-like state that enables the development of other cell types to model therapeutic effects of drugs *in vitro*.

[00462] SNP: Single Nucleotide Polymorphism, a variation in a single base pair in a DNA sequence

[00463] Molecular Biology Toolkit:

- qPCR primer probe sets:
  - RNA input normalization was assessed utilizing human glyceraldehyde 3-phosphate dehydrogenase (hGAPDH) TaqMan assay (ThermoFisher cat# 4351370) or Human Cyclophillin (IAPP) TaqMan assay (ThermoFisher cat# 4351372)
  - Total HTT detection was assessed utilizing human Htt TaqMan assay (ThermoFisher cat# 4331182)
  - Allele-specific detection of human HTT expression in HD cells containing the SNP rs362331C/T (Exon 50): for each assay, allele-specific probes to detect the SNP variant contained locked nucleic acid bases to improve allele discrimination, as compared to unmodified DNA probes.
    - 362331-F (331 forward primer): TCTCCTCCACAGAGTTTGTGA
    - 362331-R (331 reverse primer): CCTTCTTTCTGGACTAAGAAGCTG
    - 362331-C probe: TCC CTC ATC + C + AC TGT GT
    - 362331-T probe: CTC + A + T + C + T + A + C TGT GT
  - qPCR was performed using Agpath ID one-step reverse transcriptase polymerase chain reaction (RT-PCR) reagent

[00464] Protein measurements were performed via western blots probing with antibody MW1 (polyQ specific) to assess reduction of mtHTT alone. D7F7 (a.a surrounding Pro1218) was used to visualize both wtHTT and mtHTT. Lysates were standardized by DC prior to separation on a 3-8% Tris-acetate gel and transferred via wet transfer method onto nitrocellulose membranes. Blots were probed with the previously mentioned antibodies and complementary fluorescent secondary antibodies and imaged on the Li-Cor Odyssey® DLx Imaging system.

[00465] Antibody pairing of 2B7 (a.a. 1-17) and MW1 (polyQ specific) will be used to track mtHTT levels while pairing of MAB2166 (a.a. 181-810) and MAB5490 (a.a. 115-129) will be employed to track total full-length HTT.

[00466] Screening of HD molecules methods: GM09197 and/or GM04022 fibroblasts were cultured in T175 flasks incubated at 37 °C and 5% CO_2. Once confluency was reached, the media was removed, the cells were washed with 1X PBS, and cells were dissociated using TrypLE™ Express Enzyme. Media was added to the enzyme and collected into a 15-mL conical tube and centrifuged at 500xg for 5 minutes to pellet the cells. Media and enzyme were aspirated using a serological pipette. Cells were resuspended in fresh media and counted using a Countess 3 Automated Cell Counter. Cells were plated at a density of 15,000 cells/well in a tissue culture-treated polystyrene 96-well dish and incubated at 37 °C and 5% CO_2

overnight. The next day, media was removed using an 8-channel aspirator. 200 µL media/well are added back into the plate. The molecules are formulated to 1 mM and are dispensed using a Multidrop™ Pico 8 Digital Dispenser. After a 48-hour incubation with compounds, media was removed from plates, cells were washed with 1X PBS, and cells were lysed in 40 µL per well guanidinium thiocyanate buffer. RNA was isolated and purified in 382-well glass fiber column plates using chaotropic salts. Human mtHTT, wtHTT, and GAPDH mRNA were measured via RT-PCR using the ThermoFisher QuantStudio™ 7 Flex System in 384-well format. Results of HTT levels were normalized to GAPDH mRNA levels. Normalized HTT mRNA levels were expressed relative to vehicle-treated samples to assess fold change after molecule treatment.

[00467] iPSC-Neuron Duration of Action of HD molecules methods: Fibroblasts isolated from HD patients were reprogrammed into iPSCs expanded in the presence of cytokines and transduced with the Sendai virus, a cytoplasmic RNA vector. These iPSCs expressed stem cell markers and have normal karyotypes and express the pluripotent markers Nanog, Tra-1-60, and SSeA-44. iPSC-derived neuron differentiation methodology followed standard protocols for mixed cortical neuron differentiation resulting in immunohistochemical staining of iPS-Neuron of Tuj1 and Map2. iPSC-neuronal precursor cells were plated at 300,000 cells/well in a PLO/Laminin-521 coated culture-treated polystyrene 96-well dish and incubated at 37 °C and 5% CO_2. The next day, media was changed to allow neuron precursor cells to continue maturation into neurons. Four days later, media was refreshed, and cells were treated with mitotic inhibitor to remove any remaining dividing cells, resulting in a pure neuronal culture. Three days later, media was removed. 200 µL media/well were added back into the plate. The molecules were formulated to 1 mM and dispensed using a Multidrop™ Pico 8 Digital Dispenser. After 96-hour incubation with compounds, media was removed and refreshed and the cells were retreated. After 7 days of compound exposure, media was removed from plates, and cells were lysed in 60 µL of Ambion Lysis buffer. RNA is isolated using PureLink™ RNA isolation kits. cDNA is synthesized with Agilent Superscript II kit. Human mtHTT, wtHTT, and GAPDH mRNA were measured via RT-PCR using the ThermoFisher QuantStudio™ 7 Flex System in 384-well format. Results of HTT levels were normalized to GAPDH mRNA levels. Normalized HTT mRNA levels were expressed relative to vehicle-treated samples to assess fold change after treatment with the HD compounds.

[00468] E_{min} for each compound is the lowest %HTT concentration observed within a compound concentration of 0.5nM to 1000 nM.

[00469] Representative *in vitro* biochemical data is presented in **Table 6** and **Table 7**, where A > 90%, B is 90% to 80%, C < 80%.

Table 6. Representative biochemical data in GM09197 fibroblasts.

Comp. No.	GM09197 IC_{50} 48 hr C Allele (nM)	GM 09197 E min 48 Hr C Allele (%)	GM09197 IC_{50} 48 hr T Allele (nM)	Gm09197 Emin 48 hr T Allele (%)
A-1	-	A	113.637	C

A-2	-	A	31.952	C
A-3	-	A	38.667	C
A-5	-	A	9.8	C
A-7	-	A	>1000	A
A-20	>1000	A	>1000	A
A-21	>1000	A	>1000	A
A-29	>1000	A	>1000	B
A-30	>1000	A	104.895	C
A-31	>1000	B	>1000	B
A-32	>1000	B	>1000	B
A-35	>1000	A	>1000	B
A-36	>1000	A	>1000	A
A-37	>1000	B	>1000	B
A-38	>1000	A	>1000	A
A-39	>1000	B	>1000	B
A-40	>1000	A	>1000	A
A-45	>1000	C	>1000	C
A-46	>1000	B	>1000	A
A-47	>1000	C	>1000	C
A-48	>1000	A	>1000	A
A-49	>1000	A	>1000	A
A-52	>1000	C	>1000	B
A-53	>1000	A	>1000	B
A-54	>1000	B	>1000	B
A-55	>1000	A	>1000	A
A-56	>1000	C	104	C
A-58	>1000	B	>1000	B
A-59	>1000	B	>1000	B
A-68	-	A	17	C
A-71	>1000	A	>1000	A
A-72	>1000	B	>1000	A
A-73	>1000	A	>1000	A
A-74	>1000	B	>1000	A
A-75	>1000	A	>1000	A

Table 7. Representative biochemical data in GM04022 fibroblasts.

Comp. No.	GM04022 IC_{50} 48 hr C allele	GM 04022Emin 48 hr C Allele	GM04022 IC_{50} 48 hr T Allele	GM04022 Emin 48 hr T Allele
A-2	-	A	>1000	C
A-3	-	A	36	C
A-5	-	A	2	C

[00470] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of treating an individual having an expanded CAG repeat disorder, the method comprising administering an effective amount of a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded CAG repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus.

2. The method of claim 1, wherein the expanded CAG repeat disorder is Huntington's disease (HD).

3. The method of claim 1 or claim 2, wherein the first terminus comprises a polyamide.

4. The method of any one of claims 1 to 3, wherein the second terminus comprises a bromodomain binding moiety.

5. The method of any one of claims 1 to 4, wherein the second terminus comprises a moiety capable of binding to a bromodomain and extra terminal domain (BET) family member.

6. The method of claim 5, wherein the BET family member is BRD2, BRD3, BRD4, or BRDT.

7. The method of any one of claims 1 to 4, wherein the bromodomain is CBP/p300, PCAF (P300/CBP-Associated Factor), CECR2 (cat eye syndrome chromosome region candidate 2), BRPF (bromodomain and PHD finger-containing protein), ATAD2/ATAD2B (chromatin remodeling proteins), TRIM24 (Tripartite motif-containing 24), BAZ2 (Bromodomain Adjacent to Zinc finger), or TAF1 (TBP associated factors).

8. The method of any one of claims 1 to 7, wherein the gene is huntingtin (*HTT*).

9. The method of any one of claims 1 to 8, wherein the expanded CAG repeat has at least 36 repeats.

10. The method of any one of claims 1 to 9, wherein expression of the gene having the expanded nucleotide repeat is reduced.

11. The method of any one of claims 1 to 10, wherein the method reduces one or more symptoms of Huntington's disease.

12. The method of claim 11, wherein the one or more symptoms are selected from chorea, cognitive decline, abnormal libido, abnormal eye movement, abnormal sense of smell, aggression, agitation, anxiety, apathy, bradykinesia, bradyphrenia, clumsiness, delusions, depression, difficulty walking, disinhibition, dystonia, gait imbalance, muscle weakness, hallucinations, hostility, hypokinesia, irritability, memory impairment, myoclonus, obsessive-compulsive behavior, poor fine motor coordination, seizure, speech articulation difficulties, staring gaze, weight loss, abnormal cholesterol metabolism, abnormal cerebral white matter, alcoholism, Babinski sign, caudate atrophy, cerebral atrophy, choking, clonus, degeneration of the

striatum, excessive daytime sleepiness, impaired visuospatial constructive cognition, inability to walk, insomnia, mutism, oral-pharyngeal dysphagia, rigidity, suicidal ideation, cerebellar atrophy, dementia, gate ataxia, gliosis, hyperreflexia, neuronal loss, or personality changes.

13. A method of decreasing expression of a gene having an expanded CAG repeat in a cell, the method comprising contacting the cell with an effective amount of a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus.

14. The method of claim 13, wherein the method further comprises treating an expanded CAG repeat disorder.

15. The method of claim 13 or claim 14, wherein the first terminus comprises a polyamide.

16. The method of any one of claims 13 to 15, wherein the second terminus comprises a bromodomain binding moiety.

17. The method of any one of claims 13 to 16, wherein the second terminus comprises a moiety capable of binding to a bromodomain and extra terminal domain (BET) family member.

18. The method of claim 17, wherein the BET family member is BRD2, BRD3, BRD4, or BRDT.

19. The method of any one of claims 13 to 16, wherein the bromodomain is CBP/p300, PCAF (P300/CBP-Associated Factor), CECR2 (cat eye syndrome chromosome region candidate 2), BRPF (bromodomain and PHD finger-containing protein), ATAD2/ATAD2B (chromatin remodeling proteins), TRIM24 (Tripartite motif-containing 24), BAZ2 (Bromodomain Adjacent to Zinc finger), or TAF1 (TBP associated factors).

20. The method of any one of claims 13 to 19, wherein the gene is huntingtin (*HTT*).

21. The method of any one of claims 14 to 20, wherein the expanded repeat disorder is Huntington's disease (HD).

22. The method of any one of claims 13 to 21, wherein the expanded CAG repeat has at least 36 repeats.

23. The method of any one of claims 13 to 22, wherein expression of the gene having the expanded CAG repeat is reduced.

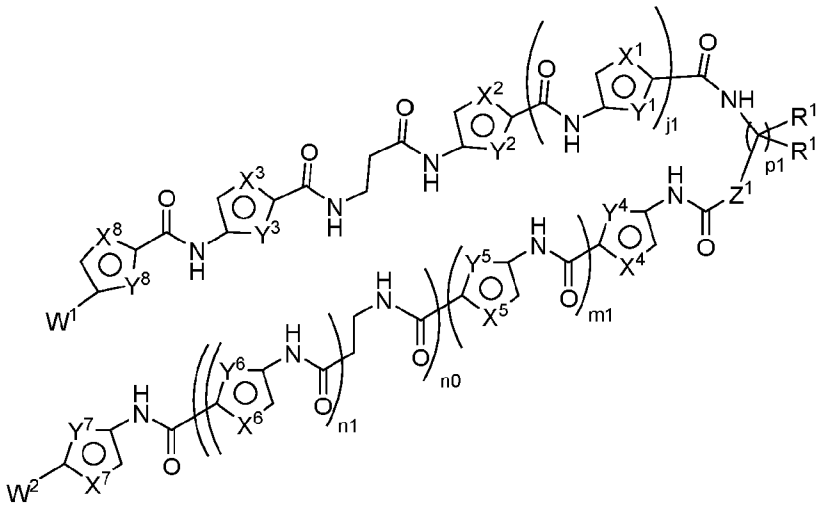
24. A transcription modulator molecule having a first terminus, a second terminus, and an oligomeric backbone moiety, wherein:

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus; and

wherein the DNA-binding moiety comprises a structure of Formula (A-1):



Formula (A-1),

or a pharmaceutically acceptable salt thereof, wherein:

Z^1 is absent, -O-, or -NH-;

each X^1, X^2, X^3, X^4, X^5, X^6, X^7, and X^8 is independently O, S, or NR^2;

each Y^1, Y^2, Y^3, Y^4, Y^5, Y^6, Y^7, and Y^8 is independently CH or N;

W^1 is deuterium, hydrogen, optionally substituted C_1-C_6 alkyl, (azaneylidene)methanediamine, (azaneylidene)-N,N,N',N'-tetramethylmethanediamine, $-C(O)-NR^{1A}R^{1B}$, $-NR^{1A}-C(O)-NR^{1A}R^{1B}$, $-Z_B-P(O)(OR^{1A})_2$, $-Z_B-(CH_2)_p-P(O)(OR^{1A})_2$, $-Z_B-(CH_2)_{p3}-O-P(O)(OR^{1A})_2$, wherein

Z_B is -N- or -O-;

p_3 is an integer from 1-10;

W^2 is an optionally substituted C_1-C_6 alkyl or $-C(O)-NR^{1A}R^{1B}$;

each R^1 is independently hydrogen, deuterium, halogen, amino, cyano, optionally C_1-C_{20} alkyl, or C_1-C_{20} heteroalkyl, or $-NHC(O)R^{1A}$; or

two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3 to 6-membered carbocyclic ring or 3 to 6-membered heterocyclic ring;

each R^2 is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino;

each R^{1A} is independently hydrogen, deuterium, or optionally substituted C_1-C_{20} alkyl;

each R^{1B} is independently hydrogen, deuterium, optionally substituted 5-membered heteroaryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_2-C_{10} heteroalkyl, or $(AA)_{p2}$, wherein each AA is an amino acid;

p_2 is an integer from 1-10;

j_1 is 0 or 1;

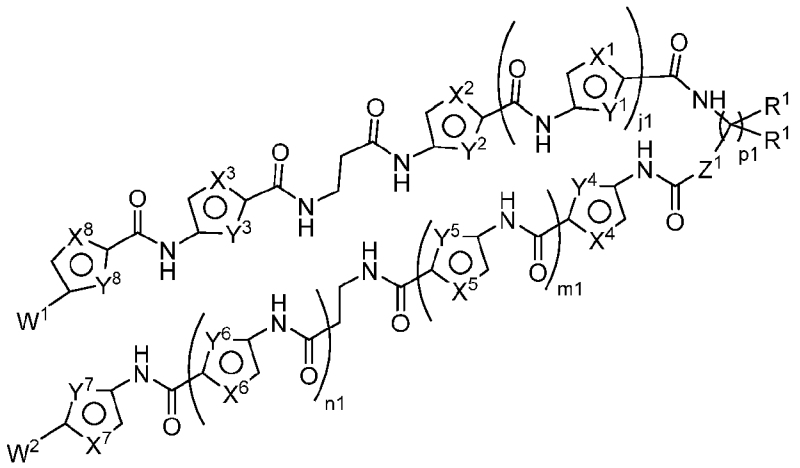
n_0 is 0 or 1;

m_1 and n_1 are each independently an integer from 0-3;

p_1 is 2 or 3, provided that when Z^1 is O or NH then p_1 is 2, and when Z^1 is absent then p_1 is 3; and

wherein one of W^1 or W^2 is attached to the oligomeric backbone.

25. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-2), or a pharmaceutically acceptable salt thereof:

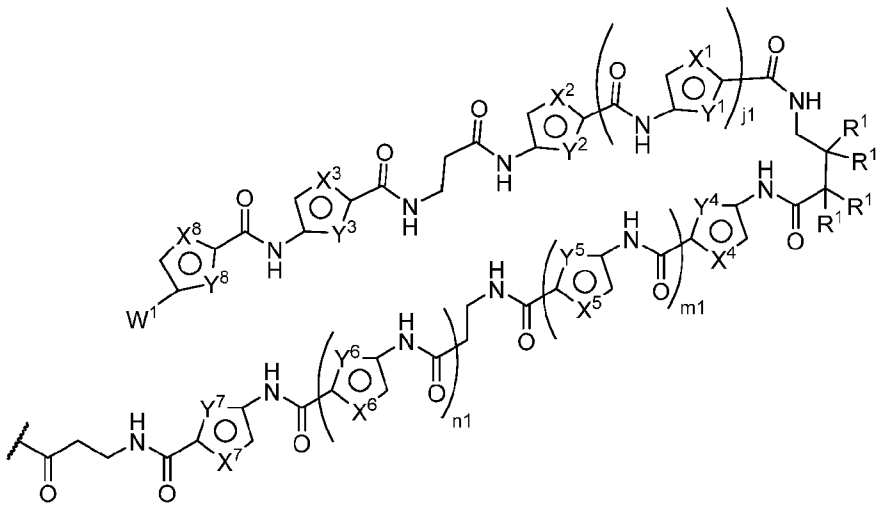


Formula (A-2).

26. The transcription modulator molecule of claim 24 or 25, or a pharmaceutically acceptable salt thereof, wherein each R^1 is hydrogen.

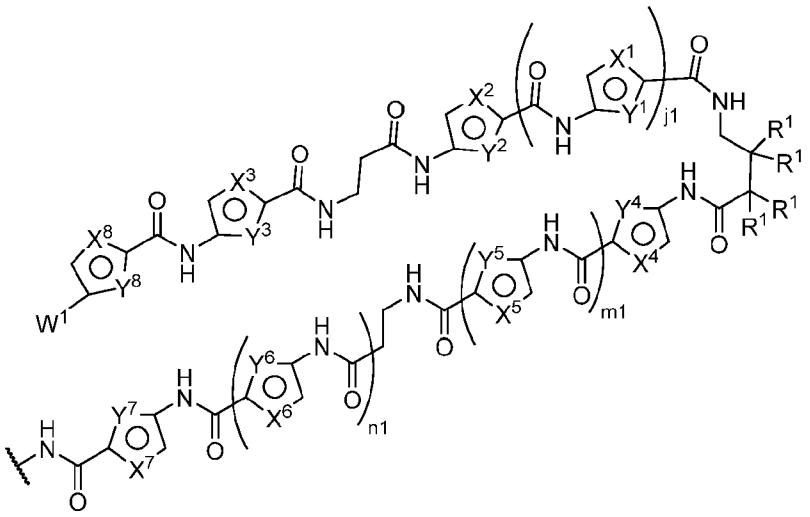
27. The transcription modulator molecule of any one of claims 24 to 26, or a pharmaceutically acceptable salt thereof, wherein W^2 is $-C(O)NR^{1A}R^{1B}$, wherein R^{1A} is hydrogen and R^{1B} is alkyl optionally substituted with oxo (=O); and wherein W^2 is attached to the oligomeric backbone.

28. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-4), or a pharmaceutically acceptable salt thereof:



Formula (A-4).

29. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-6), or a pharmaceutically acceptable salt thereof:



Formula (A-6).

30. The transcription modulator molecule of any one of claims 24 to 29, or a pharmaceutically acceptable salt thereof, wherein m_1 is 0 or 1 and p_1 is 2.

31. The transcription modulator molecule of any one of claims 24 to 30, or a pharmaceutically acceptable salt thereof, wherein Y^2, Y^4, and Y^7 are each independently N; and Y^1 and Y^3 are each independently CH.

32. The transcription modulator molecule of any one of claims 24 to 31, or a pharmaceutically acceptable salt thereof, wherein n_1 is 0.

33. The transcription modulator molecule of any one of claims 24 to 31, or a pharmaceutically acceptable salt thereof, wherein n_1 is 1.

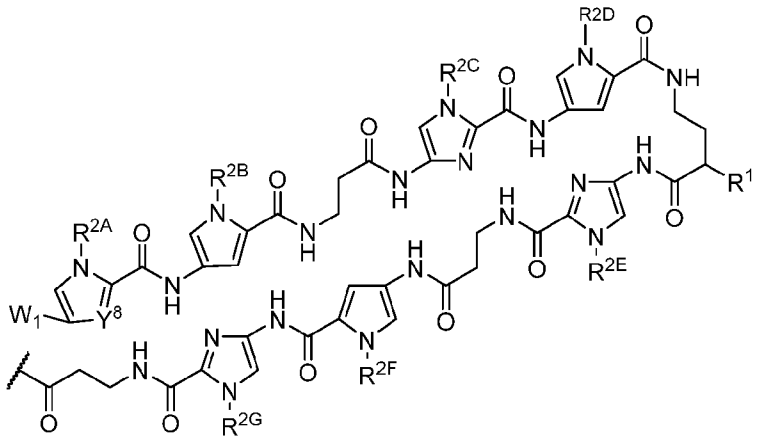
34. The transcription modulator molecule of any one of claims 24 to 33, or a pharmaceutically acceptable salt thereof, wherein j_1 is 0.

35. The transcription modulator molecule of any one of claims 24 to 33, or a pharmaceutically acceptable salt thereof, wherein j_1 is 1.

36. The transcription modulator molecule of any one of claims 24 to 31 or 33 to 35, or a pharmaceutically acceptable salt thereof, wherein Y^6 is CH.

37. The transcription modulator molecule of claim 24, or a pharmaceutically acceptable salt thereof, wherein X^1, X^2, X^3, X^4, X^5, X^6, and X^7 are each independently $-NR^2$.

38. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-8):



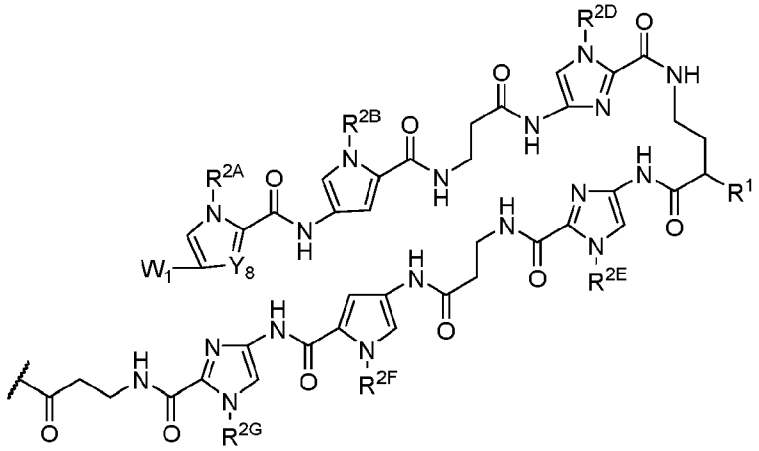
Formula (A-8),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

39. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-9):



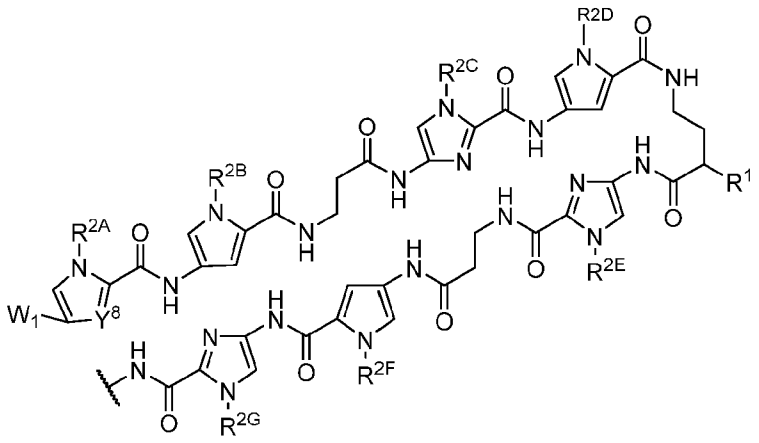
Formula (A-9),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

R^{2A}, R^{2B}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

40. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-10):



Formula (A-10),

or a pharmaceutically acceptable salt thereof, wherein:

Y^8 is CH or N; and

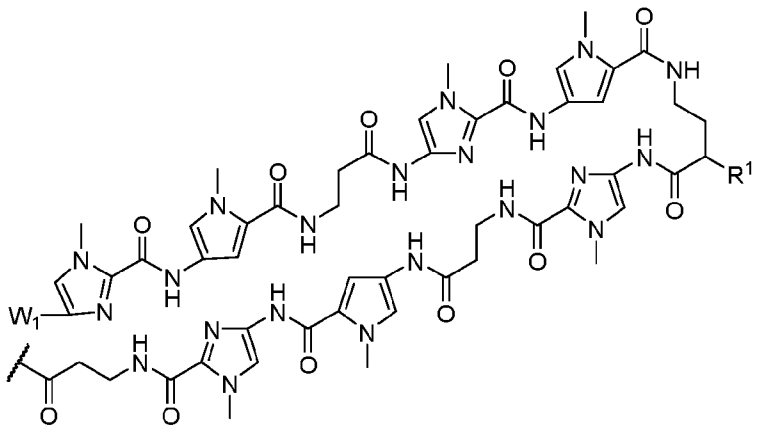
R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino.

41. The transcription modulator molecule of any one of claims 24 to 40, or a pharmaceutically acceptable salt thereof, wherein Y^8 is N.

42. The transcription modulator molecule of any one of claims 38 to 40, or a pharmaceutically acceptable salt thereof, wherein R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently hydrogen, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} heteroalkyl.

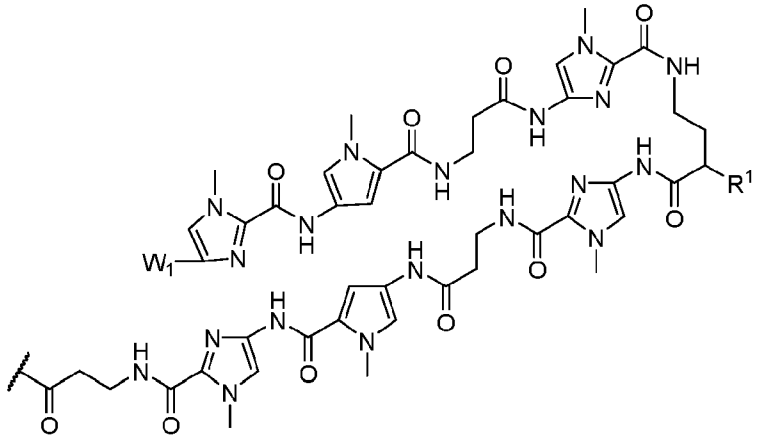
43. The transcription modulator molecule of claim 42, or a pharmaceutically acceptable salt thereof, wherein R^{2A}, R^{2B}, R^{2C}, R^{2D}, R^{2E}, R^{2F}, and R^{2G} are each independently an optionally substituted C_1-C_{20} alkyl.

44. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-11), or a pharmaceutically acceptable salt thereof:



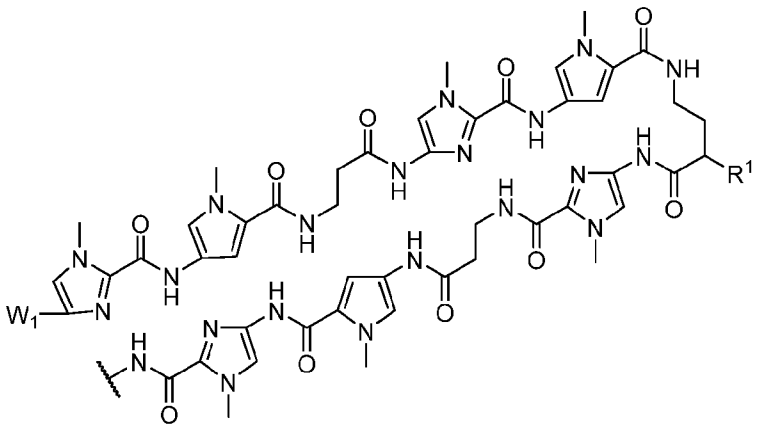
Formula (A-11).

45. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-12), or a pharmaceutically acceptable salt thereof:



Formula (A-12).

46. The transcription modulator molecule of claim 24, wherein the DNA-binding moiety comprises the structure of Formula (A-13), or a pharmaceutically acceptable salt thereof:



Formula (A-13).

47. The transcription modulator molecule of any one of claims 24 to 46, or a pharmaceutically acceptable salt thereof, wherein W^1 is hydrogen.

48. The transcription modulator molecule of any one of claims 24 to 47, or a pharmaceutically acceptable salt thereof, wherein the first terminus is capable of binding the DNA with an affinity of less than 500 nM.

49. The transcription modulator molecule of any one of claims 24 to 48, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone is a linker having a length of less than about 50 Angstroms.

50. The transcription modulator molecule of any one of claims 24 to 48, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone is a linker having a length of about 10 to 60 Angstroms.

51. The transcription modulator molecule of any one of claims 24 to 50, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone comprises a multimer having 2 to 50 spacing moieties, wherein

each spacing moiety is independently selected from the group consisting of -(($CR^{3a}R^{3b}$)$_x$-O)$_y$-, -(($CR^{3a}R^{3b}$)$_x$-NR^{4a})$_y$-, -(($CR^{3a}R^{3b}$)$_x$-CH=CH-($CR^{3a}R^{3b}$)$_x$-O)$_y$-, optionally substituted C_1-C_{12} alkyl, optionally substituted C_2-C_{10} alkenyl, optionally substituted C_2-C_{10} alkynyl, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5 to 10-membered heteroarylene, optionally substituted 4 to 10-membered heterocycloalkylene, amino acid residue, -O-, -C(O)NR^{4a}-, -NR^{4a}C(O)-, -C(O)-, -NR^{1a}-, -C(O)O-, -S-, -S(O)-, -$S(O)_2$-, -$S(O)_2NR^{4a}$-, -$NR^{4a}S(O)_2$-, and -P(O)OH-, and any combinations thereof; wherein

each x is independently 2-4;

each y is independently 1-10;

each R^{1a} is independently a hydrogen or optionally substituted C_1-C_6 alkyl;

each R^{3a} and R^{3b} is independently selected from hydrogen, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted amino, carboxyl, carboxyl ester, acyl, acyloxy, acyl amino, amino acyl, optionally substituted alkylamide, sulfonyl, optionally substituted thioalkoxy, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted cycloalkyl, and optionally substituted heterocyclyl; and

each R^{4a} is independently a hydrogen or an optionally substituted C_1-C_6 alkyl.

52. The transcription modulator molecule of claim 51, or a pharmaceutically acceptable salt thereof, wherein each spacing moiety is independently selected from the group consisting of -((CH_2)$_x$-O)$_y$-, -((CH_2)$_x$-NH)$_y$-, -O-, -C(O)NH-, -NH-, and any combinations thereof.

53. The transcription modulator molecule of any one of claims 24 to 50, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone comprises -NR^{4a}(CH_2CH_2O)$_y$(CH_2)$_x$- or -NR^{4a}-(CH_2)$_q$-C(O)NR^{4a}(CH_2CH_2O)$_y$(CH_2)$_x$-, wherein q is 2-10, x is 1-4, y is 1-50, and each R^{4a} is independently hydrogen or an optionally substituted C_1-C_6 alkyl.

54. The transcription modulator molecule of any one of claims 24 to 53, or a pharmaceutically acceptable salt thereof, wherein the second terminus comprises a bromodomain binding moiety.

55. The transcription modulator molecule of any one of claims 24 to 53, or a pharmaceutically acceptable salt thereof, wherein the second terminus comprises a moiety capable of binding to a bromodomain and extra terminal domain (BET) family member.

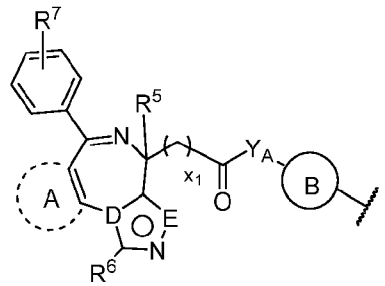
56. The transcription modulator molecule of claim 55, or a pharmaceutically acceptable salt thereof, wherein the BET family member is BRD2, BRD3, BRD4, or BRDT.

57. The transcription modulator molecule of any one of claims 24 to 54, or a pharmaceutically acceptable salt thereof, wherein the bromodomain is CBP/p300, PCAF (P300/CBP-Associated Factor), CECR2 (cat eye syndrome chromosome region candidate 2), BRPF (bromodomain and PHD finger-

containing protein), ATAD2/ATAD2B (chromatin remodeling proteins), TRIM24 (Tripartite motif-containing 24), BAZ2 (Bromodomain Adjacent to Zinc finger), or TAF1 (TBP associated factors).

58. The transcription modulator molecule of any one of claims 24 to 54, or a pharmaceutically acceptable salt thereof, wherein the second terminus has a triazolodiazepine structure.

59. The transcription modulator of any one of claims 24 to 54, wherein the second terminus comprises the structure of Formula (2-A):



Formula (2-A),

or a pharmaceutically acceptable salt thereof, wherein:

Ring A is an optionally substituted aryl or optionally substituted 5 to 6-membered heteroaryl;

Ring B is absent or an optionally substituted 6-membered monocyclic aryl or heteroaryl;

D is C or N;

E is O or N;

Y_A is -NH- or -O-;

R^5 is hydrogen, deuterium, or C_1-C_6 alkyl;

R^6 is selected from hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

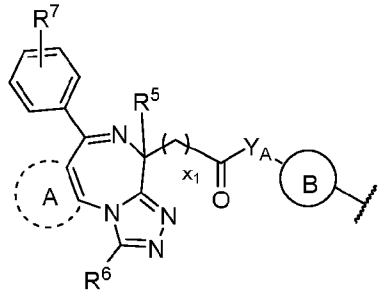
R^7 is selected from hydrogen, deuterium, halogen, -NO_2, -CN, optionally substituted aryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

or R^7 is -$NR^{7A}R^{7B}$, wherein

R^{7A} and R^{7B} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl; and

x_1 is an integer from 1-6.

60. The transcription modulator molecule of claim 59, wherein the second terminus comprises the structure of Formula (2-B):



Formula (2-B),

or a pharmaceutically acceptable salt thereof, wherein:

Ring A is an optionally substituted aryl or optionally substituted 5 to 6-membered heteroaryl;

Ring B is absent or an optionally substituted 6-membered monocyclic aryl or heteroaryl;

Y_A is -NH- or -O-;

R^5 is hydrogen, deuterium, or C_1-C_6 alkyl;

R^6 is selected from hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

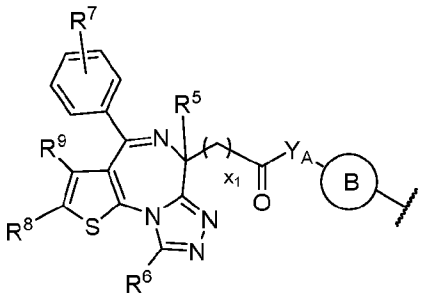
R^7 is selected from hydrogen, deuterium, halogen, -NO_2, -CN, optionally substituted aryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

or R^7 is -$NR^{7A}R^{7B}$, wherein

R^{7A} and R^{7B} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl; and

x_1 is an integer from 1-6.

61. The transcription modulator molecule of claim 59, wherein the second terminus comprises the structure of Formula (2-C):

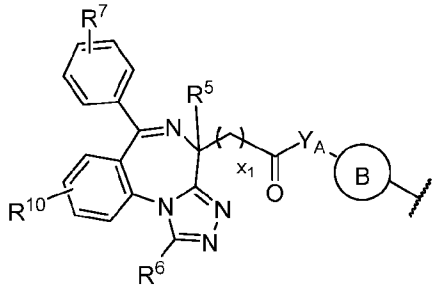


Formula (2-C),

or a pharmaceutically acceptable salt thereof, wherein:

R^8 and R^9 are each independently selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

62. The transcription modulator molecule of claim 59, wherein the second terminus comprises the structure of Formula (2-D), or a pharmaceutically acceptable salt thereof:



Formula (2-D),

or a pharmaceutically acceptable salt thereof, wherein:

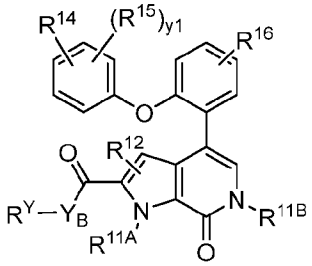
R^{10} is selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

63. The transcription modulator molecule of any one of claims 59 to 62, or a pharmaceutically acceptable salt thereof, wherein R^5 is hydrogen.

64. The transcription modulator molecule of any one of claims 59 to 63, or a pharmaceutically acceptable salt thereof, wherein Y_A is NH and x_1 is 1.

65. The transcription modulator molecule of any one of claims 59 to 64, or a pharmaceutically acceptable salt thereof, wherein Ring B is an optionally substituted aryl.

66. The transcription modulator molecule of any of claims 24 to 54, wherein the second terminus comprises the structure of Formula (3-A):



Formula (3-A),

or a pharmaceutically acceptable salt thereof, wherein,

Y_B is $-CH_2NH-$, $-CH_2O-$, -NH-, or -O-;

R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, halogen, -OH, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{14} and R^{15} are each independently hydrogen, deuterium, halogen, -CN, $-NO_2$, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is $-NR^AR^B$;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, C_1-C_6 hydroxyalkyl, $-SO_2R^A$, or $-NHSO_2R^A$;

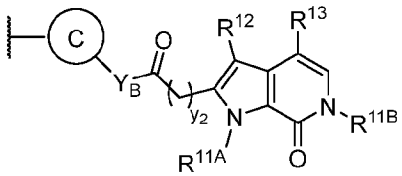
R^Y is hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted 5-6-membered monocyclic aryl or heteroaryl;

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is 1-3; and

wherein the attachment to the linker is at either R^{14} or at R^Y.

67. The transcription modulator molecule of any of claims 24 to 54, wherein the second terminus comprises the structure of Formula (3-B):



Formula (3-B),

or a pharmaceutically acceptable salt thereof, wherein,

Ring C is absent, optionally substituted 5 to 6-membered monocyclic aryl or heteroaryl or 4 to8-membered heterocycle;

Y_B is -CH_2NH-, -CH_2OH-, -NH- or -O-;

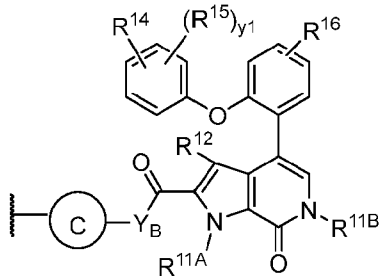
R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, $C(O)R^A$, or $C(O)NR^AR^B$; wherein each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl;

R^{13} is hydrogen, substituted aryl, substituted heteroaryl, or substituted oxydibenzene; and

y_2 is an integer from 0-2.

68. The transcription modulator molecule of claim 66, wherein the second terminus comprises the structure of Formula (3-C):



Formula (3-C),

or a pharmaceutically acceptable salt thereof, wherein:

Ring C is absent, optionally substituted 5 to 6-membered monocyclic aryl or heteroaryl or 4 to 8-membered heterocycle;

Y_B is -CH_2NH-, -CH_2O-, -NH-, or -O-;

R^{14} and R^{15} are each independently hydrogen, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is -NR^AR^B, wherein

each R^A and R^B is independently hydrogen, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl;

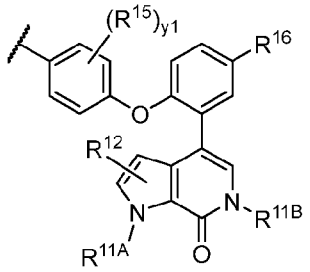
R^{16} is an substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, optionally substituted C_1-C_6 hydroxyalkyl, $-SO_2R^A$, or $-NHSO_2R^A$; and

y_1 is an integer from 1-3.

69. The transcription modulator molecule of any one of claims 66 to 68, or a pharmaceutically acceptable salt thereof, wherein Y_B is NH.

70. The transcription modulator molecule of any one of claims 66 to 69, or a pharmaceutically acceptable salt thereof, wherein Ring C is absent or optionally substituted aryl.

71. The transcription modulator molecule of claim 66, wherein the second terminus comprises the structure of Formula (3-D):



Formula (3-D),

or a pharmaceutically acceptable salt thereof, wherein:

R^{11A} and R^{11B} are each independently hydrogen or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen or an optionally substituted C_1-C_6 alkyl;

each R^{15} is independently hydrogen, halogen, -CN, $-NO_2$, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{16} is an optionally substituted optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, optionally substituted C_1-$_6$ hydroxyalkyl, $-SO_2R^A$, or $-NHSO_2R^A$; wherein

R^A is hydrogen, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is an integer from 1-3.

72. The transcription modulator molecule of any one of claims 66 to 71, or a pharmaceutically acceptable salt thereof, wherein R^{11A} and R^{11B} are each independently C_1-C_6 alkyl.

73. The transcription modulator molecule of any one of claims 24 to 72, wherein the molecule or a pharmaceutically acceptable salt thereof, is described in Table 4.

74. A pharmaceutical composition comprising a transcription modulator molecule of any one of claims 24 to 73, or a pharmaceutically acceptable salt thereof; and a pharmaceutically acceptable excipient.

75. The pharmaceutical composition of claim 74, wherein the composition is formulated for administration by intravenous administration.

76. A method of treating Huntington's disease (HD) in a patient in need thereof, the method comprising administering to the patient a transcriptional modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein

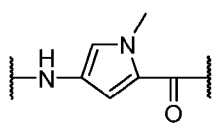
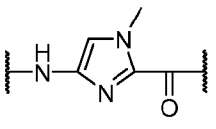
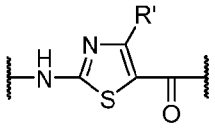
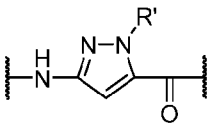
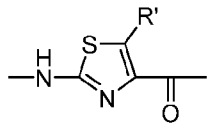
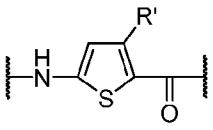
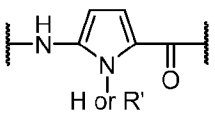
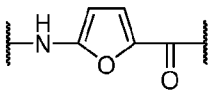
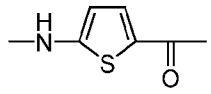
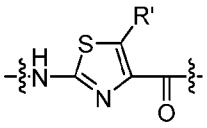
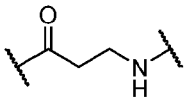
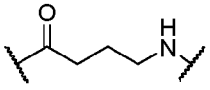
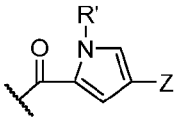
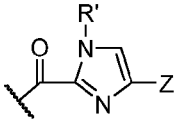
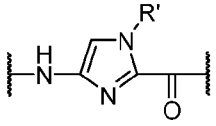
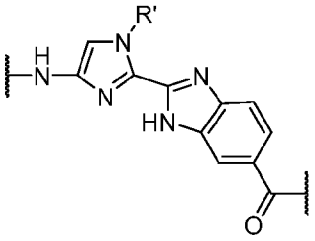
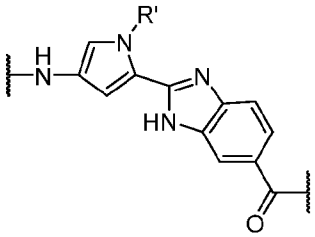
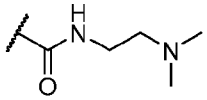
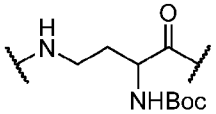
(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

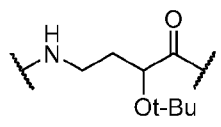
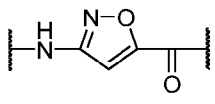
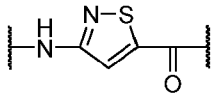
(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the expanded nucleotide repeat; and

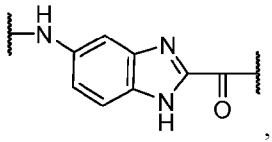
(c) the oligomeric backbone links the first terminus and the second terminus; and

wherein the DNA-binding moiety comprises -NH-Q-C(O)-, wherein Q is an optionally substituted C_6-C_{10} arylene, optionally substituted 4 to 10-membered heterocyclene, optionally substituted 5 to 10-membered heteroarylene group, or an optionally substituted alkylene group.

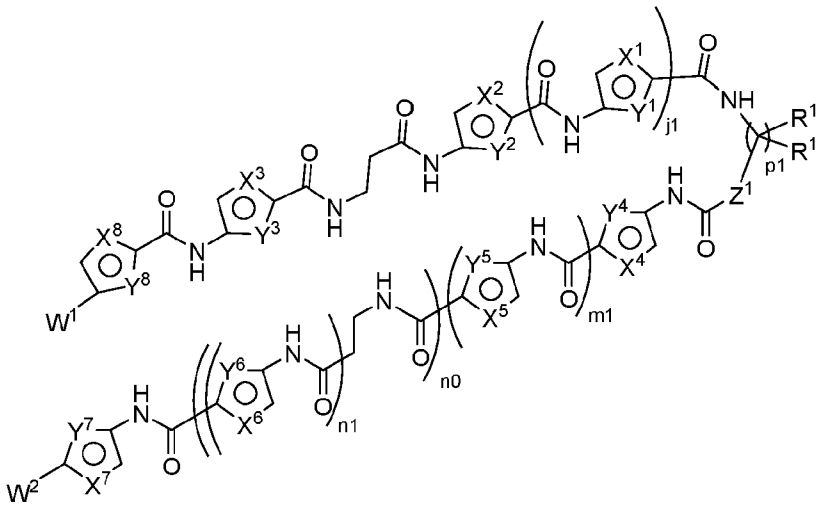
77. The method of claim 76, wherein the DNA-binding moiety comprises a polyamide of one or more of the following subunits selected from:


(Py),

(Im),

(Th),

(Pz),

(Nt),

(Tn),

(Nh),

(Fr),

(Tp),

(iNt),

(beta or β),

(gAB),

(PyT),

(ImT),

(iIm),


(ImBi),

(PyBi),

(Dp),

,

,

,

, -NH-benzopyrazinylene-C(O)-, -NH-phenylene-C(O)-, -NH-pyridinylene-C(O)-, -NH-piperidinylene-C(O)-, -NH-pyrimidinylene-C(O)-, -NH-

anthracenylene-C(O)-, -NH-quinolinylene-C(O)-, and , wherein each R' is independently hydrogen, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino; and Z is H, NH_2, C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 alkyl-NH_2.

78. The method of claim 76, wherein the DNA-binding moiety comprises the structure of Formula (A-1):



Formula (A-1),

or a pharmaceutically acceptable salt thereof, wherein:

Z^1 is absent, -O-, or -NH-;

each X^1, X^2, X^3, X^4, X^5, X^6, X^7, and X^8 is independently O, S, or NR^2;

each Y^1, Y^2, Y^3, Y^4, Y^5, Y^6, Y^7, and Y^8 is independently CH or N;

W^1 is deuterium, hydrogen, optionally substituted C_1-C_6 alkyl, (azaneylidene)methanediamine, (azaneylidene)-N,N,N',N'-tetramethylmethanediamine, -C(O)-$NR^{1A}R^{1B}$, -NR^{1A}-C(O)-$NR^{1A}R^{1B}$, -Z_B-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_p$-$P(O)(OR^{1A})_2$, -Z_B-$(CH_2)_{p3}$-O-$P(O)(OR^{1A})_2$, wherein

Z_B is N or O;

p_3 is an integer from 1-10;

W^2 is an optionally substituted C_1-C_6 alkyl or -C(O)-$NR^{1A}R^{1B}$;

each R^1 is hydrogen, halogen, amino, cyano, optionally C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, or -$NHC(O)R^{1A}$; or

two R^1 on the same or on adjacent atoms combine together with the atom(s) to which they are attached to form an optionally substituted 3 to 6-membered carbocyclic ring or 3 to 6-membered heterocyclic ring;

each R^2 is independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino;

each R^{1A} is independently hydrogen, or optionally substituted C_1-C_{20} alkyl;

each R^{1B} is independently hydrogen, optionally substituted 5-membered heteroaryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_2-C_{10} heteroalkyl, or $(AA)_{p2}$, wherein

each AA is an amino acid;

p_2 is an integer from 1-10;

j_1 is 0 or 1;

n_0 is 0 or 1;

m_1 and n_1 are each independently an integer from 0-3;

p_1 is 2 or 3, provided that wherein when Z^1 is O or NH then p_1 is 2 or when Z^1 is absent then p_1 is 3; and

wherein one of W^1 or W^2 is attached to the oligomeric backbone.

79. The method of any one of claims 76 to 78, wherein the first terminus is capable of binding the DNA with an affinity of less than 500 nM.

80. The method of any one of claims 76 to 79, wherein the oligomeric backbone is a linker having a length of less than about 50 Angstroms.

81. The method of any one of claims 76 to 79, wherein the oligomeric backbone is a linker having a length of about 10 to 60 Angstroms.

82. The method of any one of claims 76 to 81, wherein the second terminus comprises a bromodomain binding moiety.

83. The method of any one of claims 76 to 81, wherein the second terminus comprises a moiety capable of binding to a bromodomain and extra terminal domain (BET) family member.

84. The method of claim 83, wherein the BET family member is BRD2, BRD3, BRD4, or BRDT.

85. The method of any one of claims 76 to 84, wherein the second terminus is CBP/p300, PCAF (P300/CBP-Associated Factor), CECR2 (cat eye syndrome chromosome region candidate 2), BRPF (bromodomain and PHD finger-containing protein), ATAD2/ATAD2B (chromatin remodeling proteins), TRIM24 (Tripartite motif-containing 24), BAZ2 (Bromodomain Adjacent to Zinc finger), or TAF1 (TBP associated factors).

86. The method of any one of claims 76 to 85, wherein the method reduces one or more symptoms of Huntington's disease.

87. The method of claim 86, wherein the one or more symptoms are selected from chorea, cognitive decline, abnormal libido, abnormal eye movement, abnormal sense of smell, aggression, agitation, anxiety, apathy, bradykinesia, bradyphrenia, clumsiness, delusions, depression, difficulty walking, disinhibition, dystonia, gait imbalance, muscle weakness, hallucinations, hostility, hypokinesia, irritability, memory impairment, myoclonus, obsessive-compulsive behavior, poor fine motor coordination, seizure, speech articulation difficulties, staring gaze, weight loss, abnormal cholesterol metabolism, abnormal cerebral white

matter, alcoholism, Babinski sign, caudate atrophy, cerebral atrophy, choking, clonus, degeneration of the striatum, excessive daytime sleepiness, impaired visuospatial constructive cognition, inability to walk, insomnia, mutism, oral-pharyngeal dysphagia, rigidity, suicidal ideation, cerebellar atrophy, dementia, gate ataxia, gliosis, hyperreflexia, neuronal loss, or personality changes.

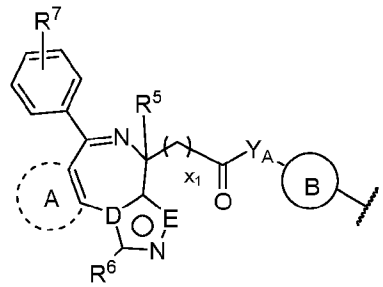
88. A transcription modulator molecule having a first terminus, a second terminus, and an oligomeric backbone, wherein:

(a) the first terminus comprises a DNA-binding moiety capable of binding a nucleotide repeat comprising CAG;

(b) the second terminus comprises a protein-binding moiety capable of binding to a regulatory molecule that modulates expression of a gene having the nucleotide repeat; and

(c) the oligomeric backbone links the first terminus and the second terminus; and

wherein second terminus comprises a structure of Formula (2-A):



Formula (2-A),

or a pharmaceutically acceptable salt thereof, wherein:

Ring A is an optionally substituted aryl or optionally substituted 5 to 6-membered heteroaryl;

Ring B is absent or an optionally substituted 6-membered monocyclic aryl or heteroaryl;

D is C or N;

E is O or N;

Y_A is -NH- or -O-;

R^5 is hydrogen, deuterium, or C_1-C_6 alkyl;

R^6 is selected from hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

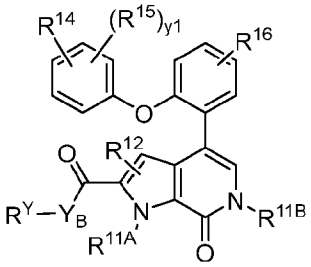
R^7 is selected from hydrogen, deuterium, halogen, -NO_2, -CN, optionally substituted aryl, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_6 haloalkyl, and optionally substituted C_1-C_6 hydroxyalkyl;

or R^7 is -$NR^{7A}R^{7B}$, wherein

R^{7A} and R^{7B} are each independently hydrogen, deuterium, optionally substituted C_1-C_{20} alkyl, or optionally substituted C_1-C_{20} heteroalkyl; and

x_1 is an integer from 1-6;

or wherein second terminus comprises a structure of Formula (3-A):



Formula (3-A),

or a pharmaceutically acceptable salt thereof, wherein,

Y_B is -CH_2NH-, -CH_2O-, -NH-, or -O-;

R^{11A} and R^{11B} are each independently hydrogen, deuterium, or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen, halogen, -OH, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{14} and R^{15} are each independently hydrogen, deuterium, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is -NR^AR^B;

R^{16} is optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, C_1-C_6 hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$;

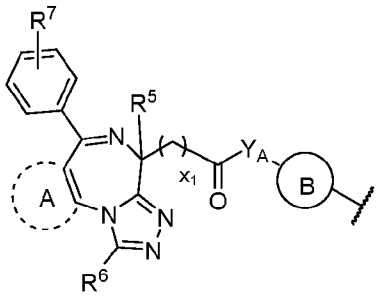
R^Y is hydrogen, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted 5-6-membered monocyclic aryl or heteroaryl;

each R^A and R^B is independently hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is 1-3; and

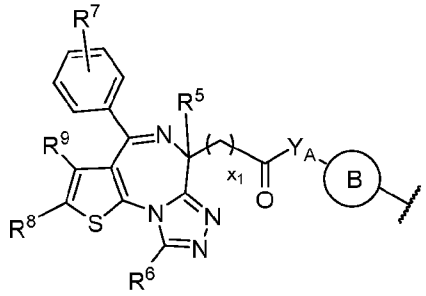
wherein the attachment to the linker is at either R^{14} or at R^Y.

89. The transcription modulator of claim 88, wherein the second terminus of Formula (2-A), comprises the structure of Formula (2-B), or a pharmaceutically acceptable salt thereof:



Formula (2-B).

90. The transcription modulator of claim 88, wherein the second terminus of Formula (2-A), comprises the structure of Formula (2-C):

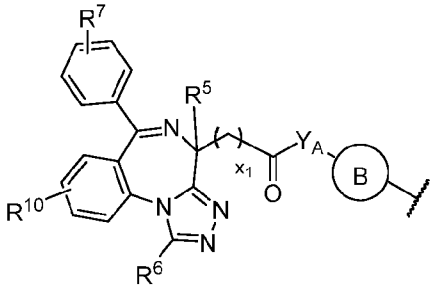


Formula (2-C),

or a pharmaceutically acceptable salt thereof, wherein:

R^8 and R^9 are each independently selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

91. The transcription modulator of claim 88, wherein the second terminus of Formula (2-A), comprises the structure of Formula (2-D):



Formula (2-D),

or a pharmaceutically acceptable salt thereof, wherein:

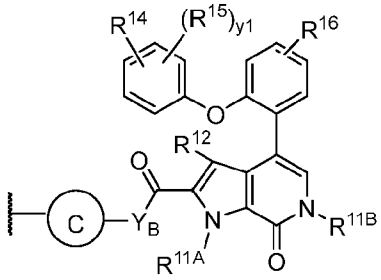
R^{10} is selected from hydrogen, deuterium, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl.

92. The transcription modulator molecule of any one of claims 88 to 91, or a pharmaceutically acceptable salt thereof, wherein R^5 is hydrogen.

93. The transcription modulator molecule of any one of claims 88 to 92, or a pharmaceutically acceptable salt thereof, wherein Y_A is NH and x_1 is 1.

94. The transcription modulator molecule of any one of claims 88 to 93, or a pharmaceutically acceptable salt thereof, wherein Ring B is an optionally substituted aryl.

95. The transcription modulator of claim 88, wherein the second terminus of Formula (3-A), comprises the structure of Formula (3-C):



Formula (3-C),

or a pharmaceutically acceptable salt thereof, wherein:

Ring C is absent, optionally substituted 5 to 6-membered monocyclic aryl or heteroaryl or 4 to 8-membered heterocycle;

Y_B is -NH-, -CH_2NH-, -CH_2O-, or -O-;

R^{14} and R^{15} are each independently hydrogen, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

or R^{14} is -NR^AR^B, wherein

each R^A and R^B is independently hydrogen, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl;

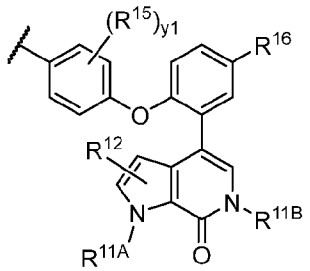
R^{16} is an substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, optionally substituted C_1-C_6 hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$; and

y_1 is an integer from 1-3.

96. The transcription modulator of claim 88 to 95, or a pharmaceutically acceptable salt thereof, wherein Y_B is NH.

97. The transcription modulator molecule of any one of claims 88, 95 or 65, or a pharmaceutically acceptable salt thereof, wherein Ring C is absent or optionally substituted aryl.

98. The transcription modulator of claim 88, wherein the second terminus of Formula (3-A), comprises the structure of Formula (3-D):



Formula (3-D),

or a pharmaceutically acceptable salt thereof, wherein:

R^{11A} and R^{11B} are each independently hydrogen or optionally substituted C_1-C_6 alkyl;

R^{12} is hydrogen or an optionally substituted C_1-C_6 alkyl;

each R^{15} is independently hydrogen, halogen, -CN, -NO_2, optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 haloalkyl, or optionally substituted C_1-C_6 hydroxyalkyl;

R^{16} is an optionally substituted optionally substituted C_1-C_6 alkyl, optionally substituted C_1-C_6 heteroalkyl, optionally substituted C_2-C_6 alkenyl, optionally substituted C_2-C_6 alkynyl, optionally substituted C_1-$_6$ hydroxyalkyl, -SO_2R^A, or -$NHSO_2R^A$; wherein

R^A is hydrogen, optionally substituted C_1-C_6 alkyl, or optionally substituted C_1-C_6 heteroalkyl; and

y_1 is an integer from 1-3.

99. The transcription modulator molecule of any one of claims 88 or 95 to 98, or a pharmaceutically acceptable salt thereof, wherein R^{11A} and R^{11B} are each independently C_1-C_6 alkyl.

100. The transcription modulator molecule of any one of claims 88 to 99, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone is a linker having a length of less than about 50 Angstroms.

101. The transcription modulator molecule of any one of claims 88 to 99, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone is a linker having a length of about 10 to 60 Angstroms.

102. The transcription modulator molecule of any one of claims 88 to 99, or a pharmaceutically acceptable salt thereof, wherein the oligomeric backbone comprises a multimer having 2 to 50 spacing moieties, wherein

each spacing moiety is independently selected from the group consisting of -$((CR^{3a}R^{3b})_x$-$O)_y$-, -$((CR^{3a}R^{3b})_x$-$NR^{4a})_y$-, -$((CR^{3a}R^{3b})_x$-CH=CH-$(CR^{3a}R^{3b})_x$-$O)_y$-, optionally substituted C_1-C_{12} alkyl, optionally substituted C_2-C_{10} alkenyl, optionally substituted C_2-C_{10} alkynyl, optionally substituted C_6-C_{10} arylene, optionally substituted C_3-C_7 cycloalkylene, optionally substituted 5 to 10-membered heteroarylene, optionally substituted 4 to 10-membered heterocycloalkylene, amino acid residue, -O-, -C(O)NR^{4a}-, -NR^{4a}C(O)-, -C(O)-, -NR^{1a}-, -C(O)O-, -S-, -S(O)-, -$S(O)_2$-, -$S(O)_2NR^{4a}$-, -$NR^{4a}S(O)_2$-, and -P(O)OH-, and any combinations thereof; wherein

each x is independently 2-4;

each y is independently 1-10;

each R^{1a} is independently a hydrogen or optionally substituted C_1-C_6 alkyl;

each R^{3a} and R^{3b} is independently selected from hydrogen, optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted alkoxy, optionally substituted amino, carboxyl, carboxyl ester, acyl, acyloxy, acyl amino, amino acyl, optionally substituted alkylamide, sulfonyl, optionally substituted thioalkoxy, optionally substituted aryl, optionally substituted heteroaryl, optionally substituted cycloalkyl, and optionally substituted heterocyclyl; and

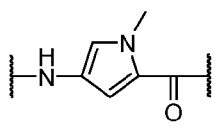
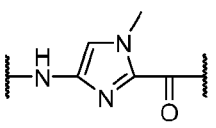
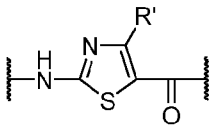
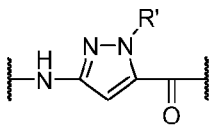
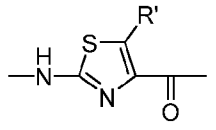
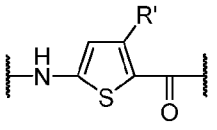
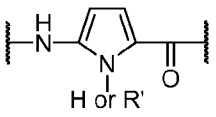
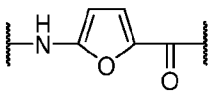
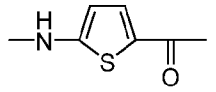
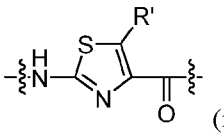
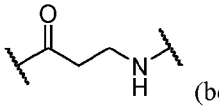
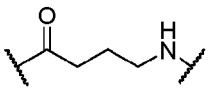
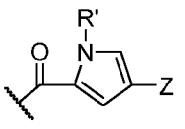
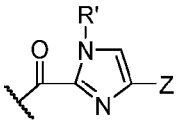
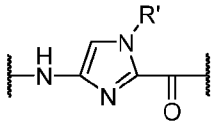
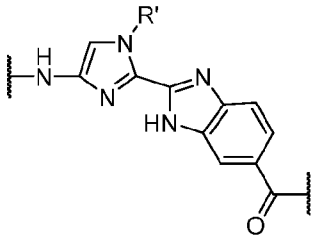
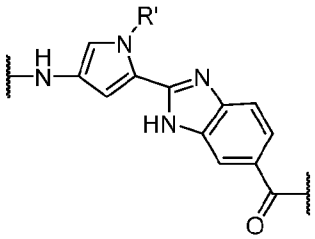
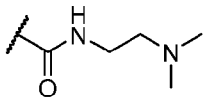
each R^{4a} is independently a hydrogen or an optionally substituted C_1-C_6 alkyl.

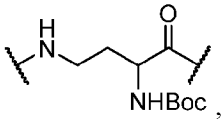
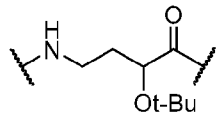
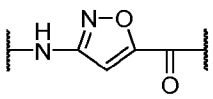
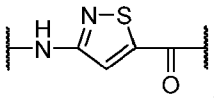
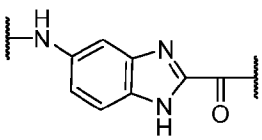
103. The transcription modulator molecule of claim 102, or a pharmaceutically acceptable salt thereof, wherein each spacing moiety is independently selected from the group consisting of -$((CH_2)_x$-$O)_y$-, -$((CH_2)_x$-$NH)_y$-, -O-, -C(O)NH-, -NH-, and any combinations thereof.

104. The transcription modulator molecule of any one of claims 88 to 103, or a pharmaceutically acceptable salt thereof, wherein the polyamide is selected from a linear polyamide, a hairpin polyamide, a H-pin polyamide, a U-pin polyamide, an overlapped polyamide, a slipped polyamide, a cyclic polyamide, a tandem polyamide, and an extended polyamide

105. The transcription modulator molecule of claim 104, or a pharmaceutically acceptable salt thereof, wherein the polyamide is a linear polyamide.

106. The transcription modulator molecule of any one of claims 88 to 105, or a pharmaceutically acceptable salt thereof, wherein the DNA-binding moiety comprises a polyamide of one or more of the following subunits selected from:


(Py),

(Im),

(Th),

(Pz),

(Nt),

(Tn),

(Nh),

(Fr),

(Tp),

(iNt),

(beta or β),

(gAB),

(PyT),

(ImT),

(iIm),


(ImBi),

(PyBi),

(Dp),

,

,

,

, -NH-benzopyrazinylene-C(O)-, -NH-phenylene-C(O)-, -NH-pyridinylene-C(O)-, -NH-piperidinylene-C(O)-, -NH-pyrimidinylene-C(O)-, -NH-anthracenylene-C(O)-, -NH-quinolinylene-C(O)-, and

, wherein each R' is

independently hydrogen, optionally substituted C_1-C_{20} alkyl, optionally substituted C_1-C_{20} heteroalkyl, optionally substituted C_1-C_{20} haloalkyl, or optionally substituted C_1-C_{20} alkylamino; and Z is H, NH_2, C_1-C_6 alkyl, C_1-C_6 haloalkyl, or C_1-C_6 alkyl-NH_2.