

COMPUESTOS ANTISENTIDO Y MÉTODOS PARA DIRIGIRSE A REPETICIONES CUG

REFERENCIA CRUZADA A LAS SOLICITUDES RELACIONADAS

5 Esta solicitud reivindica el beneficio de los números de serie de solicitud provisional de los Estados Unidos: 63/213,900, presentada el 23 de junio de 2021; 63/239,847, presentada el 1 de septiembre de 2021; 63/290,892, presentada el 17 de diciembre de 2021; 63/305,071, presentada el 31 de enero de 2022; 63/314,369, presentada el 26 de febrero de 2022; 63/316,634, presentada el 4 de marzo de 2022; 63/317,856, presentada el 8 de marzo de 2022; 63/326,201, presentada el 31 de marzo de 2022;
10 63/327,179, presentada el 4 de abril de 2022; 63/339,250, presentada el 6 de mayo de 2022; 63/362,295, presentada el 31 de marzo de 2022; 63/239,671, presentada el 1 de septiembre de 2021; 63/290,960, presentada el 17 de diciembre de 2021; 63/298,565, presentada el 11 de enero de 2022; y 63/268,577, presentada el 25 de febrero de 2022.

15 CAMPO

La presente descripción se refiere a compuestos, composiciones y métodos para modular la actividad y/o los niveles de genes que incluyen repeticiones de nucleótidos expandidas, en particular repeticiones CTG·CUG expandidas. Los compuestos y las composiciones que contienen los mismos pueden usarse para tratar enfermedades asociadas con genes que incluyen una repetición de
20 nucleótidos expandida, en particular repeticiones CTG·CUG expandidas.

INTRODUCCIÓN

Varias enfermedades están asociadas con genes que tienen repeticiones de nucleótidos expandidas, es decir, un número mayor de repeticiones de nucleótidos que el que se observa en un fenotipo sano.
25 La repetición expandida puede provocar agregación y/o nucleación del transcrito que contiene la repetición expandida y/o provocar la nucleación de proteínas que se unen al transcrito que contiene la repetición expandida. Las repeticiones expandidas pueden dar como resultado que algunas proteínas, tales como las proteínas de procesamiento de pre-ARNm, se secuestren en la repetición, lo que inhibe por tanto que las proteínas realicen sus funciones normales, tales como el
30 procesamiento de transcritos de pre-ARNm de otros genes que no contienen la repetición expandida. Existen varias enfermedades asociadas con genes que tienen repeticiones de trinucleótidos CTG·CUG expandidas (CTG se refiere a la repetición en ADN y CUG se refiere a la repetición en

el ARN correspondiente que se produce después de la transcripción). Las enfermedades asociadas con genes que tienen repeticiones de trinucleótidos CTG·CUG expandidas incluyen, pero no se limitan a, distrofia miotónica tipo 1 (DM1), ataxia espinocerebelosa-8 (SCA8), enfermedad similar a Huntington-2 (HDL2) y distrofia endotelial corneal de Fuchs (FECD).

5 La distrofia miotónica tipo 1 (DM1), la causa más frecuente de distrofia muscular en adultos, que afecta a 1 de cada 8,500 individuos en todo el mundo, está asociada con un gen que tiene una repetición de trinucleótidos expandida (Lee y Cooper. (2009) “Pathogenic mechanisms of myotonic dystrophy”, Biochem Soc Trans. 37(06): 10.1042/BST0371281). La DM1 es un trastorno que afecta al músculo esquelético y liso, así como también al ojo, al corazón, al sistema endocrino y al sistema
10 nervioso central. La DM1 es provocada por la expansión anormal de la repetición del trinucleótido CTG en la región no codificante del gen que codifica la proteína cinasa de la distrofia miotónica (DMPK). La expansión de CTG se localiza dentro de una región correspondiente a la región no traducida 3' (3'-UTR) del ARNm de DMPK. Mientras que el gen *DMPK* en individuos sanos contiene entre 5 y 40 repeticiones de trinucleótidos CTG, los pacientes con DM1 tienen desde 50 y
15 hasta varios miles de repeticiones de trinucleótidos CTG. La expansión de repeticiones de trinucleótidos CTG da como resultado la desregulación global de la expresión génica en los individuos afectados, debido a la nucleación de algunas proteínas reguladoras de la unión al ARN en la expansión de CUG en la región no traducida 3' (3'-UTR), lo que hace que las proteínas de unión al ARN, tales como la proteína de tipo muscleblind (MBNL1-3) no puedan realizar su función celular
20 normal. Las proteínas de unión al ARN nucleadas no están disponibles para unirse y efectuar la traducción de otros transcritos de ARNm. Estos agregados de ARNm-proteína con CUG expandidos forman focos nucleares distintos. La actividad de factores de corte y empalme adicionales, tales como el miembro de la familia similar a CUGBP Elav 1 (CELF1), también se interrumpe, lo que conduce al corte y empalme erróneo de un gran número de transcritos génicos aguas abajo asociados
25 con los síntomas de la DM1. La gravedad de la enfermedad aumenta y la edad de inicio disminuye con un número creciente de repeticiones (Pettersson y otros, (2015) “Molecular mechanisms in DM1 – a focus on foci”. Nucleic Acids Res. 43(4):2,433-2,441).

Las repeticiones de trinucleótidos CUG en la región no traducida 3' del ARNm de DMPK forma estructuras de horquilla estables imperfectas que se acumulan en el núcleo celular en complejos
30 ribonucleares pequeños o inclusiones visibles microscópicamente, y perjudican la función de las proteínas implicadas en la transcripción, corte y empalme o exportación del ARN. Aunque los genes *DMPK* con repeticiones CUG se transcriben en ARNm, los transcritos mutantes se secuestran en el

núcleo como agregados (focos), lo que da como resultado una disminución en los niveles de ARNm de DMPK citoplasmáticos. Estas agregaciones conducen a la desregulación del corte y empalme alternativo de muchos transcritos diferentes debido al secuestro de dos proteínas de unión al ARN: MBNL1 (de tipo muscleblind 1) y CUGBP1 (proteína de unión a CUG 1), lo que da como resultado la pérdida de función de MBNL1 y la regulación positiva de CUGBP1 (Lee y Cooper. (2009) “Pathogenic mechanisms of myotonic dystrophy”, Biochem Soc Trans. 37(06): 10.1042/BST0371281). MBNL1 y el factor 1 tipo CUGBP-ETR-3 (CELF1) son reguladores del desarrollo de eventos de corte y empalme durante la transición fetal a adulto y la modificación de sus actividades en la DM1 conduce a la expresión de un patrón de corte y empalme fetal en tejidos adultos. El impacto aguas abajo de la disminución de MBNL1 y el aumento de los niveles de CELF1 incluye la alteración del corte y empalme alternativo, la traducción del ARNm y la degradación del ARNm en proteínas tales como transcritos de la troponina T cardíaca (cTNT), el receptor de insulina (INSR), el canal de iones cloruro específico del músculo (CLCN1) y la ATPasa de calcio 1 del retículo sarcoplasmático/endoplasmático (ATP2A1), además de MBNL1 (Konieczny y otros, (2017) “Myotonic dystrophy: candidate small molecule therapeutics”. Drug Discov Today. 22(11):1,740-174).

Los posibles enfoques terapéuticos para tratar la DM1, u otras enfermedades asociadas con las repeticiones CTG·CUG expandidas, incluyen el uso de compuestos que contienen oligonucleótidos terapéuticos. Sin embargo, un problema importante asociado con el uso de compuestos de oligonucleótidos en agentes terapéuticos es su capacidad limitada para obtener acceso al compartimiento intracelular cuando se administran sistémicamente. El suministro intracelular de compuestos de oligonucleótidos puede facilitarse mediante el uso de sistemas portadores tales como polímeros, liposomas catiónicos o mediante modificación química del constructo, por ejemplo, mediante la unión covalente de moléculas de colesterol. Sin embargo, la eficiencia del suministro intracelular de los compuestos de oligonucleótidos permanece baja. Aún se requieren sistemas de suministro mejorados para aumentar la potencia de estos compuestos.

Existe una necesidad insatisfecha de composiciones efectivas para suministrar compuestos de oligonucleótidos terapéuticos a compartimientos intracelulares para tratar enfermedades que son provocadas por repeticiones CTG·CUG expandidas, tales como la DM1.

SUMARIO

Los compuestos, composiciones y métodos para tratar una enfermedad asociada con una repetición CTG·CUG expandida se describen en la presente descripción. En modalidades, esta descripción se

refiere a compuestos que incluyen un compuesto antisentido (AC) y un péptido cíclico, tal como un péptido cíclico penetrante de células (cCPP). En modalidades, el AC se une a un gen o transcrito génico que comprende una repetición CUG expandida. En modalidades, el péptido cíclico facilita la localización intracelular del AC. Los compuestos pueden comprender un vehículo de escape endosómico (EEV). El EEV puede comprender el péptido cíclico y un péptido exocíclico.

En modalidades, en la presente descripción se proporciona un compuesto que comprende: (a) al menos un péptido cíclico y (b) un compuesto antisentido (AC) que es complementario a un nucleótido objetivo. En modalidades, el nucleótido objetivo comprende al menos una repetición CUG o CTG expandida. En modalidades, el nucleótido objetivo es un gen que comprende al menos una repetición CTG expandida. En modalidades, el nucleótido objetivo es el ARN que comprende al menos una repetición CUG expandida. En modalidades, el ARN que comprende al menos una repetición CUG expandida es una secuencia de pre-ARNm. En modalidades, la repetición CUG expandida corresponde a una repetición CTG expandida en un gen a partir del cual se transcribe el pre-ARNm. En modalidades, el compuesto antisentido se une a la repetición CTG expandida o a la repetición CUG expandida. En modalidades, el AC comprende 5-40 repeticiones CAG (por ejemplo, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 repeticiones). En modalidades, el AC comprende una secuencia de un nucleótido enumerado en la Tabla 2, la Tabla 10 o la Tabla 11. En modalidades, el AC comprende una secuencia de un nucleótido enumerado en la Tabla 2.

En modalidades, el AC comprende al menos un nucleótido o ácido nucleico modificado seleccionado de un nucleótido de fosforotioato (PS), un nucleótido de morfolino fosforodiamidato (PMO), un ácido nucleico bloqueado (LNA), un ácido nucleico peptídico (PNA), un nucleótido que comprende una cadena principal modificada con 2'-O-metilo (2'-OMe), un nucleótido 2'-O-metoxi-etilo (2'-MOE), un nucleótido de etilo restringido (cEt) 2',4', un ácido 2'-desoxi-2'-fluoro-beta-D-arabinonucleico (2'-F-ANA) y sus combinaciones. En modalidades, el AC comprende un nucleótido de PMO.

En modalidades, se proporcionan compuestos que incluyen un péptido cíclico que tiene de 6 a 12 aminoácidos, en donde al menos dos aminoácidos del péptido cíclico son aminoácidos cargados, al menos dos aminoácidos del péptido cíclico son aminoácidos aromáticos hidrófobos, y al menos dos aminoácidos del péptido cíclico son aminoácidos no aromáticos no cargados. En modalidades, el compuesto antisentido (AC) es complementario a al menos una porción de una repetición CUG expandida en una secuencia de ARNm objetivo. En modalidades, el AC es un nucleótido de

morfolino fosforodiamidato (PMO).

En modalidades, al menos dos aminoácidos cargados del péptido cíclico son arginina. En modalidades, al menos dos aminoácidos aromáticos hidrófobos del péptido cíclico son fenilalanina, naftilalanina (3-naft-2-il-alanina) o una de sus combinaciones. En modalidades, al menos dos aminoácidos no aromáticos no cargados del péptido cíclico son citrulina, glicina o una de sus combinaciones. En modalidades, el compuesto es un péptido cíclico que tiene de 6 a 12 aminoácidos, en donde dos aminoácidos del péptido cíclico son arginina, al menos dos aminoácidos son aminoácidos aromáticos hidrófobos seleccionados de fenilalanina, naftilalanina y sus combinaciones, y al menos dos aminoácidos son aminoácidos no aromáticos no cargados seleccionados de citrulina, glicina y sus combinaciones.

En modalidades, el compuesto comprende un vehículo de escape endosómico que comprende un péptido cíclico y un péptido exocíclico (EP). En modalidades, el EP se conjuga con un enlazador en un grupo amino. El enlazador puede ser un enlazador como se describe en la presente descripción. En modalidades, el EP se conjuga al péptido cíclico mediante el enlazador. En modalidades, el EP se conjuga al AC mediante el enlazador. En modalidades, el EP se conjuga al enlazador que conjuga el AC al péptido cíclico.

En modalidades, el EP comprende de 2 a 10 aminoácidos. En modalidades, el EP comprende de 4 a 8 residuos de aminoácidos. En modalidades, el EP comprende 1 o 2 aminoácidos que comprenden una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta. En modalidades, el EP comprende 1, 2, 3 o 4 residuos de lisina. En modalidades, el grupo amino en la cadena lateral de cada residuo de lisina se sustituye con un grupo trifluoroacetilo (-COCF₃), aliloxicarbonilo (Alloc), 1-(4,4-dimetil-2,6-dioxociclohexiliden)etilo (Dde) o (4,4-dimetil-2,6-dioxociclohex-1-iliden-3)-metilbutilo (ivDde). En modalidades, el EP comprende al menos 2 residuos de aminoácido con una cadena lateral hidrófoba. En modalidades, el residuo de aminoácido con una cadena lateral hidrófoba se selecciona de valina, prolina, alanina, leucina, isoleucina y metionina. En modalidades, el péptido exocíclico comprende una de las siguientes secuencias: PKKKRKV; KR; RR; KKK; KGK; KBK; KBR; KRK; KRR; RKK; RRR; KKKK; KKRK; KRKK; KRRK; RKKR; RRRR; KGKK; KKGK; KKKKK; KKKRK; KBKKBK; KKKRKV; PGKKRKV; PKGKRKV; PKKGRKV; PKKKGKV; PKKKRGV; o PKKKRKG. En modalidades, el péptido exocíclico consiste en una de las siguientes secuencias: PKKKRKV; KR; RR; KKK; KGK; KBK; KBR; KRK; KRR; RKK; RRR; KKKK; KKRK; KRKK; KRRK; RKKR; RRRR; KGKK; KKGK; KKKKK; KKKRK; KBKKBK; KKKRKV; PGKKRKV; PKGKRKV; PKKGRKV; PKKKGKV;

PKKKRGV; o PKKKRKG. En modalidades, el péptido exocíclico tiene la estructura: Ac-P-K-K-K-R-K-V-.

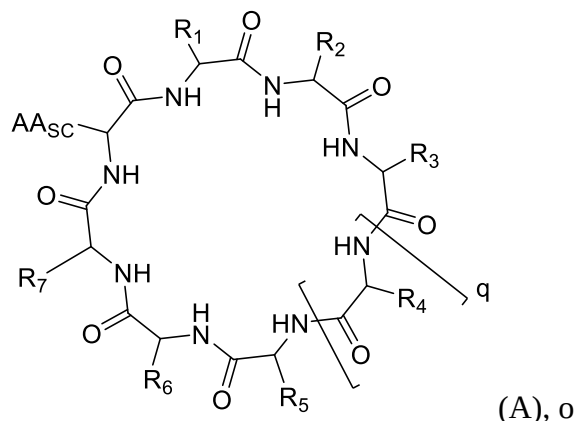
En modalidades, el péptido cíclico comprende de 4 a 12 aminoácidos. En modalidades, el péptido cíclico comprende de 6 a 12 aminoácidos. En modalidades, al menos dos aminoácidos del péptido cíclico son aminoácidos cargados, al menos dos aminoácidos del péptido cíclico son aminoácidos aromáticos hidrófobos, y al menos dos aminoácidos del péptido cíclico son aminoácidos no aromáticos no cargados. En modalidades, al menos dos aminoácidos cargados del péptido cíclico son arginina, al menos dos aminoácidos aromáticos hidrófobos del péptido cíclico son fenilalanina, naftilalanina o sus combinaciones, y al menos dos aminoácidos no aromáticos no cargados son citrulina, glicina o sus combinaciones.

En modalidades, el péptido cíclico tiene de 4 a 12 aminoácidos, en donde al menos dos aminoácidos son arginina y al menos dos aminoácidos comprenden una cadena lateral hidrófoba, siempre y cuando el péptido cíclico no sea un péptido cíclico que tenga una secuencia de SEQ ID NO: 89-117. En modalidades, el péptido cíclico no es un péptido cíclico que tiene una secuencia de SEQ ID NO: 89-117.

Secuencias de CPP y SEQ ID NO					
FΦRRRQ	89	RRFRΦRQ	99	FΦRRRRQK	109
FΦRRRC	90	FRRRRΦQ	100	FΦRRRRQC	110
FΦRRRU	91	rRFRΦRQ	101	fΦRrRrRQ	111
RRRΦFQ	92	RRΦFRRQ	102	FΦRRRRRQ	112
RRRRΦF	93	CRRRRFWQ	103	RRRRΦFDΩC	113
FΦRRRR	94	FfΦRrRrQ	104	FΦRRR	114
FφrRrRq	95	FFΦRRRRQ	105	FWRRR	115
FφrRrRQ	96	RFRFRΦRQ	106	RRRΦF	116
FΦRRRRQ	97	URRRRFWQ	107	RRRWF	117
fΦRrRrQ	98	CRRRRFWQ	108		

donde F es L-fenilalanina, f es D-fenilalanina, Φ es L-3-(2-naftil)-alanina, Φ es D-3-(2-naftil)-alanina, R es L-arginina, r es D-arginina, Q es L-glutamina, q es D-glutamina, C es L-cisteína, U es L-selenocisteína, W es L-triptófano, K es L-lisina, D es ácido L-aspartico y Ω es L-norleucina.

En modalidades, el péptido cíclico tiene la siguiente estructura:



una forma protonada de esta,

en donde:

R_1 , R_2 y R_3 son cada uno independientemente H o una cadena lateral aromática o heteroaromática de un aminoácido;

al menos uno de R_1 , R_2 y R_3 es una cadena lateral aromática o heteroaromática de un aminoácido;

R_4 , R_5 , R_6 , R_7 son independientemente H o una cadena lateral de aminoácido;

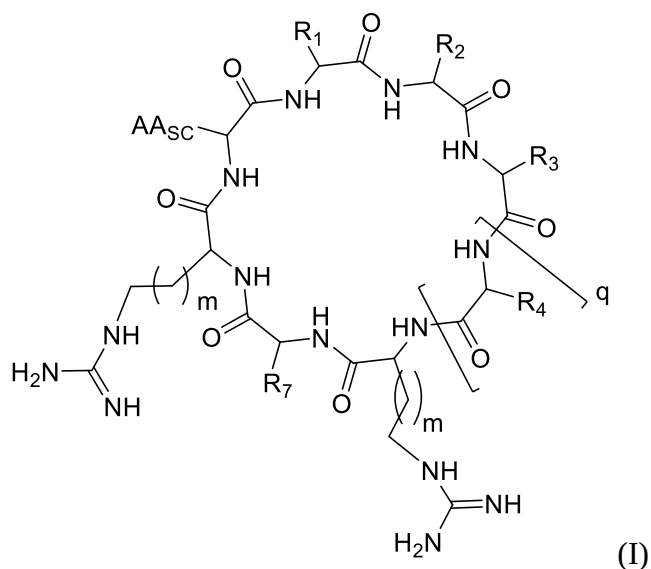
al menos uno de R_4 , R_5 , R_6 , R_7 es la cadena lateral del ácido 3-guanidino-2-aminopropiónico, ácido 4-guanidino-2-aminobutanoico, arginina, homoarginina, N-metilarginina, N,N-dimetilarginina, ácido 2,3-diaminopropiónico, ácido 2,4-diaminobutanoico, lisina, N-metillisina, N,N-dimetillisina, N-etillisina, N,N,N-trimetillisina, 4-guanidinofenilalanina, citrulina, N,N-dimetillisina, β -homoarginina, 3-(1-piperidinil)alanina;

AA_{SC} es una cadena lateral de aminoácido a la que se conjuga el compuesto antisentido; y

q es 1, 2, 3 o 4.

En modalidades, al menos uno de R_4 , R_5 , R_6 , R_7 es independientemente una cadena lateral no aromática no cargada de un aminoácido. En modalidades, al menos uno de R_4 , R_5 , R_6 , R_7 es independientemente H o una cadena lateral de la citrulina.

En modalidades, el péptido cíclico tiene la estructura de fórmula I:



o una forma protonada de esta,

en donde:

R₁, R₂ y R₃ son cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

al menos uno de R₁, R₂ y R₃ es una cadena lateral aromática o heteroaromática de un aminoácido;

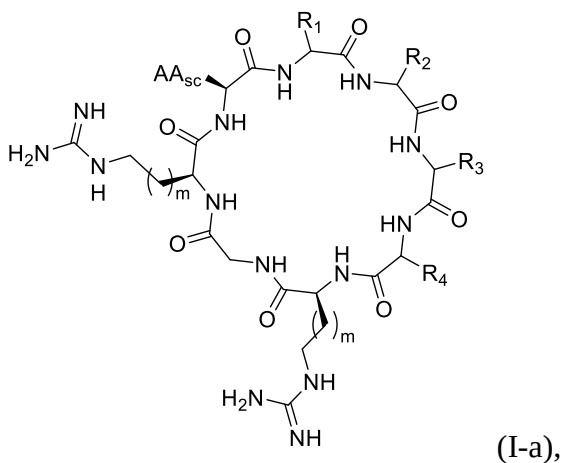
R₄ y R₇ son independientemente H o una cadena lateral de aminoácido;

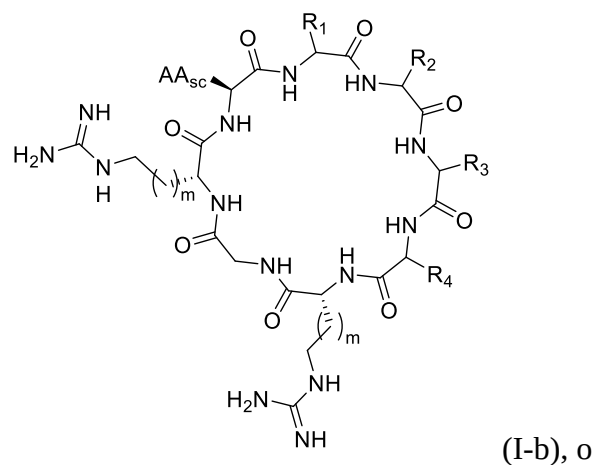
AA_{sc} es una cadena lateral de aminoácido a la que se conjuga el compuesto antisentido;

q es 1, 2, 3 o 4; y

cada m es independientemente un número entero de 0, 1, 2 o 3.

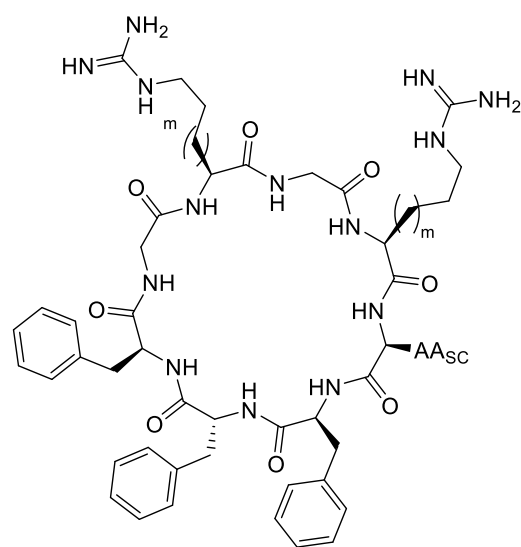
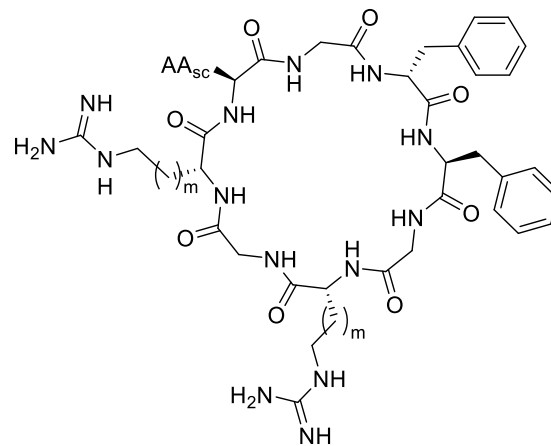
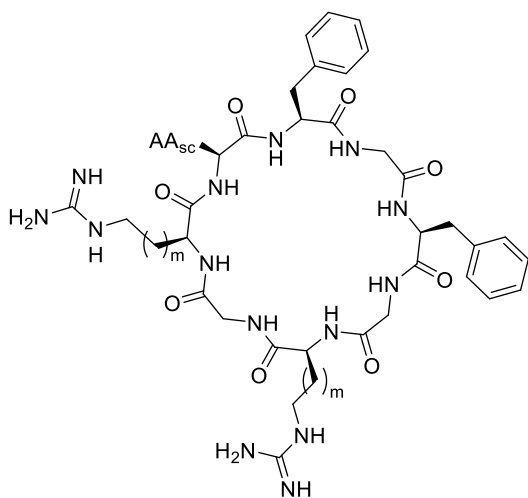
En modalidades, el péptido cíclico de fórmula (I) tiene una de las siguientes estructuras:

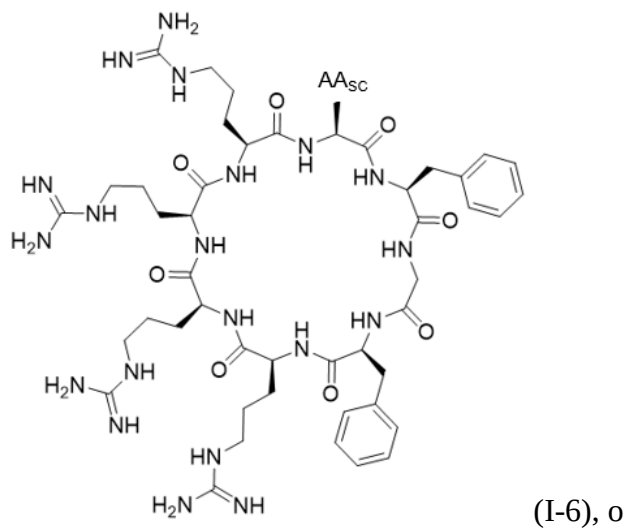
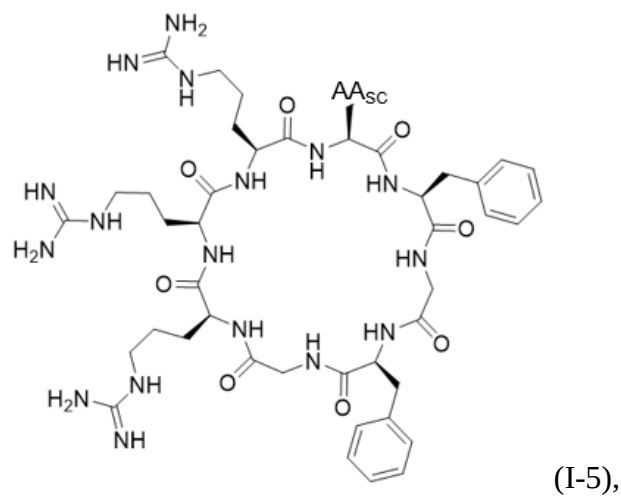
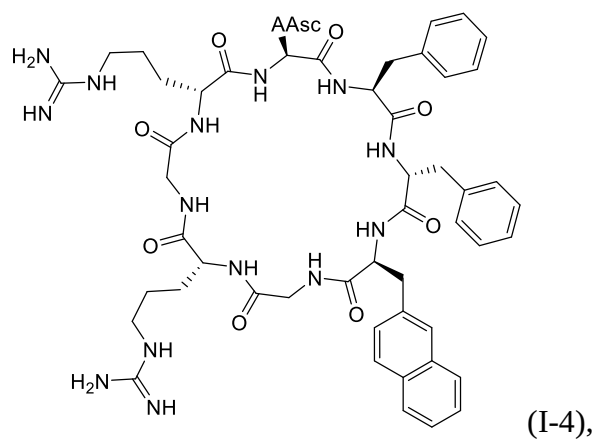




una forma protonada de estas.

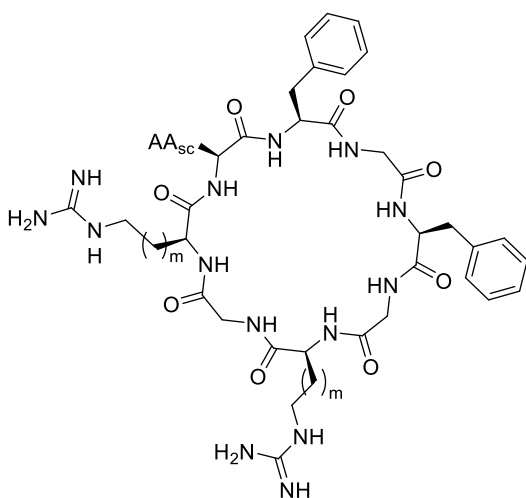
En modalidades, el péptido cíclico de fórmula (I) tiene una de las siguientes estructuras:



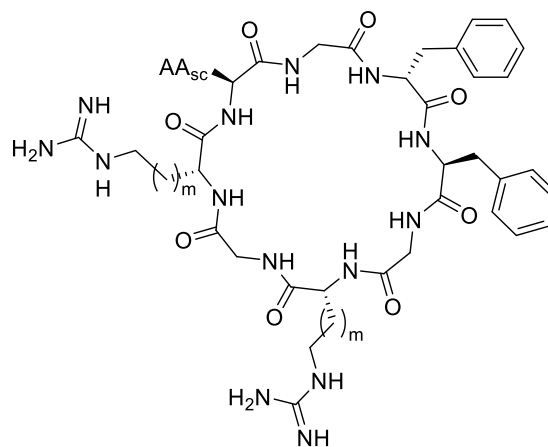


una forma protonada de estas.

En modalidades, el péptido cíclico de fórmula (I) tiene una de las siguientes estructuras:



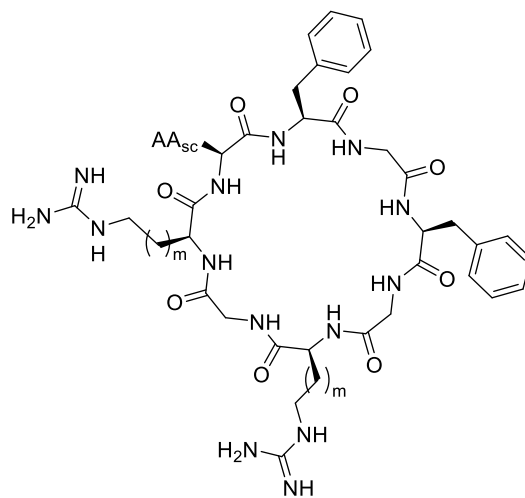
(I-1),



(I-2), o

una forma protonada de estas.

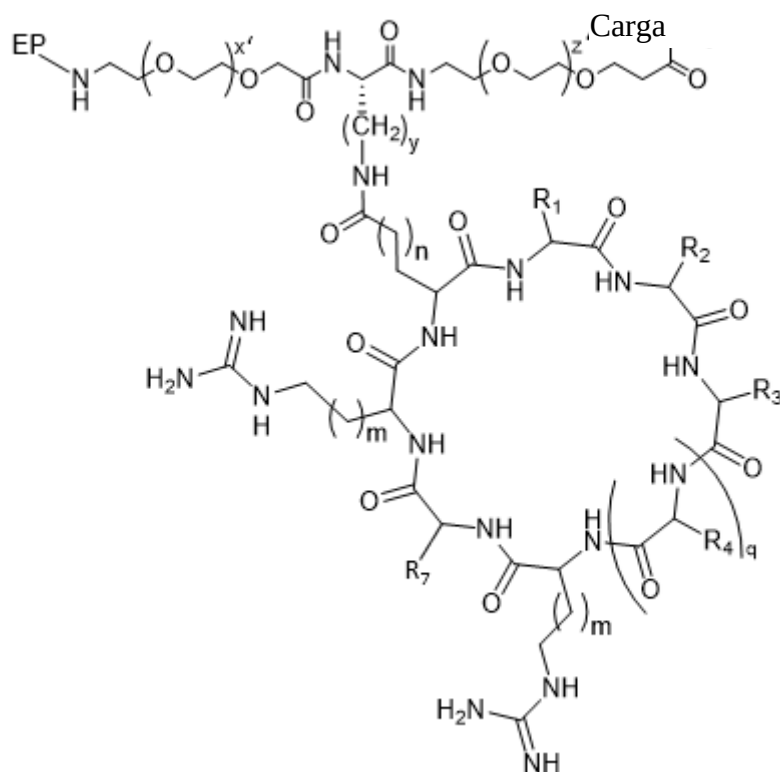
En modalidades, el péptido cíclico de fórmula (I) tiene la siguiente estructura:



(I-1), o

una forma protonada de estas.

En modalidades, el compuesto tiene una estructura de fórmula C:



(C), o

una forma protonada o sal de esta,

en donde:

R₁, R₂ y R₃ son cada uno independientemente H o una cadena lateral que comprende un grupo arilo o heteroarilo, en donde al menos uno de R₁, R₂ y R₃ es una cadena lateral que comprende un grupo arilo o heteroarilo;

R₄ y R₇ son independientemente H o una cadena lateral de aminoácido;

EP es el péptido exocíclico;

cada m es independientemente un número entero de 0-3;

n es un número entero de 0-2;

x' es un número entero de 1-23;

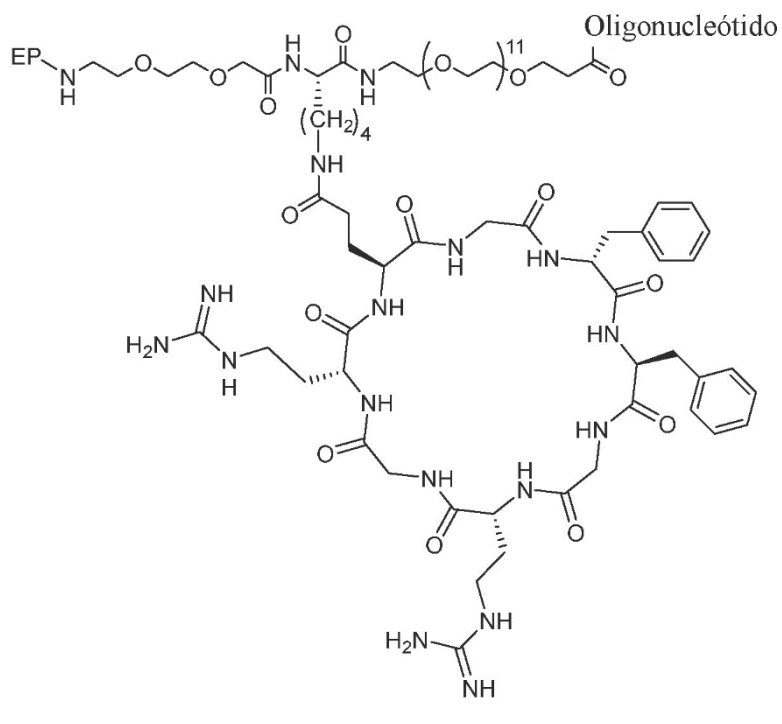
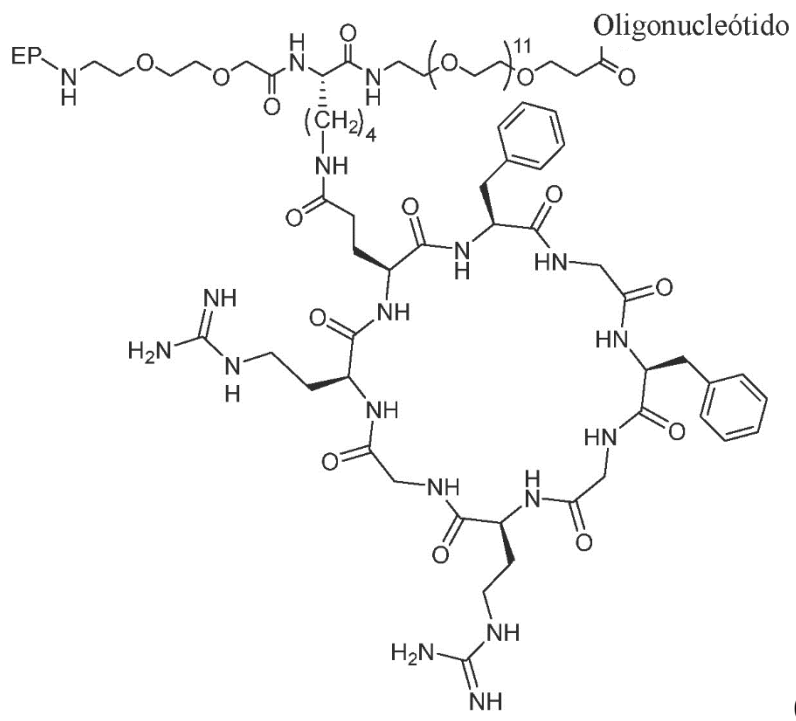
y es un número entero de 1-5;

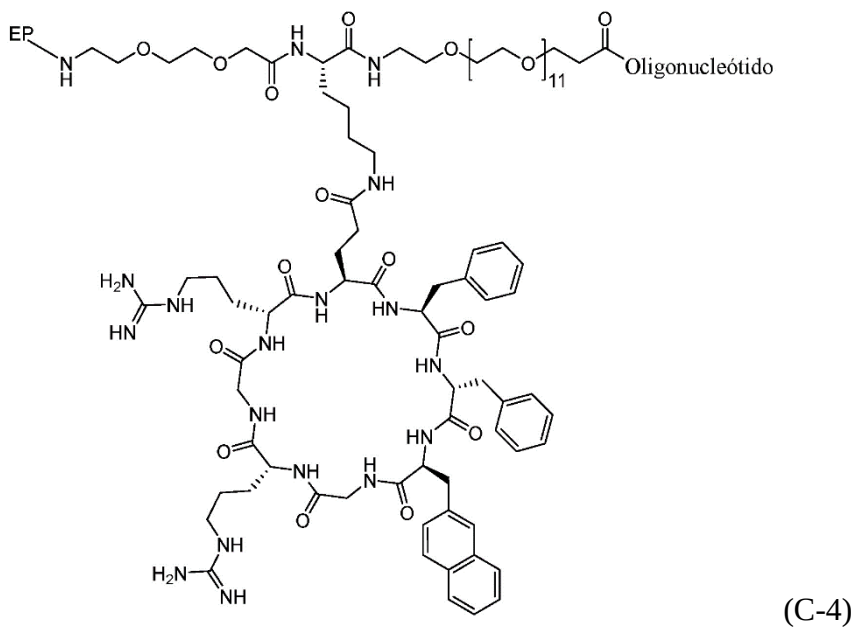
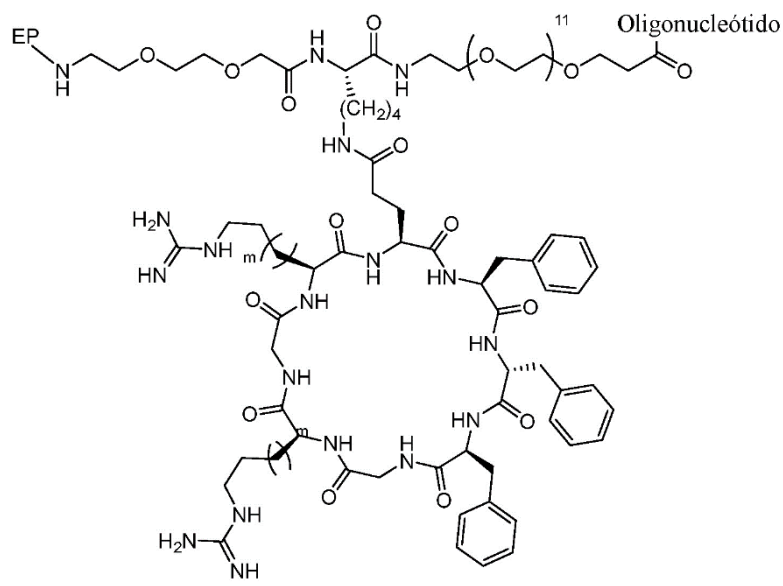
q es un número entero de 1-4;

z' es un número entero de 1-23, y

Carga es el compuesto antisentido.

En modalidades, el compuesto tiene una de las siguientes estructuras:





o una forma protonada o sal de estas,
 en donde EP es el péptido exocíclico, y
 el oligonucleótido es el compuesto antisentido.

En modalidades, el oligonucleótido del compuesto de fórmula (C-1), (C-2), (C-3) o (C-4) comprende la siguiente secuencia: 5'-CAG CAG CAG CAG CAG CAG CAG-3'.

En modalidades, el EP del compuesto de fórmula (C-1), (C-2), (C-3) o (C-4) comprende la siguiente secuencia: PKKKRKV.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La Figura 1 es un esquema que muestra múltiples estrategias para dirigirse a las repeticiones CUG en el ARNm.

La Figura 2 muestra nucleótidos modificados usados en los oligonucleótidos antisentido descritos en la presente descripción. Las estructuras 1-3 (1 = Fosforotioato; 2 = (S_{C5}-R_p)- α,β -CAN; 3 = PMO) son modificaciones de la cadena principal del fosfato; 4 (2-tio-dT) es una modificación de base; 5-8 (5 = 2'-OMe-ARN; 6 = 2'-O-MOE-ARN; 7 = 2'-F-ARN; 8 = 2'-F-ANA) son modificaciones del azúcar 2'; 9-11 son nucleótidos restringidos; 12-14 (9 = LNA; 10 = (S)-cET; 11 = tcADN; 12 = FHNA; 13 = (S)5'-C-metilo; 14 = UNA) son modificaciones adicionales del azúcar; y 15-18 (15 = E-VP; 16 = fosfonato de metilo; 17 = fosforotioato 5'; 18 = (S)-5'-C-metilo con fosfato) son modificaciones de estabilización del fosfato 5'; 19 es un azúcar morfolino. Reformateado de Khvorova, A., y otros, Nat. Biotechnol. Marzo de 2017; 35(3): 238–248.

Las Figuras 3A-3D ilustran las químicas de conjugación para conectar un AC a un péptido penetrante de células cíclico. La Figura 3A muestra la formación de enlaces amida entre péptidos con un grupo de ácido carboxílico o con éster activado TFP y residuos de amina primaria en el extremo 5' de un AC. La Figura 3B muestra la conjugación del AC modificado con amina secundaria o amina primaria en 3' y el éster péptido-TFP a través de la formación de enlaces amida. La Figura 3C muestra la conjugación de una péptido-azida al AC modificado con ciclooctino en 5' mediante la cicloadición de azida-alquino sin cobre. La Figura 3D demuestra otra conjugación ilustrativa entre un AC modificado con ciclooctino en 3' o AC modificado con azida en 3' y un CPP que contiene un resto enlazador-azida o enlazador-alquino/ciclooctino, mediante una cicloadición de azida-alquino sin cobre o cicloadición de azida-alquino catalizada por cobre, respectivamente (reacción clic).

La Figura 4 muestra la química de conjugación para conectar un AC y un CPP con una modalidad de enlazador adicional que contiene un resto de polietilenglicol (PEG).

Las Figuras 5A-5D proporcionan estructuras de monómeros de subunidades de morfolino de adenina (5A), citosina (5B), guanina (5C) y timina (5D) que pueden usarse para sintetizar oligómeros de morfolino enlazados a fosforodiamidato (PMO).

La Figura 6A-6F muestra el análisis por RT-PCR de eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de MBNL1 (exón 5; Figura 6A, 6B, 6E) y CLASP1 (exón 19; Figura 6B, 6D, 6F) 24 horas (6A-6B) y 48 horas (6C-6D) después de que células HeLa-48 se trataron con 1 μ M, 3 μ M o 10 μ M de diversos compuestos PMO o PMO-EEV mediante el uso del agente de transfección Endo-Porter (6A-6C) o sin el agente Endo-Porter (6E-6F). La línea celular HeLa parental y la línea celular HeLa-480 tratada con (6A-6D) o sin (6E-6F) el agente Endo-Porter,

se incluyeron como controles.

La Figura 7A-7B muestra el análisis por RT-PCR de eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de MBNL1 (exón 5; Figura 7A) y CLASP1 (exón 19; Figura 7B) 48 horas después de que los mioblastos DM1 se trataron con 1 μ M de diversos compuestos PMO o PMO-EEV sin el reactivo de transfección Endo-Porter. Dos controles, DM-04 sin tratamiento con endo-porter y DM-05 sin endo-porter se incluyeron como controles.

La Figura 8A-8D muestra el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de Atp2a1 (exón 22; Figura 8A), Nfix (exón 7; Figura 8B), Clcn1 (exón 7a; Figura 8C) y Mbnl1 (exón 5; Figura 8D) de tejidos del músculo gastrocnemio una semana después de que ratones HSA-LR (modelo de ratón DM1) se trataron con un PMO, 20 mpk de PMO-EEV 221-1106 o 40 mpk de PMO-EEV 221-1106. Se incluyeron ratones FVB/NJ (ratón endogámico de tipo salvaje) y HSA-LR (sin tratamiento) como grupos de control.

La Figura 9A-9D muestra el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de Atp2a1 (exón 22; Figura 9A), Nfix (exón 7; Figura 9B), Clcn1 (exón 7a; Figura 9C) y Mbnl1 (exón 5; Figura 9D) de tejidos del músculo cuádriceps una semana después de que ratones HSA-LR (modelo de ratón DM1) se trataron con un PMO, 20 mpk de PMO-EEV 221-1106 o 40 mpk de PMO-EEV 221-1106. Se incluyeron ratones FVB/NJ (ratón endogámico de tipo salvaje) y HSA-LR (sin tratamiento) como grupos de control.

La Figura 10A-10D muestra el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de Atp2a1 (exón 22; Figura 10A), Nfix (exón 7; Figura 10B), Clcn1 (exón 7a; Figura 10C) y Mbnl1 (exón 5; Figura 10D) de tejidos del músculo tibial anterior una semana después de que ratones HSA-LR (modelo de ratón DM1) se trataron con un PMO, 20 mpk de PMO-EEV 221-1106 o 40 mpk de PMO-EEV 221-1106. Se incluyeron ratones FVB/NJ (ratón endogámico de tipo salvaje) y HSA-LR (sin tratamiento) como grupos de control.

La Figura 11A-11F muestra el análisis por RT-PCR de eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de MBNL1 (exón 5, Figura 11A), SOS1 (exón 25, Figura 11B), IR (exón 11, Figura 11C), DMD (exón 78, Figura 11D), BIN1 (exón 11, Figura 11E) y LDB3 (exón 11, Figura 11F) después de que células musculares derivadas de pacientes con DM1 se trataron con diferentes concentraciones (10 μ m, 3 μ m, 1 μ m, 0.3 μ m) de EEV-PMO dirigidos a CUG en DMPK (CUG^{exp} 197-777 y CUG^{exp} 221-1106). Las células musculares de dos grupos, personas sanas (control negativo) y pacientes con DM1 (control positivo), se sometieron a prueba

para determinar los eventos de corte y empalme alternativo de ARN como control. Todos los datos se recopilaron a partir de tres experimentos individuales (n=3). Se realizó una prueba T de miotubos de DM1 tratados frente a no tratados; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

La Figura 12A-12F muestra el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de MBL1 (exón 5, Figura 12A), SOS1 (exón 25, Figura 12B), INSR (exón 11, Figura 12C), DMD (exón 78, Figura 12D), BIN1 (exón 11, Figura 12E) y LDB3 (exón 11, Figura 12F) después de que los mioblastos y miotubos de DM1 derivados de pacientes se trataran con 10 μ m, 3 μ m o 1 μ m de EEV-PMO 197-777 dirigido a CUG en DMPK. Se usaron células de pacientes sanos y células de DM1 como controles. Todos los datos se recopilaron a partir de tres experimentos individuales (n=3). Se realizó una prueba T de miotubos de DM1 tratados frente a no tratados; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

La Figura 13A-13B muestra los niveles relativos de ARNm después de que ratones HSA-LR se trataron con diversas concentraciones de PMO-EEV 221-1120. La Figura 13A muestra el nivel relativo de ARNm para el gastrocnemio, tríceps, tibial anterior y el diafragma. La Figura 13B muestra los niveles relativos de ARNm en el diafragma.

La Figura 14A-14C muestra los niveles relativos de ARNm en los tejidos del cuádriceps (14A), gastrocnemio (14B), tríceps (14C) y tibial anterior (14D) después de que ratones HSA-LR se trataron con diversas concentraciones de PMO-EEV 221-1120.

La Figura 15A-15D muestra el índice de corte y empalme de DM1 en ratón (mDSI) para diversos genes en tejidos del cuádriceps (Figura 15A), gastrocnemio (Figura 15B), tríceps (Figura 15C), y tibial anterior (Figura 15D) después de que ratones HSA-LR se trataron con diversas concentraciones de PMO-EEV 221-1120.

La Figura 16A-16C muestra la prevalencia de focos de ARN en el músculo tibial anterior después de que ratones HSA-LR no se trataron o se trataron con EEV-PMO 221-1120 (EEV-PMO-DM1-3; DM1-3). La Figura 16A-16B muestra imágenes de tejido del músculo tibial anterior teñido para detectar los focos de ARN CUG (rojo) y los núcleos (azul). La Figura 16C es un gráfico que cuantifica el por ciento de núcleos que tienen un foco de CUG a partir de datos asociados con las imágenes en la Figura 16A-16B.

La Figura 17A-17F son gráficos que muestran una respuesta dependiente de la dosis para los niveles del fármaco en los tejidos del cuádriceps (17A), tríceps (17B), corazón (17C), gastrocnemio (17D), tibial anterior (17E), diafragma (17F), cerebro (17H), hígado (17I) y riñones (17J) después de que ratones HSA-LR se trataron con diversas concentraciones de EEV-PMO-DM1-3. La Figura 17K

muestra la exposición al fármaco de diversos tejidos a un nivel de dosificación de 60 mpk.

La Figura 18 muestra una reducción de la miotonía dependiente de la dosis en ratones HSA-LR 7 días después del tratamiento con EEV-PMO-DM1-3 a 15, 30, 60 y 90 mpk.

La Figura 19A-19D son gráficos que muestran los resultados de un análisis de componente principal que compara la expresión génica en ratones sin enfermedad (WT), ratones DM1 (HSA-LR) y ratones HSA-LR tratados con PMO-EEV 221-1120. La Figura 19A y 19C son gráficos que muestran tres componentes principales y la Figura 19B y 19D son gráficos que muestran dos componentes principales.

La Figura 20A-20B muestra mapas térmicos de genes expresados diferencialmente entre ratones sin enfermedad (WT), ratones DM1 (HSA-LR) y ratones HSA-LR tratados con 60 mpk de PMO-EEV 221-1120. La Figura 20A es un mapa térmico agrupado que muestra 513 genes expresados diferencialmente. La Figura 20B es un mapa térmico agrupado que muestra 40 genes que se conoce que tienen repeticiones CTG·CUG.

La Figura 21 es un gráfico de volcán que muestra el cambio transcripcional global en los ratones HSA-LR no tratados y en los ratones tratados con PMO-EEV 221-1120.

La Figura 22A-22E son gráficos que muestran el resultado de un análisis de componente principal para los genes Scube2 (22A), Greb1 (22B), Ttc7 (22C), Txlnb(CUG)9 (22D) y Ndr3 (22E) de ratones sin enfermedad, ratones HSA-LR y ratones HSA-LR tratados con PMO-EEV 221-1120.

La Figura 23A-23D muestra datos de secuenciación de ARN (ARNsec) para Atp2a1 (23A; el exón 22 está en la caja), Clcn1 (23B; el exón 7a está en la caja), Nfix (23C; el exón 7 está en la caja) y Mbn1 (23D; el exón 5 está en la caja) para ratones sin enfermedad (WT-solución salina), ratones HSA-LR (HSA-LR-solución salina) y ratones HSA-LR tratados con PMO-EEV 221-1120. Se muestran dos lecturas para cada grupo de tratamiento.

La Figura 24 muestra el índice del por ciento de empalme (PSI) de exones individuales para diversos genes de interés para ratones sin enfermedad (WT-solución salina), ratones HSA-LR (HSA-LR-solución salina) y ratones HSA-LR tratados con PMO-EEV 221-1120.

La Figura 25A-25D muestran los niveles de fármaco en ratones HSA-LR tratados con 80 mpk (60 mpk de oligo, 80 mpk de fármaco completo) de EEV-PMO-DM1-3 después de 1 semana a 4 semanas en los tejidos del tibial anterior (25A), gastrocnemio (25B), tríceps (25C) y cuádriceps (25D).

La Figura 26A-26D muestra los niveles de fármaco en ratones después de que ratones HSA-LR se trataron con una única dosis de 80 mpk de EEV-PMO-DM1-3. La Figura 26A-26B muestra los niveles de fármaco en el hígado desde 1 semana hasta 12 semanas después del tratamiento. La Figura

26C-26D muestra los niveles de fármaco en el riñón desde 1 semana hasta 12 semanas después del tratamiento.

La Figura 27A-27C son gráficos que muestran el nivel de inclusión de exones para MBNL1 (exón 5; 26A), SOS1 (exón 25; 26B) y NFIX (exón 7; 26C) después de que células musculares derivadas de pacientes con DM1 se trataran con EEV-PMO-DM1-3 30 μ M.

La Figura 28A-28C muestra que EEV-PMO-DM1-3 reduce los focos nucleares de CUG (verde) en el núcleo (azul) en células musculares derivadas de pacientes con DM1. La Figura 28A-28B son imágenes de células musculares derivadas de pacientes con DM1 que no se trataron o se trataron con EEV-PMO-DM1-3 o no tratadas. La Figura 28C es la cuantificación del número de focos de CUG por núcleo para los datos asociados con las imágenes en la Figura 28A.

La Figura 29A-29B muestra los datos sin procesar (29A) y los datos normalizados (29B) de un ensayo de viabilidad luminiscente CELLTITER-GLO donde las células RPTEC se trataron con diversas concentraciones de PMO-DM1 o EEV-PMO-DM1-3. La melitina se usó como un control positivo.

Las Figuras 30A-30C muestran imágenes que representan los focos de ARN con repeticiones CUG en células derivadas de pacientes con DM1 (30A) y células derivadas de pacientes con DM1 tratadas con un EEV-PMO 221-1113 (30B). Las células se tiñeron para detectar los núcleos (azul; Hoechst) y los focos de ARN CUG (verde). La Figura 30C es un gráfico de los focos de ARN CUG por área nuclear para los datos asociados con las imágenes de la Figura 30A-30B.

La Figura 31A-31B muestra la prevalencia de los focos de ARN CUG7 en células HeLa, células HeLa480 no tratadas y células HeLa480 tratadas con EEV-PMO 221-1113. La Figura 31A muestra imágenes de células teñidas para detectar los focos de ARN CUG7 (verde) y los núcleos (azul). La Figura 31B es un gráfico que cuantifica los focos de CUG7 por área nuclear para los datos asociados con las imágenes en la Figura 31A.

Las Figuras 32A-32C son gráficos que muestran el por ciento de inclusión del exón 5 en MBNL1 (32A), el exón 25 en SOS1 (32B) y el exón 7 en NFIX (32C) después de que células derivadas de pacientes con DM1 se trataran con 30 μ M de EEV-PMO 221-1113.

Las Figuras 33A-33E muestran el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión de exones) de MBNL1 (exón 5; 33A), SOS1 (exón 25; 33B), CLASP1 (exón 19, 33C), NFIX (exón 7, 33D) e INSR (exón 11, 33E) después de que células musculares derivadas de pacientes con DM1 se trataran con diversas concentraciones de PMO-EEV 221-1113. Se usó la prueba T para determinar la significación; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Las Figuras 34A-34D muestran el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión de exones) Atp2a1 (exón 22, 34A), Nfix (exón 7, 34B), Clcn1 (exón 7a, 34C) y Mbnl1 (exón 5, 34D) en el tejido del gastrocnemio de ratones tratados con diversas concentraciones de PMO 221 o EEV-PMO 221-1106.

- 5 Las Figuras 35A-35C muestran el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión o exclusión de exones) de Mbnl1 (exón 5, 35A), Nfix (exón 7, 35B) y Atp2a1 (exón 22, 35C) en el tejido del tibial anterior de ratones HSA-LR tratados con PMO-EEV 0221-1121 (21-mer) o PMO-EEV 0325-1121 (24-mer).

- 10 Las Figuras 36A-36C muestran el análisis por RT-PCR de los eventos de corte y empalme alternativo de ARN (por ejemplo, inclusión de exones) de Mbnl1 (exón 5, 36A), Nfix (exón 7, 36B) y Atp2a1 (exón 22, 36C) en el tejido del gastrocnemio de ratones HSA-LR tratados con PMO-EEV 0221-1121 (21-mer) o PMO-EEV 0325-1121 (24-mer).

La Figura 37 muestra a PMO-0221a, el metabolito principal de PMO-EEV 220-1120 detectado *in vivo*.

- 15 Las Figuras 38A-38B muestran el por ciento de inclusión de exones en el tibial anterior (38A) y el gastrocnemio (38B) para MBNL1 (exón 5) después de que células Hela480 se trataran con diversas concentraciones de EEV-PMO 221-1120.

- 20 Las Figuras 39A-39B muestran el por ciento de inclusión de exones en el tibial anterior (39A) y el gastrocnemio (39B) para NFIX (exón 7) después de que células Hela480 se trataran con diversas concentraciones de EEV-PMO 221-1120.

Las Figuras 40A-40B muestran el por ciento de inclusión de exones en el tibial anterior (40A) y el gastrocnemio (40B) para Atp2a1 (exón 22) después de que células Hela480 se trataran con diversas concentraciones de EEV-PMO 221-1120.

- 25 La Figura 41 muestra imágenes (41A) que representan los focos de ARN con repeticiones CUG en células Hela480 después del tratamiento con diversas concentraciones de EEV-PMO 221-1120. La Figura 41B es un gráfico de los focos de ARN por área nuclear para los datos asociados con las imágenes de la Figura 41A.

- 30 Las Figuras 42A-42D muestran los niveles relativos de ARNm con repeticiones r(CUG480), (42A), los niveles relativos de ARNm de DMPK (42B), el por ciento de inclusión del exón 5 de MBNL1 (42C) y el por ciento de inclusión del exón 25 en SOS1 (42D) en células HeLa480 después del tratamiento con diversas concentraciones de EEV-PMO 221-1120.

La Figura 43 es un gráfico de barras que muestra ejemplos de genes expresados en tejido muscular

que se conoce que tienen repeticiones CTG·CUG.

Las Figuras 44A-44D muestran una reducción de la miotonía fenotípica en el modelo de ratón HSA-LR tratado con 20 mpk de PMO-EEV 221-1106. Las Figuras 44A y 44C muestran gráficos de relajación. La Figura 44B muestra un ejemplo de trazado de fuerza sin procesar. La Figura 44D muestra trazados de electromiografía representativos.

DESCRIPCIÓN DETALLADA

Compuestos

En modalidades, se proporcionan compuestos que modulan el nivel y/o la actividad de un transcrito génico que tiene una repetición de trinucleótidos CUG expandida. En modalidades, los compuestos de la presente descripción incluyen al menos un péptido penetrante de células cíclico (cCPP) y un resto terapéutico (TM). El cCPP facilita la entrada del TM a la célula. En modalidades, el compuesto incluye un vehículo de escape endosómico (EEV) que comprende el cCPP y un péptido exocíclico (EP). El cCPP o el EEV pueden permitir que el TM entre al citosol o a un compartimiento celular para interactuar con el transcrito objetivo.

Restos terapéuticos

Generalmente, el TM es el resto efector que provoca una respuesta. En modalidades, el TM provoca una respuesta mediante la modulación de la expresión, la actividad y/o el nivel de un transcrito objetivo y/o una proteína objetivo. En modalidades, el transcrito objetivo incluye una repetición de trinucleótidos CUG expandida. En modalidades, el TM modula los niveles de un transcrito objetivo y/o proteína objetivo dentro de una célula. En modalidades, el TM disminuye el nivel del transcrito objetivo y/o proteína objetivo dentro de una célula.

En modalidades, el TM modula la actividad del transcrito objetivo al reducir la afinidad entre el transcrito objetivo y una o más proteínas que se unen al transcrito objetivo. Al reducir la afinidad entre el transcrito objetivo y la una o más proteínas, el TM puede modular efectivamente la actividad de la una o más proteínas que de cualquier otra manera se asociarían con el transcrito objetivo. Por ejemplo, si la una o más proteínas no se unen al transcrito objetivo, están disponibles para llevar a cabo sus funciones en otras moléculas. Por ejemplo, si la una o más proteínas están involucradas en el procesamiento de pre-ARNm, reducir la afinidad de la una o más proteínas por un transcrito que comprende una repetición CUG expandida puede permitir que la una o más proteínas procesen transcritos de pre-ARNm que no comprenden repeticiones CUG expandidas. Como tal, el TM puede modular la actividad, la expresión y/o los niveles de los genes aguas abajo (genes que no contienen

la repetición CTG expandida) que están regulados por la una o más proteínas cuya interacción con el transcrito objetivo se interrumpe por el TM.

En modalidades, el TM comprende un oligonucleótido, un péptido, un anticuerpo y/o una molécula pequeña. La clase y la identidad del TM dependen del mecanismo que se usa para modular el nivel y/o la actividad del transcrito objetivo que incluye una repetición de trinucleótidos CUG expandida.

Compuesto antisentido

En diversas modalidades, los compuestos descritos en la presente descripción comprenden un péptido penetrante de células (CPP) conjugado a un compuesto antisentido (AC).

El término “compuesto antisentido” se refiere a una secuencia de oligonucleótidos que es complementaria, o al menos parcialmente complementaria, a una secuencia de nucleótidos objetivo.

Un AC es un oligonucleótido que incluye bases de ADN naturales, bases de ADN modificadas, bases de ARN naturales, bases de ARN modificadas, azúcares de ARN naturales, azúcares de ARN modificados, azúcares de ADN naturales, azúcares de ADN modificados, enlaces internucleósidos naturales, enlaces internucleósidos modificados, o cualquiera de sus combinaciones. Los AC incluyen, pero no se limitan a, ARNi de oligonucleótidos antisentido, microARN, antagomirs, aptámeros, ribozimas, oligonucleótidos inmunoestimuladores, oligonucleótidos señuelo, supermir, imitadores de miARN, inhibidores de miARN, adaptadores U1 y sus combinaciones.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a un transcrito objetivo que tiene una repetición de trinucleótidos CUG expandida.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una repetición de trinucleótidos CUG expandida en una secuencia de ARNm objetivo. Varias enfermedades se asocian con repeticiones de trinucleótidos CUG expandidas, por ejemplo, la distrofia miotónica tipo 1 (DM1), la distrofia endotelial corneal de Fuchs (FECD), la ataxia espinocerebelosa-8 (SCA8) y la enfermedad similar a Huntington (HDL2). La Tabla 1 proporciona ejemplos de trastornos por repeticiones de nucleótidos, y características de los genes con repeticiones de nucleótidos expandidas asociadas con dichos trastornos. El siguiente documento describe oligonucleótidos ilustrativos para tratar enfermedades de repeticiones en tándem y se incorpora como referencia en la presente descripción en su totalidad: Zain y otros, Neurotherapeutics. 2019; 16(2): 248-262; Zarouchlioti y otros, Am J Hum Genet. 2018;102(4):528-539; Fautsch y otros, Prog Retin Eye Res. 2021; 81:100883.

Tabla 1: Enfermedades asociadas con repeticiones de trinucleótidos CUG expandidas

Enfermedad (abreviatura)	Gen	Longitud de la repetición normal	Longitud de la repetición expandida	Producto génico	Secuencia de la repetición	Ubicación de la repetición
DM1	DMPK	5-35	> 50	Proteína cinasa de la distrofia miotónica	CTG•CAG	3' UTR
FECD	TCF4	< 30	> 40	Factor de transcripción 4	CTG•CAG	Intrón 3
SCA8	ATXN8OS y/o ATXN8	15-50	> 50	Ataxina 8 y hebra opuesta de ataxina 8	CTG•CAG	3' UTR
HDL2	JPH3	6-27	>40	Junctofilina 3	CTG•CAG	3' UTR

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una secuencia de nucleótidos que está dentro de un transcrito de ARNm objetivo que incluye una repetición de trinucleótidos CTG•CUG expandida. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una repetición de trinucleótidos CTG•CUG expandida en un transcrito de ARNm objetivo.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una secuencia de nucleótidos que está dentro de un transcrito objetivo de DMPK1 que incluye una repetición de trinucleótidos CTG•CUG expandida. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una secuencia de nucleótidos que está dentro de un transcrito objetivo de TCF4 que incluye una repetición de trinucleótidos CTG•CUG expandida. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una secuencia de nucleótidos que está dentro de un transcrito objetivo de ATXN8OS/ATXN8 que incluye una repetición de trinucleótidos CTG•CUG expandida. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una secuencia de nucleótidos que está dentro de un transcrito objetivo de JPH3 que incluye una repetición de trinucleótidos CTG•CUG expandida.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una repetición de trinucleótidos en una UTR 3' de un transcrito de ARNm objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente

complementaria a una repetición de trinucleótidos CTG·CUG expandida en una UTR 3' de un transcrito objetivo de DMPK1. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a una repetición de trinucleótidos CTG·CUG expandida en una UTR 3' de un transcrito objetivo de ATXN8OS/ATXN8. En modalidades, el AC incluye una
5 secuencia de nucleótidos que es al menos parcialmente complementaria a una repetición de trinucleótidos CTG·CUG expandida en una UTR 3' de un transcrito objetivo de JPH3.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a repeticiones de trinucleótidos, tales como las repeticiones CTG·CUG. En modalidades, la secuencia de nucleótidos objetivo comprende al menos una repetición de
10 trinucleótidos expandida (por ejemplo, repeticiones CTG·CUG). En modalidades, la secuencia de nucleótidos objetivo comprende al menos 40, al menos 45, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 150, al menos 200, al menos 300, al menos 400, al menos 500, al menos 600, al menos 700, al menos 800, al menos 900, al menos 1,000 o al menos 2,000 repeticiones de trinucleótidos CTG·CUG. En modalidades, la repetición de trinucleótidos
15 expandida está en la UTR 3' de la secuencia de nucleótidos objetivo.

En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en el transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al
20 menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, y hasta 50, hasta 100, hasta 150, hasta 200, hasta 300, hasta 400, hasta 500, hasta 600, hasta 700, hasta 800, hasta 900, hasta 1,000 o hasta 2,000 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia
25 de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 5 a 10 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente
30 complementaria a, y puede hibridarse con, 5 a 9 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 5 a 8 repeticiones de trinucleótidos en un transcrito

objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 5 a 7 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 5 a 6 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 5 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 6 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 7 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, 8 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es complementaria al menos parcialmente a, y puede hibridarse con, 9 repeticiones de trinucleótidos en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que

En modalidades, el AC puede incluir una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en cualquier ubicación en un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en la UTR 3' de un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en la UTR 3' del transcrito objetivo de DMPK1, SCA8 y/or HDL2. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en un intrón de un transcrito objetivo. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en el intrón 3 de un transcrito objetivo de TCF4. En modalidades, el AC incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en el locus CTG18.1 del transcrito de TCF4. En modalidades, el AC

incluye una secuencia de nucleótidos que es al menos parcialmente complementaria a, y puede hibridarse con, al menos una porción de las repeticiones de trinucleótidos expandidas contiguas presentes en un exón de un transcrito objetivo.

En modalidades, el AC tiene 5 o más, 10 o más, 15 o más, 20 o más, 25 o más, 30 o más, 35 o más, 40 o más, o 45 o más ácidos nucleicos de longitud. En modalidades, el AC tiene 50 o menos, 45 o menos, 40 o menos, 35 o menos, 30 o menos, 25 o menos, 20 o menos, 15 o menos, o 10 o menos ácidos nucleicos de longitud. En modalidades, el AC tiene de 5 a 50, 5 a 45, 5 a 40, 5 a 35, 5 a 30, 5 a 25, 5 a 20, 5 a 15 o 5 a 10 ácidos nucleicos de longitud. En modalidades, el AC tiene de 10 a 50, 10 a 45, 10 a 40, 10 a 35, 10 a 30, 10 a 25, 10 a 20 o 10 a 15 ácidos nucleicos de longitud. En modalidades, el AC tiene de 15 a 50, 15 a 45, 15 a 40, 15 a 35, 15 a 30, 15 a 25 o 15 a 20 ácidos nucleicos de longitud. En modalidades, el AC tiene de 20 a 50, 20 a 45, 20 a 40, 20 a 35, 20 a 30 o 20 a 25 ácidos nucleicos de longitud. En modalidades, el AC tiene de 25 a 50, 25 a 45, 25 a 40, 25 a 35 o 25 a 30 ácidos nucleicos de longitud. En modalidades, el AC tiene de 30 a 50, 30 a 45, 30 a 40 o 30 a 35 ácidos nucleicos de longitud. En modalidades, el AC tiene de 35 a 50, 35 a 45 o 35 a 40 ácidos nucleicos de longitud. En modalidades, el AC tiene de 40 a 50 o de 40 a 45 ácidos nucleicos de longitud. En modalidades, el AC tiene de 45 a 50 ácidos nucleicos de longitud. En modalidades, el AC tiene 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 o 50 ácidos nucleicos de longitud.

En modalidades, el AC tiene una complementariedad del 100 % con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC no tiene una complementariedad del 100 % con respecto a una secuencia de nucleótidos objetivo. Como se usa en la presente, el término "por ciento de complementariedad" se refiere al número de nucleobases (por ejemplo, nucleobase natural o nucleobase modificada) de un AC que tienen complementariedad de nucleobase con una nucleobase correspondiente de un compuesto oligomérico o ácido nucleico (por ejemplo, una secuencia de nucleótidos objetivo) dividido por la longitud total (número de nucleobases) del AC. Un experto en la técnica reconoce que es posible la inclusión de errores de emparejamiento sin eliminar la actividad del compuesto antisentido.

En modalidades, el AC incluye 20 % o menos, 15 % o menos, 10 % o menos, 5 % o menos, o cero errores de emparejamiento con respecto a la secuencia de nucleótidos objetivo. En algunas modalidades, el AC incluye 5 % o más, 10 % o más, o 15 % o más errores de emparejamiento. En modalidades, el AC incluye cero a 5 %, cero a 10 %, cero a 15 %, o cero a 20 % de errores de

emparejamiento con respecto a la secuencia de nucleótidos objetivo. En modalidades, el AC incluye 5 % a 10 %, 5 % a 15 %, o 5 % a 20 % de errores de emparejamiento con respecto a la secuencia de nucleótidos objetivo. En modalidades, el AC incluye 10 % a 15 % o 10 % a 20 % de errores de emparejamiento con respecto a la secuencia de nucleótidos objetivo. En modalidades, el AC incluye de 10 % a 20 % de errores de emparejamiento con respecto a la secuencia de nucleótidos objetivo. En modalidades, el AC tiene 80 % o más, 85 % o más, 90 % o más, 95 % o más, 96 % o más, 97 % o más, 98 % o más, o 99 % o mayor complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene 100 % o menos, 99 % o menos, 98 % o menos, 97 % o menos 96 % o menos 95 % o menos, 90 % o menos, 85 % o menos complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 80 % a 100 %, de 80 % a 99 %, de 80 % a 98 %, de 80 % a 97 % de 80 % a 96 %, de 80 % a 95 %, de 80 % a 90 % u de 80 % a 85 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 85 % a 100 %, de 85 % a 99 %, de 85 % a 98 %, de 85 % a 97 % de 85 % a 96 %, de 85 % a 95 %, u de 85 % a 90 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 90 % a 100 %, de 90 % a 99 %, de 90 % a 98 %, de 90 % a 97 %, de 90 % a 96 %, o de 90 % a 95 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 95 % a 100 %, de 95 % a 99 %, de 95 % a 98 %, de 95 % a 97 %, o de 95 % a 96 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 96 % a 100 %, de 96 % a 99 %, de 96 % a 98 %, o de 96 % a 97 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 97 % a 100 %, de 97 % a 99 %, o de 97 % a 98 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene de 98 % a 100 % o de 98 % a 99 % de complementariedad con respecto a una secuencia de nucleótidos objetivo. En modalidades, el AC tiene una complementariedad del 99 % al 100 % con respecto a una secuencia de nucleótidos objetivo.

En modalidades, la incorporación de modificaciones de la afinidad de los nucleótidos permite un mayor número de errores de emparejamiento en comparación con un compuesto no modificado. De manera similar, ciertas secuencias de oligonucleótidos pueden ser más tolerantes a los errores de emparejamiento que otras secuencias de oligonucleótidos. Un experto en la técnica es capaz de determinar un número apropiado de errores de emparejamiento entre un AC y una secuencia de nucleótidos objetivo, tal como mediante la determinación de la temperatura de fusión térmica (T_m). La T_m o ΔT_m pueden calcularse mediante técnicas que son familiares para un experto en la técnica.

Por ejemplo, las técnicas descritas en Freier y otros, (Nucleic Acids Research, 1997, 25, 22: 4,429-4,443) permiten a un experto en la técnica evaluar las modificaciones de nucleótidos para determinar su capacidad de aumentar la temperatura de fusión de un dúplex ARN:ADN.

En modalidades, el AC incluye una secuencia de nucleótidos que en sí misma es una repetición de trinucleótidos, es decir, una repetición de trinucleótidos CAG. El complemento inverso de un 5'-CAG-3' tiene una complementariedad del 100 % y puede hibridarse con una repetición de trinucleótidos 5'-CUG-3'. En modalidades, el AC incluye de una a 50 repeticiones CAG. En modalidades, las repeticiones CAG son contiguas. En modalidades, las repeticiones CAG no son contiguas. En modalidades, el AC incluye una secuencia de nucleótidos que incluye repeticiones CAG incompletas en el extremo 5' o 3'. Por ejemplo, en modalidades, el AC incluye una secuencia tal como AG(CAG)_n, G(CAG)_n, (CAG)_nAG o (CAG)_nA, donde n es un número entero de 1 a 50. En modalidades, el AC incluye una secuencia de nucleótidos que incluye uno o más, 2 o más, 3 o más, 4 o más, 5 o más, 6 o más, 7 o más, 8 o más, 9 o más, 10 o más, 20 o más, 30 o más, o 50 o más repeticiones CAG. En modalidades, el AC incluye una secuencia de nucleótidos que incluye 50 o menos, 40 o menos, 30 o menos, 20 o menos, 10 o menos, 9 o menos, 8 o menos, 7 o menos, 6 o menos, 5 o menos, 4 o menos, 3 o menos, o 2 o menos repeticiones CAG. En modalidades, el AC incluye una secuencia de nucleótidos que incluye de 2 a 50, de 2 a 20, de 2 a 10, de 4 a 10, de 5 a 10, de 6 a 10, de 6 a 9, de 6 a 8 o de 6 a 7 repeticiones CAG. En modalidades, el AC incluye una cualquiera de las secuencias de nucleótidos en la Tabla 2 (SEQ ID NO: 151-291).

Tabla 2: Secuencias de nucleótidos de AC con repeticiones CAG

Secuencia de AC (5' a 3')	SEQ ID NO:
CAG	NA
CAG-CAG	NA
CAG-CAG-CAG	NA
CAG-CAG-CAG-CAG	151
CAG-CAG-CAG-CAG-CAG	152
CAG-CAG-CAG-CAG-CAG-CAG	153
CAG-CAG-CAG-CAG-CAG-CAG-CAG	154
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	155
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	156
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	157
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	158
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	159
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	160

5	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	161
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	162
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	163
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	164
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	165
10	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	166
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	167
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	168
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	169
15	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	170
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	171
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	172
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	173
20	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	174
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	175
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	176
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	177
25	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	178
	CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	179

	GCA	NA
	GCAGCA	NA
	GCAGCAGCA	NA
	GCAGCAGCAGCA	245
	GCAGCAGCAGCAGCA	246
5	GCAGCAGCAGCAGCAGCA	247
	GCAGCAGCAGCAGCAGCAGCA	248
	GCAGCAGCAGCAGCAGCAGCAGCA	249
	GCAGCAGCAGCAGCAGCAGCAGCAGCA	250
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	251
10	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	252
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	253
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	254
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	255
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	256
15	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	257
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	258
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	259
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	260
	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	261
20	GCA	262
	GCA	263
	GCA	264
	GCA	265
	GCA	266
25	GCA	267
	GCA	268
	GCA	269
	GCA	
	GCA	
30	GCA	
	GCA	

ácido nucleico objetivo del transcrito objetivo que está separada de la repetición CUG expandida del transcrito objetivo.

En modalidades, un AC de la descripción es un oligonucleótido gapmer como se describe en la patente de Estados Unidos núm. 9,550,988, cuya descripción se incorpora como referencia en la presente descripción.

En modalidades, un AC de la descripción comprende la secuencia y/o estructura de uno cualquiera de los AC dirigidos a DMPK descritos en la publicación de patente de Estados Unidos núm. 2017/0260524, cuya descripción se incorpora como referencia en la presente descripción.

En modalidades, un AC de la descripción comprende la secuencia y/o estructura de uno cualquiera de los AC u oligonucleótidos descritos en las publicaciones de patente de Estados Unidos US20030235845A1, US20060099616A1, US 2013/0072671 A1, US 2014/0275212 A1, US 2009/0312532 A1, US20100125099A1, US 2010/0125099 A1, US 2009/0269755 A1, US 2011/0294753 A1, US 2012/0022134 A1, US 2011/0263682 A1, US 2014/0128592 A1, US 2015/0073037 A1 y US20120059042A1, cuyos contenidos se incorporan en la presente descripción en su totalidad para todos los propósitos.

Cuando se usa un AC para dirigirse y/o hibridarse a una repetición CTG·CUG expandida, se debe tener cuidado de evitar efectos fuera del objetivo donde el AC se una no intencionalmente a transcritos que no son el objetivo que incluyen repeticiones CTG·CUG (por ejemplo, un transcrito que incluye repeticiones CTG·CUG que no son una repetición CTG·CUG expandida). Un análisis informático del genoma humano revela que, en total, 63 genes humanos tienen repeticiones CTG·CUG (Uhlen, y otros, *Science* 2015 347(6220):1260419)). Los 63 genes pueden clasificarse por la expresión de ARNm más la cantidad de proteína expresada en el músculo total (músculo cardiaco, esquelético y liso). El nivel de expresión puede cuantificarse mediante el uso de RPM (lecturas por millón) la expresión de ARNm es FPKM (fragmentos por kilobase de transcripción por millón de fragmentos mapeados) y la expresión de proteínas es pTPM (transcritos por millón de genes codificantes de proteínas), mediante el uso de más de 10 RPM como valor de corte para la expresión no insignificante. La Figura 43 muestra los resultados de dicho análisis informático. Treinta y seis genes muestran un nivel de expresión de > 10 RPM. De los 36 genes, solo tres genes (además de DMPK) tenían > 10 repeticiones CTG·CUG. Los genes con ≤ 10 repeticiones CTG·CUG representan el menor riesgo de unión fuera del objetivo y de toxicidad. No obstante, el número de repeticiones CTG·CUG (11-24) en estos 3 genes (TCF4, CASK, MAP3k4) es significativamente inferior al observado en pacientes con DM1 clásica y congénita. Por ejemplo, los pacientes con DM1

de inicio tardío tienen 100-600 repeticiones CTG·CUG en la DMPK, los pacientes con DM1 clásica tienen 250-750 repeticiones CTG·CUG en la DMPK y los pacientes con DM1 congénita tienen 750-1,400 repeticiones CTG·CUG en la DMPK. Se puede realizar el mismo análisis informático para el hígado y el riñón. CASK es el único gen significativo con > 10 repeticiones CTG·CUG en el riñón.

5 Ningún gen con > 10 repeticiones CTG·CUG fue significativo en el hígado.

Los AC descritos en la presente descripción pueden contener uno o más centros asimétricos y, por tanto, dan lugar a enantiómeros, diastereómeros y otras configuraciones estereoisoméricas que pueden definirse, en términos de estereoquímica absoluta, como (R) o (S); α o β ; o como (D) o (L). En los compuestos antisentido proporcionados en la presente descripción se incluyen todos dichos posibles isómeros, así como también sus formas racémicas y ópticamente puras.

10 La eficacia de los AC puede evaluarse mediante la evaluación de la actividad antisentido efectuada por su administración. Como se usa en la presente, el término "actividad antisentido" se refiere a cualquier actividad detectable y/o medible atribuible a la hibridación de un compuesto antisentido a su secuencia de nucleótidos objetivo. Dicha detección y/o medición puede ser directa o indirecta. En 15 modalidades, la actividad antisentido se evalúa al detectar y/o medir la cantidad de la proteína expresada a partir del transcrito de interés. En modalidades, la actividad antisentido se evalúa al detectar y/o medir la cantidad del transcrito de interés. En modalidades, la actividad antisentido se evalúa al detectar y/o medir la cantidad de ARN empalmado alternativamente y/o la cantidad de isoformas de proteínas traducidas a partir del transcrito objetivo.

20 Mecanismos de modulación del AC

En modalidades, el AC puede modular la actividad y/o el nivel del transcrito objetivo dentro de la célula. La Figura 1 muestra mecanismos ilustrativos de cómo un AC puede modular el nivel y/o la actividad de un transcrito objetivo.

En modalidades, el AC puede modular el nivel del transcrito objetivo dentro de una célula. Por 25 ejemplo, en modalidades donde el AC es un gapmer, la unión del AC al transcrito objetivo induce la degradación del transcrito objetivo mediante las vías de la ARNasa H (Figura 1, flechas A y B). En modalidades, el gapmer se hibrida con una porción del transcrito objetivo que es distinta de la repetición del trinucleótido CUG expandida, y de esta manera, induce la degradación del transcrito objetivo mediante las vías de la ARNasa H (Figura 1, flecha A). En modalidades, el gapmer se 30 hibrida a al menos una porción de la repetición CUG expandida dentro del transcrito objetivo, y de esta manera, induce la degradación del transcrito objetivo mediante las vías de la ARNasa H (Figura 1, flecha B).

En modalidades, el AC puede modular la actividad del transcrito objetivo. La modulación de la actividad puede incluir aumentar o disminuir la capacidad del transcrito objetivo para unirse con un ligando. En modalidades, el AC puede modular la actividad del transcrito objetivo al disminuir la capacidad del transcrito objetivo para unirse con una o más proteínas que pueden asociarse con el transcrito, particularmente proteínas que se asocian con al menos una porción de la repetición CUG expandida de un transcrito objetivo (Figura 1, flecha C). En modalidades, disminuir la capacidad del transcrito objetivo para unirse a una o más proteínas, incluye disminuir la afinidad del transcrito objetivo por la una o más proteínas. En modalidades, disminuir la capacidad del transcrito objetivo para unirse a una o más proteínas, incluye el bloqueo estérico parcial o completo del transcrito objetivo para que no se una a una o más proteínas. Por ejemplo, el AC puede ocupar al menos una porción del sitio de unión que puede estar ocupado por la una o más proteínas si no se bloquean estéticamente. En modalidades, el sitio de unión de la una o más proteínas al transcrito objetivo incluye al menos una porción de la repetición de trinucleótidos expandida del transcrito objetivo. Como tal, en modalidades, el AC puede ocupar al menos una porción de la repetición de trinucleótidos expandida (por ejemplo, la repetición CUG expandida) que puede estar ocupada por la una o más proteínas si no se bloquean estéticamente. El bloqueo estérico parcial del transcrito objetivo puede dar como resultado una menor afinidad entre el transcrito objetivo y la una o más proteínas. Por ejemplo, en modalidades, el AC se une a al menos una porción de la repetición CUG expandida del transcrito objetivo, bloqueando y/o disminuyendo de esta manera la afinidad del transcrito objetivo por una proteína que puede unirse a la repetición CUG expandida (Figura 1, flecha C). El siguiente artículo de revisión describe aplicaciones adicionales para oligonucleótidos antisentido de bloqueo estérico y se incorpora como referencia en la presente descripción en su totalidad: Roberts y otros, *Nature Reviews Drug Discovery* (2020) 19: 673-694.

Las repeticiones CUG de una repetición CUG expandida pueden formar una estructura de horquilla bicatenaria. En el estado de enfermedad, las proteínas se unen a la estructura de horquilla bicatenaria y quedan secuestradas e incapaces de realizar otras funciones. En modalidades, el AC se une a al menos una porción de la estructura de horquilla bicatenaria, bloqueando y/o disminuyendo de esta manera la afinidad de la estructura de horquilla bicatenaria por un ligando de proteína. En modalidades, el AC se une a al menos una porción de una repetición CUG expandida monocatenaria, inhibiendo de esta manera la formación de la estructura de horquilla bicatenaria y, por tanto, inhibe que una o más proteínas se unan a la estructura de horquilla bicatenaria. En modalidades, la hibridación del AC a la estructura de horquilla doble bloquea estéricamente y/o disminuye la afinidad

por una o más proteínas por unirse a la estructura de horquilla doble. En modalidades, la hibridación del AC a al menos una porción de una región monocatenaria de la repetición de trinucleótidos expandida, inhibe la formación de una estructura de horquilla bicatenaria.

La disminución de la capacidad del transcrito objetivo para unirse con la una o más proteínas, puede permitir que la una o más proteínas lleven a cabo otras funciones tales como, por ejemplo, regular el corte y empalme de transcritos aguas abajo (transcritos que no contienen una repetición CUG expandida). Como tal, en modalidades, disminuir la capacidad del transcrito objetivo para unirse a la una o más proteínas, puede aumentar el nivel de la una o más proteínas dentro de la célula que están disponibles para proporcionar otras funciones o funciones en otros transcritos. En modalidades, disminuir la capacidad del transcrito objetivo para unirse a la una o más proteínas, puede aumentar el nivel citosólico de la una o más proteínas dentro de la célula que están disponibles para proporcionar otras funciones o funciones en otros transcritos. Por lo tanto, en modalidades, la unión del AC al transcrito objetivo puede dar como resultado la modulación del nivel y/o la actividad de la una o más proteínas que interactúan con el transcrito objetivo.

En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida, disminuye la afinidad y o bloquea estéricamente la unión de MNBL1 al transcrito objetivo. MNBL1 es un factor de corte y empalme que regula el corte y empalme de transcritos génicos aguas abajo. En un fenotipo de enfermedad DM1, MNBL1 se une a la repetición CUG expandida de un transcrito objetivo. Aunque se une al transcrito objetivo, MNBL1 se secuestra en el núcleo y no es capaz de regular el corte y empalme de transcritos génicos aguas abajo (transcritos que no contienen una repetición CUG expandida). En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en el transcrito objetivo, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito objetivo, permitiendo de esta manera regular el corte y empalme de transcritos génicos aguas abajo. En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en el transcrito objetivo, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito objetivo, aumentando de esta manera la cantidad de MNBL1 libre (por ejemplo, no unido a un transcrito que tiene una repetición de CUG). En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en el transcrito objetivo, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito objetivo, disminuyendo de esta manera la cantidad de MBNL1 unido a y secuestrado por el transcrito objetivo.

En modalidades donde el transcrito objetivo es DMPK, la hibridación del AC a al menos una porción de la repetición CUG expandida, disminuye la afinidad y/o bloquea estéricamente la unión de

MNBL1 al transcrito objetivo. MNBL1 es un factor de corte y empalme que regula el corte y empalme de transcritos génicos aguas abajo. En un fenotipo de enfermedad DM1, MNBL1 se une a la repetición CUG expandida de DMPK1. Aunque se une a DMPK1, MNBL1 se secuestra en el núcleo y no es capaz de regular el corte y empalme de transcritos génicos aguas abajo (transcritos que no contienen una repetición CUG expandida). En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en la DMPK, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito de DMPK, permitiendo de esta manera regular el corte y empalme de transcritos génicos aguas abajo. En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en la DMPK, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito de DMPK, aumentando de esta manera la cantidad de MNBL1 libre (por ejemplo, no unido a un transcrito que tiene una repetición CUG). En modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida en la DMPK, bloquea estéricamente y/o disminuye la afinidad de MNBL1 por el transcrito de DMPK, disminuyendo de esta manera la cantidad de MBNL1 unido a y secuestrado por el transcrito de DMPK1.

En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC a al menos una porción de la repetición CUG expandida, da como resultado la disminución de los niveles de CUGBP1. En el estado de enfermedad DM1, el nivel de MBNL1 libre (capaz de funcionar) disminuye mientras que el nivel de CUGBP1 libre (capaz de funcionar) aumenta. Un aumento en los niveles de CUGBP1 se asocia con el estado de enfermedad. Como tal, en modalidades, la hibridación del AC a al menos una porción de la repetición CUG expandida da como resultado niveles aumentados de MBNL1 libre (capaz de funcionar) y/o una disminución en los niveles de CUGBP1 libre (capaz de funcionar).

La disminución de la capacidad del transcrito objetivo para unirse a la una o más proteínas, puede reducir o inhibir la formación de focos de repeticiones CUG. Los transcritos que incluyen repeticiones de nucleótidos expandidas (por ejemplo, repeticiones CUG expandidas) pueden transcribirse y después secuestrarse en el núcleo. Dentro del núcleo, los transcritos secuestrados pueden formar agregados. Las proteínas que se unen al transcrito pueden entonces nuclearse en el transcrito secuestrado y/o el agregado de transcrito secuestrado, formando de esta manera focos de repeticiones de nucleótidos expandidas (por ejemplo, repeticiones CUG). Los focos de repeticiones CUG pueden ser visibles mediante el uso de microscopía. En modalidades, disminuir la capacidad del transcrito objetivo para unirse a la una o más proteínas, puede reducir o inhibir la formación de agregados que incluyen el transcrito objetivo. En modalidades, disminuir la capacidad del transcrito

objetivo para unirse a la una o más proteínas, puede reducir o inhibir la nucleación de la una o más proteínas en el transcrito objetivo, en una región de horquilla bicatenaria de un transcrito o en un agregado de transcritos objetivo. En modalidades donde el transcrito objetivo es la DMPK, disminuir la capacidad del transcrito objetivo de DMPK1 para unirse a MNBL1 puede reducir o inhibir la nucleación de MNBL1 en un transcrito objetivo de DMPK1 o un agregado de transcrito objetivo de DMPK1. En modalidades, la hibridación del AC al transcrito objetivo puede dar como resultado la inhibición o reducción en la formación de focos nucleares de repeticiones CUG. En modalidades, la hibridación del AC al transcrito objetivo puede dar como resultado la inhibición o reducción en la formación de focos nucleares de repeticiones CUG formados a partir de un transcrito objetivo de DMPK, TCF4, JPH3 y/o ATXN8OS /ATXN8.

En modalidades, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación del nivel, la expresión y/o la actividad de uno o más genes aguas abajo. Por ejemplo, la hibridación del AC al transcrito objetivo puede usarse para inducir la degradación del transcrito objetivo o bloquear estéricamente o disminuir la afinidad del transcrito objetivo por una o más proteínas, permitiendo de esta manera que la una o más proteínas que se secuestraron por el transcrito objetivo regulen la expresión, el nivel y/o la actividad de los genes aguas abajo. Por ejemplo, en modalidades, la una o más proteínas pueden incluir una proteína que está involucrada en la regulación del corte y empalme de uno o más transcritos aguas abajo (transcritos que no contienen una repetición CUG expandida). En modalidades, el corte y empalme de los transcritos aguas abajo se altera cuando la proteína involucrada en el corte y empalme se une y se secuestra en el transcrito objetivo. Por ejemplo, la alteración del corte y empalme puede incluir la exclusión de uno o más exones o la inclusión de uno o más intrones en un transcrito, conduciendo de esta manera a la expresión de diversas isoformas de proteínas. En modalidades, la alteración del corte y empalme puede dar como resultado la inclusión de un exón y/o intrón que incluye un codón de parada prematuro, resultando de esta manera en una isoforma truncada que puede tener o no tener actividad perjudicial. La alteración del corte y empalme del transcrito génico aguas abajo puede conducir a un cambio en el nivel, el plegado y/o la actividad del producto génico aguas abajo que puede asociarse con un fenotipo de enfermedad. Cuando no se une al transcripto objetivo que comprende la repetición CUG expandida, la proteína involucrada en el corte y empalme es libre de regular el corte y empalme, lo que puede dar como resultado una corrección (o rescate) del corte y empalme del transcripto génico aguas abajo, restaurando de esta manera al menos parcialmente el nivel, el plegado y/o la actividad de proteína del producto génico aguas abajo asociado con un fenotipo sano.

En modalidades, la hibridación de AC al transcrito objetivo puede dar como resultado la modulación del corte y empalme de transcritos génicos aguas abajo que está regulado por proteínas que son secuestradas por el transcrito objetivo en un estado de enfermedad asociado con repeticiones de nucleótidos expandidas (por ejemplo, repeticiones de trinucleótidos expandidas). En la enfermedad asociada con repeticiones de trinucleótidos expandidas, los transcritos génicos aguas abajo se procesan con frecuencia erróneamente, por ejemplo, se empalman erróneamente. Este corte y empalme erróneo de los transcritos génicos aguas abajo puede conducir a productos génicos que se destruyen antes de la traducción o se traducen a proteínas que tienen estructura y/o función aberrantes. Por ejemplo, el secuestro de proteínas que regulan el procesamiento de transcritos génicos aguas abajo puede conducir a la inclusión de exones y/o intrones con codones de parada prematuros, la inclusión de intrones, la exclusión de exones y/o la inclusión de exones alternativos, que pueden conducir a un transcrito y/o producto génico que se destruye antes de la traducción o que se traduce en un producto génico con función aberrante. El cambio en los niveles del transcrito génico y/o el producto génico aguas abajo y/o la estructura y/o función aberrantes de los productos génicos aguas abajo se asocia con fenotipos de enfermedades de trinucleótidos expandidos. En modalidades, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación de la inclusión de exones, exclusión de exones, inclusión de intrones y/o exclusión de intrones en transcritos aguas abajo cuyo corte y empalme está regulado por proteínas que se secuestran al transcrito objetivo durante un estado de enfermedad. Como tal, la hibridación del AC al transcrito objetivo puede dar como resultado la regulación positiva de los isómeros de proteínas y/o transcritos aguas abajo asociados con un fenotipo sano. De manera similar, la hibridación del AC al transcrito objetivo puede dar como resultado la regulación negativa (por ejemplo, supresión) de los transcritos y/o isómeros de proteínas aguas abajo asociados con un fenotipo de enfermedad.

En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación del corte y empalme de los transcritos génicos aguas abajo que están regulados por proteínas que son secuestradas por el transcrito objetivo de DMPK durante un estado de enfermedad. En la DM1, varios transcritos génicos aguas abajo se conducen a un empalme erróneo. Los genes con empalme erróneo se asocian con el fenotipo de la enfermedad. Como tal, la modulación del corte y empalme del gen puede incluir corregir (por ejemplo, rescatar) el corte y empalme de genes para dar como resultado productos génicos de genes aguas abajo que se asocian con un fenotipo sano. En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación del corte y

empalme de transcritos génicos aguas abajo que son regulados por MNBL1, un regulador del corte y empalme que es secuestrado por el transcrito objetivo de DMPK durante un estado de enfermedad. En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación del corte y empalme de transcritos génicos aguas abajo que son regulados por CUGBP1, una proteína cuya actividad se ve afectada por repeticiones CUG expandidas. En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC al transcrito objetivo puede dar como resultado el procesamiento correcto (por ejemplo, corte y empalme) de genes aguas abajo que son regulados por MNBL1 y/o CUGBP1. En modalidades donde el transcrito objetivo es la DMPK, la hibridación del AC al transcrito objetivo puede dar como resultado la modulación del corte y empalme de genes aguas abajo que incluyen, pero no se limitan a, 4833439L19Rik, Abcc9, Atp2a1, Arhgef10, Arhgap28, Armcx6, Angel1, Best3, Bin1, Brd2, Cacna1s, Cacna2d1, Cpd, Cpeb3, Ccp1, Clasp1, Clc-1, Clcn1, Clk4, Cpeb2, Camk2g, Capzb, Copz2, Coch, cTNT, Ctu2, Cyp2s1, Dctn4, Dnm1l, Eya4, Efna3, Efna2, Fbxo31, Fbxo21, Frem2, Fgd4, Fuca1, Fn1, Gogla4, Gpr37l1, Greb1, Heg1, Insr, Impdh2, IR, Itgav, Jag2, Klcl1, Kcan6, Kif13a, Ldb3, Lrrfip2, Mapt, Macf1, Map3k4, Mapkap1, Mbnl1, Mllt3, Mbnl2, Mef2c, Mpdz, Mrpl1, Mxra7, Mybpc1, Myo9a, Ncapd3, Ngfr, Ndr3, Ndufv3, Neb, Nfix, Numa1, Opa1, Pacsin2, Pcolce, Pdlm3, Pla2g15, Phactr4, Phka1, Phtf2, Ppp1r12b, Ppp3cc, Ppp1cc, Ramp2, Rapgef1, Rur1, Ryr1, Sorcs2, Spsb4, Scube2, Sema6c, Sfc8a3, Slain2, Sorbs1, Spag9, Tmem28, Tacc1, Tacc2, Ttc7, Tnik, Tnfrsf22, Tnfrsf25, Trappc9, Trim55, Ttn, Txnl4a, Txlnb, Ube2d3, Vsp39, o cualquiera de sus combinaciones.

El corte y empalme erróneo de muchos de los transcritos génicos aguas abajo mencionados anteriormente da como resultado fenotipos específicos de la enfermedad DM1. Por ejemplo, MNBL1 es un factor de corte y empalme con pérdida de función en DM1 debido a la inclusión del exón 5. MNBL1 es secuestrado por la expansión de CUG en DMPK y forma focos nucleares de ARN. Adicionalmente, SOS1 promueve la activación de Ras para regular positivamente la vía de señalización de RAS/MAPK. En la DM1, se excluye el exón 25 de SOS1, lo que conduce a la inhibición de las vías de hipertrofia muscular. En la DM1, IR/INSR tiene exclusión del exón 11, lo que da como resultado niveles más altos de la isoforma no muscular de baja señalización y disminución de la respuesta metabólica a la insulina en la DM1 (resistencia a la insulina). De manera similar, se observa la exclusión del exón 78 de DMD en la DM. La exclusión del exón 78 da como resultado un transcrito fuera de marco en el dominio C-terminal. Esta proteína mutada se expresa en pacientes con DM1 y se asocia con un mecanismo responsable del desgaste muscular en los

pacientes. Se requiere BIN1 para una formación adecuada de los túbulos T musculares (acoplamiento de EC). La exclusión del exón 11 produce una isoforma inactiva y se encuentra en los pacientes con DM1. LDB3 interactúa con la α -actinina en el disco Z en el músculo estriado y mantiene la estructura muscular. La inclusión del exón 11 de LDB3 detectada en la DM1 da como resultado la reducción de la afinidad por la proteína cinasa C (PKC). En consecuencia, la PKC se vuelve hiperactiva en la DM1. En modalidades, la modulación de uno o más genes aguas abajo da como resultado la corrección, o rescate, del corte y empalme de transcritos que se asocia con un fenotipo sano. Como tal, en modalidades, la hibridación del AC al transcrito objetivo de DMPK, da como resultado el rescate del corte y empalme erróneo de genes/transcritos aguas abajo, reduciendo de esta manera el nivel de genes/transcritos aguas abajo asociados con un fenotipo de enfermedad. Como tal, en modalidades, la hibridación del AC al transcrito objetivo de DMPK, da como resultado el rescate del corte y empalme erróneo de genes/transcritos aguas abajo, aumentando de esta manera el nivel de genes/transcritos aguas abajo asociados con un fenotipo sano.

En modalidades, el AC inhibe la expresión del transcrito objetivo. En modalidades, el AC inhibe la expresión del transcrito objetivo al bloquear la maquinaria de procesamiento de pre-ARNm y/o la maquinaria de traducción para acceder y/o completar la traducción y/o el procesamiento de pre-ARNm. En modalidades, el AC inhibe la expresión del transcrito objetivo al inducir la degradación del transcrito objetivo, por ejemplo, a través de las vías de la ARNasa H.

Estructura del AC

El AC incluye un oligonucleótido y/o un oligonucleósido. Los oligonucleótidos y/u oligonucleósidos son nucleótidos o nucleósidos enlazados a través de enlaces internucleósidos. Los nucleósidos incluyen un azúcar pentosa (por ejemplo, ribosa o desoxirribosa) y una base nitrogenada unida covalentemente al azúcar. Las bases de origen natural (o bases tradicionales) que se encuentran en el ADN y/o ARN son adenina (A), guanina (G), timina (T), citosina (C) y uracilo (U). Los azúcares de origen natural (o azúcares tradicionales) que se encuentran en el ADN y/o ARN son desoxirribosa (ADN) y ribosa (ARN). El enlace de nucleósidos de origen natural (o enlace internucleósido tradicional) es un enlace fosfodiéster. En modalidades, los AC de la presente descripción pueden tener todos azúcares, bases y enlaces internucleósidos naturales.

Los nucleósidos modificados químicamente se usan rutinariamente para su incorporación en compuestos antisentido para potenciar una o más propiedades, tales como resistencia a nucleasas, farmacocinética, o afinidad por un ARN objetivo. En modalidades, los AC de la presente descripción pueden tener uno o más nucleósidos modificados. En modalidades, los AC de la presente descripción

pueden tener uno o más azúcares modificados. En modalidades, las AC de la presente descripción pueden tener una o más bases modificadas. En modalidades, las AC de la presente descripción pueden tener uno o más enlaces internucleósidos modificados.

En general, una nucleobase es cualquier grupo que contiene uno o más átomos o grupos de átomos capaces de unirse por un hidrógeno a una base de otro ácido nucleico. Además de las nucleobases "no modificadas" o "naturales" (A, G, T, C y U) muchas nucleobases modificadas o miméticos de nucleobases que son conocidos para los expertos en la técnica son compatibles con los compuestos descritos en la presente descripción. Generalmente, una nucleobase modificada se refiere a una nucleobase que es bastante similar en estructura a la nucleobase parental, tal como, por ejemplo, una 7-deaza purina, una 5-metil citosina, 2-tio-dT (Figura 2) o una abrazadera G. Generalmente, un mimético de nucleobase es una nucleobase que incluye una estructura que es más complicada que una nucleobase modificada, tal como, por ejemplo, un mimético de nucleobase de fenoxazina tricíclica. Los métodos para la preparación de las nucleobases modificadas mencionadas anteriormente se conocen bien por los expertos en la técnica.

En modalidades, el AC puede incluir uno o más nucleósidos que tienen un resto de azúcar modificado. En modalidades, el azúcar furanosilo de un nucleósido natural puede tener una modificación 2', modificaciones para fabricar un nucleósido restringido, y otras (ver la Figura 2). Por ejemplo, en modalidades, el anillo de azúcar furanosilo de un nucleósido natural puede modificarse de varias maneras que incluyen, pero no se limitan a, la adición de un grupo sustituyente, la formación de puentes de dos átomos del anillo no geminales para formar un ácido nucleico bicíclico (BNA) o un ácido nucleico bloqueado; el intercambio del oxígeno del anillo furanosilo con C o N; y/o la sustitución de un átomo o grupo tal (ver Figura 2). Los azúcares modificados se conocen bien y pueden usarse para aumentar o disminuir la afinidad del AC por su secuencia de nucleótidos objetivo. Los azúcares modificados también pueden usarse para aumentar la resistencia del AC a las nucleasas. Los azúcares también pueden reemplazarse por grupos miméticos de azúcar, entre otros. En modalidades, uno o más azúcares de los nucleósidos del AC se reemplazan con un anillo metilenmorfolino como se muestra como 19 en la Figura 2.

En modalidades, el AC incluye uno o más nucleósidos que incluyen un azúcar modificado bicíclico (BNA; a veces llamados ácidos nucleicos con puente). Los ejemplos de BNA incluyen, pero no se limitan a, LNA (puente 4'-(CH₂)-O-2'), 2'-tio-LNA (puente 4'-(CH₂)-S-2'), 2'-amino-LNA (puente 4'-(CH₂)-NR-2'), ENA (puente 4'-(CH₂)₂-O-2'), BNA con puente 4'-(CH₂)₃-2', BNA" con puente 4'-(CH₂CH(CH₃))-2' cEt (puente 4'-(CH(CH₃)-O-2') y BNA cMOE (puente 4'-(CH(CH₂OCH₃)-O-2').

Los BNA se han preparado y descrito en la literatura de patentes, así como también en la literatura científica (Srivastava, y otros, J. Am. Chem. Soc. (2007), publicación en línea de ACS Advanced, 10.1021/ja071106y; Albaek y otros, J. Org. Chem. (2006), 71, 7731 -7740; Fluiter, y otros, Chembiochem (2005), 6, 1,104-1,109; Singh y otros, Chem. Commun. (1998), 4, 455-456; Koshkin y otros, Tetrahedron (1998), 54, 3,607-3,630; Wahlestedt y otros, Proc. Natl. Acad. Sci. U.S.A. (2000), 97, 5,633-5,638; Kumar y otros, Bioorg. Medio Chem. Lett. (1998), 8, 2,219-2,222; documento WO 94/14226; documento WO 2005/021570; Singh y otros, J. Org. Chem. (1998), 63, 10,035-10,039, documento WO 2007/090071; Patentes de Estados Unidos núms. 7,053,207; 6,268,490; 6,770,748; 6,794,499; 7,034,133; y 6,525,191; y publicación antes de la concesión de Estados Unidos núms. 2004-0171570; 2004-0219565; 2004-0014959; 2003-0207841; 2004-0143114; y 20030082807).

En modalidades, el AC incluye uno o más nucleósidos que incluyen un ácido nucleico bloqueado (LNA). En los LNA, el grupo hidroxilo 2' del anillo de azúcar ribosilo se une al átomo de carbono 4' del anillo de azúcar formando de esta manera un enlace 2'-C,4'-C-oximetileno para formar el resto de azúcar bicíclico (revisado en Elayadi y otros, Curr. Opinion Invens. Drugs (2001), 2, 558-561; Braasch y otros, Chem. Biol. (2001), 8 1-7; y Orum y otros, Curr. Opinion Mol. Ther. (2001), 3, 239-243; ver también las patentes de Estados Unidos: 6,268,490 y 6,670,461). El enlace puede ser un grupo metileno (-CH₂-) que enlaza el átomo de oxígeno 2' y el átomo de carbono 4', para lo cual se usa el término LNA para el resto bicíclico; en el caso de un grupo etileno en esta posición, se usa el término ENATM (Singh y otros, Chem. Commun. (1998), 4, 455-456; ENATM; Morita y otros, Bioorganic Medicinal Chemistry (2003), 11, 2,211-2,226). El LNA y otros análogos de azúcar bicíclicos muestran estabilidades térmicas del dúplex muy altas con ADN y ARN complementarios (T_m = +3 a +10 °C), estabilidad hacia la degradación exonucleolítica 3' y buenas propiedades de solubilidad. Se han descrito oligonucleótidos antisentido potentes y no tóxicos que contienen LNA (Wahlestedt y otros, Proc. Natl. Acad. Sci. U.S.A. (2000), 97, 5,633-5,638).

Un isómero de LNA que también se ha estudiado es el alfa-L-LNA, que se ha demostrado que tiene una estabilidad superior contra una 3'-exonucleasa. Los alfa-L-LNA se incorporaron en gapmers y quimeras antisentido que mostraron una potente actividad antisentido (Frieden y otros, Nucleic Acids Research (2003), 21, 6,365-6,372).

Se ha descrito la síntesis y preparación de los monómeros de LNA adenina, citosina, guanina, 5-metil-citosina, timina y uracilo, junto con su oligomerización y propiedades de reconocimiento de ácidos nucleicos (Koshkin y otros, Tetrahedron, 1998, 54, 3,607-3,630). Los LNA y la preparación

de estos se describe además en los documentos WO 98/39352 y WO 99/14226.

También se han preparado los análogos de LNA fosforotioato-LNA y 2'-tio-LNA (Kumar y otros, Bioorg. Med. Chem. Lett., 1998, 8, 2,219-2,222). También se ha descrito la preparación de análogos de LNA que contienen varios dúplex de oligodesoxirribonucleótidos como sustratos para polimerasas de ácidos nucleicos (Wengel y otros, documento WO 99/14226). Además, se ha descrito la síntesis del 2'-amino-LNA, un análogo de oligonucleótidos de alta afinidad restringido conformacionalmente (Singh y otros, J. Org. Chem. (1998), 63, 10,035-10,039). Además, se han preparado 2'-amino- y 2'-metilamino-LNA, y se ha informado previamente la estabilidad térmica de sus dúplex con hebras complementarias de ARN y ADN.

Los métodos para las preparaciones de azúcares modificados se conocen bien por los expertos en la técnica. Algunas patentes y publicaciones representativas que enseñan la preparación de dichos azúcares modificados incluyen, pero no se limitan a, las patentes de Estados Unidos: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; 5,700,920; y 6,600,032; y el documento WO 2005/121371.

Enlaces internucleósidos

En la presente descripción se describen grupos de enlace internucleósidos que enlazan los nucleósidos o unidades de monómero de nucleósido modificados de cualquier otra manera entre sí, formando de esta manera un oligonucleótido y/o un oligonucleótido que contiene un AC. Los AC pueden incluir enlaces internucleósidos de origen natural, enlaces internucleósidos no naturales, o ambos.

En el ADN y ARN de origen natural, el grupo de enlace internucleósido es un fosfodiéster que enlaza covalentemente nucleósidos adyacentes entre sí para formar un compuesto polimérico lineal. En el ADN y ARN de origen natural, el fosfodiéster se enlaza al resto hidroxilo 2', 3' o 5' del azúcar. Dentro de los oligonucleótidos, los grupos fosfato se denominan comúnmente como que forman la cadena principal internucleósidos del oligonucleótido. En el ADN y ARN de origen natural, el enlace o cadena principal de ARN y ADN, es un enlace fosfodiéster de 3' a 5'. En modalidades, los grupos de enlace internucleósidos de los AC son fosfodiésteres. En modalidades, los grupos de enlace internucleósidos de los AC son enlaces fosfodiéster 3' a 5'.

Las dos clases principales de grupos de enlace internucleósidos no naturales se definen por la presencia o ausencia de un átomo de fósforo. Los enlaces internucleósidos representativos que contienen fósforo incluyen, pero no se limitan a, fosfotriésteres, metilfosfonatos, fosforamidato y

fosforotioatos. Los grupos de enlace internucleósidos representativos que no contienen fósforo incluyen, pero no se limitan a, metilenmetilimino (-CH₂-N(CH₃)-O-CH₂-), tiodiéster (-O-C(O)-S-), tionocarbamato (-O-C(O)(NH)-S-); siloxano (-O-Si(H₂-O-); y N,N'-dimetilhidrazina (-CH₂-N(CH₃)-N(CH₃)-). Los AC que tienen grupos de enlace internucleósidos con fósforo se denominan oligonucleótidos. Los compuestos antisentido que tienen grupos de enlace internucleósidos que no contienen fósforo se denominan oligonucleósidos. Los enlaces internucleósidos modificados, en comparación con los enlaces fosfodiéster naturales, pueden usarse para alterar, típicamente, aumentar, la resistencia a las nucleasas del compuesto antisentido. Los enlaces internucleósidos que tienen un átomo quiral pueden prepararse como racémicos, quirales, o como una mezcla. Los enlaces internucleósidos quirales representativos incluyen, pero no se limitan a, alquilfosfonatos y fosforotioatos. Los métodos de preparación de enlaces que contienen fósforo y que no contienen fósforo se conocen bien por los expertos en la técnica.

En modalidades, dos o más nucleósidos que tienen azúcares modificados y/o nucleobases modificadas pueden unirse mediante el uso de un fosforamidato. En modalidades, dos o más nucleósidos que tienen un anillo metilenmorfolino pueden conectarse a través de un enlace internucleósido de fosforamidato.

Los compuestos antisentido que incluyen nucleobases con un anillo metilenmorfolino que se enlazan a través del enlace internucleósido de fosforamidato pueden denominarse oligómeros de morfolino fosforamidato (PMO).

Grupos conjugados

En modalidades, los AC se modifican mediante la unión covalente de uno o más grupos conjugados. En general, los grupos conjugados modifican una o más propiedades del AC unido, lo que incluye, pero no se limita a, farmacodinámica, farmacocinética, unión, absorción, distribución celular, captación celular, carga y aclaramiento. Los grupos conjugados se usan rutinariamente en las técnicas químicas y se enlazan directamente o mediante un resto de enlace o grupo de enlace opcional a un compuesto parental tal como un AC. Los grupos conjugados incluyen, sin limitación, intercaladores, moléculas indicadoras, poliaminas, poliamidas, polietilenglicoles, tioéteres, poliéteres, colesteroles, tiocolesteroles, restos de ácido cólico, folato, lípidos, fosfolípidos, biotina, fenazina, fenantridina, antraquinona, adamantano, acridina, fluoresceínas, rodaminas, cumarinas y colorantes. En modalidades, el grupo conjugado es un polietilenglicol (PEG), y el PEG se conjuga con ya sea el AC o el CPP (CPP analizado en otra parte en la presente descripción).

En modalidades, los grupos conjugados incluyen restos lipídicos tales como un resto de colesterol

(Letsinger y otros, Proc. Natl. Acad. Sci. USA (1989), 86, 6,553); ácido cólico (Manoharan y otros, Bioorg. Med. Chem. Lett. (1994), 4, 1,053); un tioéter, por ejemplo, hexil-S-tritilol (Manoharan y otros, Ann. N.Y. Acad. Sci. (1992), 660, 306; Manoharan y otros, Bioorg. Med. Chem. Lett. (1993), 3, 2,765); un tiocolesterol (Oberhauser y otros, Nucl. Acids Res. (1992), 20, 533); una cadena alifática, por ejemplo, residuos dodecandiol o undecilo (Saison-Behmoaras y otros, EMBO J. (1991), 10, 111; Kabanov y otros, FEBS Lett. (1990), 259, 327; Svinarchuk y otros, Biochimie (1993), 75, 49); un fosfolípido, por ejemplo, di-hexadecil-rac-glicerol o trietilamonio-1,2-di-O-hexadecil-rac-glicerol-3-H-fosfonato (Manoharan y otros, Tetrahedron Lett. (1995), 36, 3,651; Shea y otros, Nucl. Acids Res. (1990), 18, 3777); una poliamina o una cadena de polietilenglicol (Manoharan y otros, Nucleosides & Nucleotides (1995), 14, 969); ácido acético adamantano (Manoharan y otros, Tetrahedron Lett. (1995), 36, 3,651); un resto palmitilo (Mishra y otros, Biochim. Biophys. Acta. (1995), 1,264, 229); o un resto octadecilamina o hexilamino-carbonil-oxicolesterol (Crooke y otros, J. Pharmacol. Exp. Ther. (1996), 277,923).

Tipos de compuestos antisentido

Pueden usarse diversos tipos de AC, por ejemplo, lo que incluye un oligonucleótido antisentido, ARNip, microARN, antagomir, aptámero, ribozima, supermir, imitador de miARN, inhibidor de miARN, o sus combinaciones.

Oligonucleótidos antisentido

En diversas modalidades, el compuesto antisentido (AC) es un oligonucleótido antisentido (ASO) que es complementario a una secuencia de nucleótidos objetivo. El término "oligonucleótido antisentido (ASO)" o simplemente "antisentido" pretende incluir oligonucleótidos que son complementarios a una secuencia de nucleótidos objetivo. El término abarca además los ASO que pueden no ser totalmente complementarios a la secuencia de nucleótidos objetivo deseada. Los ASO incluyen hebras únicas de ADN y/o ARN que son complementarias a una secuencia de nucleótidos objetivo elegida o a un gen objetivo. Los ASO pueden incluir una o más bases de ADN y/o ARN modificados, azúcares modificados y/o enlaces internucleósidos no naturales. En modalidades, los ASO pueden incluir uno o más enlaces internucleósidos fosforamidatos. En modalidades, el ASO es un oligómero de morfolino fosforamidato (PMO). Los ASO pueden tener cualquier característica, ser de cualquier longitud, unirse a cualquier secuencia de nucleótidos objetivo y/o elemento de secuencia, y efectuar cualquier mecanismo como se describe con relación a un AC.

Se ha demostrado que los oligonucleótidos antisentido son efectivos como inhibidores dirigidos de la síntesis de proteínas y, en consecuencia, pueden usarse para inhibir específicamente la síntesis de

proteínas por un gen dirigido. La eficacia del ASO para inhibir la síntesis de proteínas está bien establecida. Hasta la fecha, estos compuestos han demostrado ser prometedores en varios modelos in vitro e in vivo, lo que incluye modelos de enfermedad inflamatoria, cáncer y VIH (Agrawal, Trends in Biotech. (1996), 14:376-387). Los antisentido también pueden afectar la actividad celular al hibridarse específicamente con ADN cromosómico.

Los métodos para producir oligonucleótidos antisentido se conocen en la técnica, y pueden adaptarse fácilmente para producir un oligonucleótido antisentido que se dirige a cualquier secuencia de polinucleótidos. La selección de secuencias de oligonucleótidos antisentido específicas para una secuencia objetivo dada se basa en el análisis de la secuencia objetivo elegida y la determinación de la estructura secundaria, la T_m , la energía de unión y la estabilidad relativa. Los oligonucleótidos antisentido pueden seleccionarse basado en su incapacidad relativa para formar dímeros, horquillas u otras estructuras secundarias que reducirían o prohibirían la unión específica al ARNm objetivo en una célula huésped. Las regiones objetivo del ARNm incluyen aquellas regiones en o cerca del codón de inicio de la traducción AUG y aquellas secuencias que son sustancialmente complementarias a las regiones 5' del ARNm. Estos análisis de la estructura secundaria y las consideraciones de selección del sitio objetivo pueden realizarse, por ejemplo, mediante el uso de la v.4 del software de análisis de cebadores OLIGO (Molecular Biology Insights) y/o el software del algoritmo BLASTN 2.0.5 (Altschul y otros, Nucleic Acids Res. 1997, 25(17):3,389-402).

Interferencia del ARN

En modalidades, el AC incluye una molécula que media la interferencia del ARN (ARNi). Como se usa en la presente, la frase "media el ARNi" se refiere a la capacidad de silenciar, de una manera específica de secuencia, un transcrito objetivo. Aunque no se desea limitarse por la teoría, se cree que el silenciamiento usa la maquinaria o proceso de ARNi y un ARN guía, por ejemplo, un compuesto de ARNip de aproximadamente 21 a aproximadamente 23 nucleótidos. En modalidades, el AC se dirige al transcrito objetivo para su degradación. Como tal, en modalidades, la molécula de ARNi puede usarse para interrumpir la expresión de un gen o polinucleótido de interés. En modalidades, la molécula de ARNi se usa para inducir la degradación del transcrito objetivo, tal como un pre-ARNm o un ARNm maduro.

En modalidades, el AC incluye un ARN de interferencia pequeño (ARNip) que induce una respuesta de ARNi.

Los ARN de interferencia pequeños (ARNip) son dúplex de ácidos nucleicos normalmente de aproximadamente 16 a aproximadamente 30 nucleótidos de largo que pueden asociarse con un

complejo multiproteínas citoplasmático conocido como complejo de silenciamiento inducido por ARNi (RISC). El RISC cargado con ARNip media la degradación de los transcritos homólogos, por lo tanto, el ARNip puede diseñarse para atenuar la expresión de proteínas con alta especificidad. A diferencia de otras tecnologías antisentido, el ARNip funciona a través de un mecanismo natural que evolucionó para controlar la expresión génica a través de ARN no codificante. Una variedad de reactivos de ARNi, lo que incluye ARNip dirigidos a objetivos clínicamente relevantes, están actualmente en desarrollo farmacéutico, como se describe, por ejemplo, en de Fougères, A. y otros, *Nature Reviews* (2007) 6:443-453.

Aunque las primeras moléculas de ARNi descritas fueron híbridos ARN:ARN que incluyen tanto una cadena sentido de ARN como una cadena antisentido de ARN, ahora se ha demostrado que los híbridos ADN sentido:ARN antisentido, los híbridos ARN sentido:ADN antisentido y los híbridos ADN:ADN son capaces de mediar ARNi (Lamberton, J.S. y Christian, A.T., *Molecular Biotechnology* (2003), 24:111-119). En modalidades, se usan moléculas de ARNi que incluyen cualquiera de estos tipos diferentes de moléculas bicatenarias. Además, se entiende que las moléculas de ARNi pueden usarse e introducirse en las células en una variedad de formas. En consecuencia, como se usa en la presente, las moléculas de ARNi abarcan todas y cada una de las moléculas capaces de mediar un ARNi en las células, lo que incluye, pero no se limita a, oligonucleótidos bicatenarios que incluyen dos hebras separadas, es decir, una hebra sentido y una hebra antisentido, por ejemplo, ARN de interferencia pequeño (ARNip); oligonucleótido bicatenario que incluye dos hebras separadas que se unen entre sí mediante un enlazador no nucleotídico; oligonucleótidos que incluyen un lazo en horquilla de secuencias complementarias, que forma una región bicatenaria, por ejemplo, moléculas de ARNihp, y vectores de expresión que expresan uno o más polinucleótidos capaces de formar un polinucleótido bicatenario solo o en combinación con otro polinucleótido.

Un "compuesto de ARNip monocatenario", como se usa en la presente, es un compuesto de ARNip que se compone de una sola molécula. Puede incluir una región dúplex, formada por emparejamiento intracatenario, por ejemplo, puede ser, o incluir, una estructura en horquilla o de mango de sartén. Los compuestos de ARNip monocatenarios pueden ser antisentido con respecto a la molécula objetivo.

Un compuesto de ARNip monocatenario puede ser lo suficientemente largo que pueda entrar en el RISC y participar en la escisión mediada por RISC de un ARNm objetivo. Un compuesto de ARNip monocatenario tiene al menos aproximadamente 14, al menos aproximadamente 15, al menos aproximadamente 20, al menos aproximadamente 25, al menos aproximadamente 30, al menos

aproximadamente 35, al menos aproximadamente 40 o hasta aproximadamente 50 nucleótidos de longitud. En ciertas modalidades, el ARNip monocatenario tiene menos de aproximadamente 200, aproximadamente 100 o aproximadamente 60 nucleótidos de longitud.

5 Los compuestos de ARNip en horquilla pueden tener una región dúplex igual o al menos aproximadamente de 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24 o aproximadamente 25 pares de nucleótidos. La región dúplex puede ser igual o menor que aproximadamente 200, aproximadamente 100 o aproximadamente 50 pares de nucleótidos de longitud. En ciertas modalidades, los intervalos para la región dúplex tienen de aproximadamente 15
10 a aproximadamente 30, de aproximadamente 17 a aproximadamente 23, de aproximadamente 19 a aproximadamente 23, y de aproximadamente 19 a aproximadamente 21 pares de nucleótidos de longitud. La horquilla puede tener un saliente monocatenario o región terminal no emparejada. En ciertas modalidades, los salientes tienen de aproximadamente 2 a aproximadamente 3 nucleótidos de longitud. En modalidades, el saliente está en el mismo lado de la horquilla y en modalidades, en
15 el lado antisentido de la horquilla.

Un "compuesto de ARNip bicatenario", como se usa en la presente, es un compuesto de ARNip que incluye más de una, y en algunos casos dos, hebras en las que la hibridación entre cadenas puede formar una región de estructura dúplex.

La hebra antisentido de un compuesto de ARNip bicatenario puede ser igual o al menos
20 aproximadamente 14, aproximadamente 15, aproximadamente 16 aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 40 o aproximadamente 60 nucleótidos de longitud. Puede ser igual o menor que aproximadamente 200, aproximadamente 100 o aproximadamente 50 nucleótidos de longitud. Los intervalos pueden tener de aproximadamente 17 a aproximadamente
25 25, de aproximadamente 19 a aproximadamente 23, y de aproximadamente 19 a aproximadamente 21 nucleótidos de longitud. Como se usa en la presente, el término "hebra antisentido" significa la hebra de un compuesto de ARNip que es suficientemente complementaria a una molécula objetivo, por ejemplo, la secuencia de nucleótidos objetivo de un transcrito objetivo.

La hebra sentido de un compuesto de ARNip bicatenario puede ser igual o al menos
30 aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 40 o aproximadamente 60 nucleótidos de longitud. Puede

ser igual o menor que aproximadamente 200, aproximadamente 100 o aproximadamente 50, nucleótidos de longitud. Los intervalos pueden tener de aproximadamente 17 a aproximadamente 25, de aproximadamente 19 a aproximadamente 23, y de aproximadamente 19 a aproximadamente 21 nucleótidos de longitud.

- 5 La porción bicatenaria de un compuesto de ARNip bicatenario puede ser igual o al menos aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24, aproximadamente 25, aproximadamente 30, aproximadamente 40 o aproximadamente 60 pares de nucleótidos de longitud.
- 10 Puede ser igual o menor que aproximadamente 200, aproximadamente 100 o aproximadamente 50, pares de nucleótidos de longitud. Los intervalos pueden ser de aproximadamente 15 a aproximadamente 30, de aproximadamente 17 a aproximadamente 23, de aproximadamente 19 a aproximadamente 23, y de aproximadamente 19 a aproximadamente 21 pares de nucleótidos de longitud.
- 15 En modalidades, el compuesto de ARNip es lo suficientemente grande como para que pueda escindirse por una molécula endógena, por ejemplo, por Dicer, para producir compuestos de ARNip más pequeños, por ejemplo, agentes ARNip.
- Las hebras sentido y antisentido pueden elegirse de manera que el compuesto de ARNip bicatenario incluya una sola hebra o región no emparejada en uno o ambos extremos de la molécula. Por tanto,
- 20 un compuesto de ARNip bicatenario puede contener hebras sentido y antisentido, emparejadas para contener un saliente, por ejemplo, uno o dos salientes 5' o 3', o un saliente 3' de 1 a 3 nucleótidos. Los salientes pueden ser el resultado de que una hebra sea más larga que la otra, o el resultado de escalonar dos hebras de la misma longitud. Algunas modalidades tendrán al menos un saliente en 3'. En modalidades, ambos extremos de una molécula de ARNip tendrán un saliente en 3'. En
- 25 modalidades, el saliente es de 2 nucleótidos.
- En modalidades, la longitud para la región dúplex es de aproximadamente 15 a aproximadamente 30, o aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, o aproximadamente 23 nucleótidos de longitud, por ejemplo, en el intervalo del compuesto de ARNipmc (ARNip con salientes pegajosos) analizado anteriormente. Los
- 30 compuestos de ARNipmc pueden parecerse en longitud y estructura a los productos naturales procesados con Dicer a partir de ARNipbc largos. También se incluyen modalidades en las que las dos hebras del compuesto de ARNipmc se enlazan, por ejemplo, se enlazan covalentemente. En

modalidades, se incluyen horquillas u otras estructuras monocatenarias que proporcionan una región bicatenaria, y salientes en 3'.

Los compuestos de ARNip descritos en la presente descripción, lo que incluye compuestos de ARNip bicatenario y compuestos de ARNip monocatenario pueden mediar el silenciamiento de un ARN objetivo, por ejemplo, ARNm, por ejemplo, un transcrito de un gen que codifica una proteína. Por conveniencia, dicho ARNm también se denomina en la presente descripción ARNm a silenciar. Dicho gen también se denomina gen objetivo. En general, el ARN a silenciar es un gen endógeno. En modalidades, un compuesto de ARNip es "suficientemente complementario" a un transcrito objetivo, de manera que el compuesto de ARNip silencia la producción de la proteína codificada por el ARNm objetivo. En modalidades, el compuesto de ARNip es "suficientemente complementario" a al menos una porción de un transcrito objetivo, de manera que el compuesto de ARNip silencia la producción del producto génico codificado por el transcrito objetivo. En otra modalidad, el compuesto de ARNip es "exactamente complementario" a una secuencia de nucleótidos objetivo (por ejemplo, una porción de un transcrito objetivo) de manera que la secuencia de nucleótidos objetivo y el compuesto de ARNip se hibridan, por ejemplo, para formar un híbrido formado exclusivamente por pares de bases Watson-Crick en la región de complementariedad exacta. Un "suficientemente complementario" a una secuencia de nucleótidos objetivo puede incluir una región interna (por ejemplo, de al menos aproximadamente 10 nucleótidos) que es exactamente complementaria a una secuencia de nucleótidos objetivo. Además, en ciertas modalidades, el compuesto de ARNip discrimina específicamente una diferencia de un solo nucleótido. En este caso, el compuesto de ARNip media solo el ARNi si se encuentra una complementariedad exacta en la región (por ejemplo, dentro de 7 nucleótidos de) la diferencia de un solo nucleótido.

Las aplicaciones terapéuticas del ARNi son extremadamente amplias, ya que los constructos de ARNip y miARN pueden sintetizarse con cualquier secuencia de nucleótidos dirigida contra una proteína objetivo. Hasta la fecha, los constructos de ARNip han mostrado la capacidad de regular negativamente de manera específica las proteínas objetivo tanto en modelos in vitro como in vivo, así como también en estudios clínicos

MicroARN

En modalidades, el AC incluye una molécula de microARN. Los microARN (miARN) son una clase altamente conservada de moléculas de ARN pequeñas que se transcriben a partir del ADN en los genomas de plantas y animales, pero no se traducen en proteínas. Los miARN procesados son moléculas de ARN monocatenario de 17-25 nucleótidos que se incorporan en el complejo de

silenciamiento inducido por ARN (RISC) y se han identificado como reguladores clave del desarrollo, la proliferación celular, la apoptosis y la diferenciación. Se cree que desempeñan un papel en la regulación de la expresión génica mediante la unión a la región no traducida 3' de ARNm específicos. El RISC media la regulación negativa de la expresión génica a través de la inhibición traslacional, la escisión del transcrito, o ambas. El RISC también está implicado en el silenciamiento transcripcional en el núcleo de una amplia gama de eucariotas.

Antagomirs

En modalidades, el AC es un antagomir. Los antagomirs son oligonucleótidos similares al ARN que albergan diversas modificaciones para la protección ante ARNasa y propiedades farmacológicas, tales como una captación celular y tisular potenciada. Se diferencian del ARN normal en, por ejemplo, la 2'-O-metilación completa del azúcar, la cadena principal de fosforotioato y, por ejemplo, un resto de colesterol en el extremo 3'. Los antagomirs pueden usarse para silenciar eficientemente los miARN endógenos mediante la formación de dúplex que incluyen el antagomir y el miARN endógeno, evitando de esta manera el silenciamiento génico inducido por miARN. Un ejemplo de silenciamiento de miARN mediado por un antagomir es el silenciamiento de miR-122, descrito en Krutzfeldt y otros, Nature (2005), 438: 685-689, que se incorpora expresamente como referencia en la presente descripción en su totalidad. Los ARN de antagomir pueden sintetizarse mediante el uso de protocolos estándar de síntesis de oligonucleótidos en fase sólida (solicitud de patente de Estados Unidos núms. ser. 11/502,158 y 11/657,341; cuyas descripciones se incorporan en la presente descripción como referencia).

Un antagomir puede incluir subunidades de monómeros conjugadas con ligando y monómeros para la síntesis de oligonucleótidos. Los monómeros se describen en la solicitud de Estados Unidos núm. 10/916,185. Un antagomir puede tener una estructura ZXY, tal como se describe en la solicitud PCT núm. PCT/US2004/07070. Un antagomir puede formar un complejo con un resto anfipático. Los restos anfipáticos para su uso con agentes oligonucleotídicos se describen en la solicitud PCT núm. PCT/US2004/07070.

Aptámeros

En modalidades, el AC incluye un aptámero. Los aptámeros son moléculas de ácido nucleico o péptido que se unen a una molécula de interés en particular con alta afinidad y especificidad (Tuerk y Gold, Science 249:505 (1990); Ellington y Szostak, Nature 346:818 (1990)). Se han producido con éxito aptámeros de ADN o ARN que se unen a muchas entidades diferentes de proteínas grandes a moléculas orgánicas pequeñas (Eaton, Curr. Opin. Chem. Biol. (1997), 1: 10-16; Famulok, Curr.

Opin. Struct. Biol. (1999), 9:324-9; y Hermann y Patel, Science (2000), 287:820-5). Los aptámeros pueden basarse en ARN o ADN y pueden incluir un ribointerruptor. Un ribointerruptor es una parte de una molécula de ARNm que puede unirse directamente a una molécula objetivo pequeña, y cuya unión al objetivo afecta la actividad del gen. Por tanto, un ARNm que contiene un ribointerruptor está directamente involucrado en la regulación de su propia actividad, en dependencia de la presencia o ausencia de su molécula objetivo. Generalmente, los aptámeros se manipulan por ingeniería genética a través de rondas repetidas de selección in vitro o de manera equivalente, SELEX (evolución sistémica de ligandos por enriquecimiento exponencial) para unirse a diversos objetivos moleculares tales como moléculas pequeñas, proteínas, ácidos nucleicos e incluso células, tejidos y organismos. El aptámero puede prepararse mediante cualquier método conocido, lo que incluye métodos sintéticos, recombinantes y de purificación, y puede usarse solo o en combinación con otros aptámeros específicos para el mismo objetivo. Además, el término "aptámero" también incluye "aptámeros secundarios", que contienen una secuencia consenso derivada de comparar dos o más aptámeros conocidos para un objetivo dado. En modalidades, el aptámero es un "aptámero intracelular", o "intramer", que reconoce específicamente objetivos intracelulares (Famulok y otros, Chem Biol. (2001),8(10):931-939; Yoon y Rossi, Adv. Drug Deliv. Rev. (2018), 134:22-35; cada uno incorporado como referencia en la presente descripción).

Ribozimas

En modalidades, el AC es una ribozima. Las ribozimas son complejos de moléculas de ARN que tienen dominios catalíticos específicos que poseen actividad endonucleasa (Kim y Cech, Proc. Natl. Acad. Sci. USA (1987),84(24):8,788-92; Forster y Symons, Cell (1987) 24, 49(2):211-20). Por ejemplo, un gran número de ribozimas aceleran las reacciones de transferencia de fosfoéster con un alto grado de especificidad, con frecuencia al escindir solo uno de varios fosfoésteres en un sustrato oligonucleotídico (Cech y otros, Cell (1981), 27(3 Pt 2):487-96; Michel y Westhof, J. Mol. Biol. (1990), 5, 216(3):585-610; Reinhold-Hurek y Shub, Nature (1992), 14, 357(6374): 173-6). Esta especificidad se ha atribuido a la necesidad de que el sustrato se una mediante interacciones específicas de emparejamiento de bases a la secuencia guía interna (IGS) de la ribozima antes de la reacción química.

Actualmente se conocen al menos seis variedades básicas de ARN enzimáticos de origen natural. Cada uno puede catalizar la hidrólisis de los enlaces fosfodiéster del ARN in trans (y por tanto puede escindir otras moléculas de ARN) en condiciones fisiológicas. En general, los ácidos nucleicos enzimáticos actúan al unirse primero a un ARN objetivo. Dicha unión se produce a través de la

porción de unión al objetivo de un ácido nucleico enzimático que se mantiene muy cerca de una porción enzimática de la molécula que actúa para escindir el ARN objetivo. Por tanto, el ácido nucleico enzimático primero reconoce y después se une a un ARN objetivo a través del emparejamiento de bases complementarias, y una vez unido al sitio correcto, actúa enzimáticamente para cortar el ARN objetivo. La escisión estratégica de dicho ARN objetivo destruirá su capacidad para dirigir la síntesis de una proteína codificada. Después de que un ácido nucleico enzimático se ha unido y escindido su objetivo de ARN, se libera de ese ARN para buscar otro objetivo y puede unirse y escindir repetidamente nuevos objetivos.

La molécula de ácido nucleico enzimático puede formarse en un motivo de cabeza de martillo, horquilla, un virus de la hepatitis δ , intrón del grupo I o ARN de ARNasaP (en asociación con una secuencia guía de ARN) o ARN de *Neurospora* VS, por ejemplo. Los ejemplos específicos de motivos de cabeza de martillo se describen por Rossi y otros, *Nucleic Acids Res.* (1992), 20(17):4,559-65. Los ejemplos de motivos de horquilla se describen por Hampel y otros. (Publicación de solicitud de patente europea núm. EP 0360257), Hampel y Tritz, *Biochemistry* (1989), 28(12):4,929-33; Hampel y otros, *Nucleic Acids Res.* (1990), 18(2):299-304 y patente de Estados Unidos núm. 5,631,359. Un ejemplo del motivo del virus de la hepatitis se describe por Perrotta y Been, *Biochemistry* (1992), 31(47): 11,843-52; un ejemplo del motivo de ARNasaP se describe por Guerrier-Takada y otros, *Cell* (1983), 35(3 Pt 2):849-57; el motivo de ribozima de ARN de *Neurospora* VS se describe por Collins (Saville y Collins, *Cell* (1990), 61(4):685-96; Saville y Collins, *Proc. Natl. Acad. Sci. USA* (1991), 88(19):8,826-30; Collins y Olive, *Biochemistry* (1993), 32(11):2,795-9); y un ejemplo del intrón del Grupo I se describe en la patente de Estados Unidos 4,987,071. En modalidades, las moléculas de ácido nucleico enzimático tienen un sitio de unión al sustrato específico que es complementario a una o más de las regiones de ADN o ARN del gen objetivo, y que tienen secuencias de nucleótidos dentro o alrededor de ese sitio de unión al sustrato que imparten una actividad de escisión del ARN a la molécula. Por tanto, los constructos de ribozima no necesitan limitarse a los motivos específicos mencionados en la presente descripción.

Las ribozimas pueden diseñarse como se describe en la publicación de solicitud de patente internacional núm. WO 93/23569 y la publicación de solicitud de patente internacional núm. WO 94/02595, cada uno incorporado específicamente en la presente descripción como referencia, y sintetizado para someterse a prueba in vitro e in vivo, como se describe en la misma. En modalidades, la ribozima se dirige a una secuencia de nucleótidos objetivo de un transcrito objetivo.

La actividad de la ribozima puede aumentarse al alterar la longitud de los brazos de unión de la

ribozima o mediante la síntesis química de ribozimas con modificaciones que evitan su degradación por las ribonucleasas séricas (ver, por ejemplo, la publicación de solicitud de patente internacional núm. WO 92/07065; publicación de solicitud de patente internacional núm. WO 93/15187; publicación de solicitud de patente internacional núm. WO 91/03162; publicación de solicitud de patente europea núm. 92110298.4; patente de Estados Unidos núm. 5,334,711; y publicación de solicitud de patente internacional núm. WO 94/13688, que describen diversas modificaciones químicas que pueden realizarse a los restos de azúcar de las moléculas de ARN enzimático), modificaciones que potencian su eficacia en células y la eliminación de bases madre II para acortar los tiempos de síntesis del ARN y reducir los requerimientos químicos.

10 Supermir

En modalidades, el AC es un supermir. Un supermir se refiere a un oligómero monocatenario, bicatenario, o parcialmente bicatenario o polímero de ARN, polímero de ADN, o ambos, o modificaciones de estos, que tiene una secuencia de nucleótidos que es sustancialmente idéntica a un miARN y que es antisentido con respecto a su objetivo, Este término incluye oligonucleótidos compuestos por nucleobases de origen natural, azúcares y enlaces internucleósidos covalentes (cadena principal) y que contienen al menos una porción que no es de origen natural que funciona de manera similar. Dichos oligonucleótidos modificados o sustituidos tienen propiedades convenientes tales como, por ejemplo, mejor captación celular, mayor afinidad por el objetivo del ácido nucleico y mayor estabilidad en presencia de nucleasas. En modalidades, el supermir no incluye una hebra sentido, y en otra modalidad, el supermir no se autohibrida en un grado significativo. Un supermir puede tener una estructura secundaria, pero es sustancialmente monocatenario en condiciones fisiológicas. Un supermir que es esencialmente monocatenario es monocatenario en la medida en que menos de aproximadamente 50 % (por ejemplo, menos de aproximadamente 40 %, aproximadamente 30 %, aproximadamente 20 %, aproximadamente 10 % o aproximadamente 5 %) del supermir forma un dúplex con sí mismo. El supermir puede incluir un segmento en horquilla, por ejemplo, la secuencia, por ejemplo, en el extremo 3' puede autohibridarse y formar una región dúplex, por ejemplo, una región dúplex de al menos aproximadamente 1, aproximadamente 2, aproximadamente 3, o aproximadamente 4 o menos de aproximadamente 8, aproximadamente 7, aproximadamente 6, o aproximadamente 5 nucleótidos o aproximadamente 5 nucleótidos. La región dúplex puede conectarse mediante un enlazador, por ejemplo, un enlazador de nucleótidos, por ejemplo, aproximadamente 3, aproximadamente 4, aproximadamente 5 o aproximadamente 6 dT, por ejemplo, dT modificadas. En otra modalidad, el supermir forma un

dúplex con un oligo más corto, por ejemplo, de aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9 o aproximadamente 10 nucleótidos de longitud, por ejemplo, en uno o ambos de los extremos 3' y 5' o en un extremo y en el no terminal o en el medio del supermir.

5 Imitadores de miARN

En modalidades, el AC es un imitador de miARN. Los imitadores de miARN representan una clase de moléculas que pueden usarse para imitar la capacidad de silenciamiento génico de uno o más miARN. Por tanto, el término "imitador de microARN" se refiere a ARN sintéticos no codificantes (es decir, el miARN no se obtiene mediante purificación a partir de una fuente del miARN endógeno) que es capaz de entrar en la vía del ARNi y regular la expresión génica. Los imitadores de miARN pueden diseñarse como moléculas maduras (por ejemplo, monocatenarios) o precursores de imitadores (por ejemplo, pri- o pre-miARN). Los imitadores de miARN pueden incluir ácido nucleico (ácidos nucleicos modificados o modificados) lo que incluye oligonucleótidos que incluyen, sin limitación, ARN, ARN modificado, ADN, ADN modificado, ácidos nucleicos bloqueados, o ácidos nucleicos con puente de 2'-0,4'-C-etileno (ENA), o cualquier combinación de los anteriores (lo que incluidos los híbridos ADN-ARN). Además, los imitadores de miARN pueden incluir conjugados que pueden afectar el suministro, la compartimentación intracelular, la estabilidad, la especificidad, la funcionalidad, el uso de la hebra y/o la potencia. En un diseño, los imitadores de miARN son moléculas bicatenarias (por ejemplo, con una región dúplex de entre aproximadamente 16 y aproximadamente 31 nucleótidos de longitud) y contienen una o más secuencias que tienen identidad con la hebra madura de un miARN dado. Las modificaciones pueden incluir modificaciones en 2' (lo que incluyen modificaciones 2'-0 metilo y modificaciones 2' F) en una o ambas hebras de la molécula y modificaciones internucleósidos (por ejemplo, modificaciones de fosforotioato) que potencian la estabilidad y/o especificidad del ácido nucleico. Además, los imitadores de miARN pueden incluir salientes. Los salientes pueden incluir de aproximadamente 1 a aproximadamente 6 nucleótidos ya sea en el extremo 3' o 5' de cualquiera de las hebras y pueden modificarse para potenciar la estabilidad o funcionalidad. En modalidades, un imitador de miARN incluye una región dúplex de aproximadamente 16 a aproximadamente 31 nucleótidos y uno o más de los siguientes patrones de modificación química: la hebra sentido contiene modificaciones 2'-0-metilo de los nucleótidos 1 y 2 (contando desde el extremo 5' del oligonucleótido sentido), y todas las C y los U; las modificaciones de la hebra antisentido pueden incluir la modificación 2' F de todas las C y los U, fosforilación del extremo 5' del oligonucleótido, y enlaces internucleósidos

estabilizados asociados con un saliente 3' de 2 nucleótidos.

Inhibidor de miARN

En modalidades, el AC es un inhibidor de miARN. Los términos "antimir" "inhibidor de microARN", "inhibidor de miR" o "inhibidor de miARN" son sinónimos y se refieren a oligonucleótidos u oligonucleótidos modificados que interfieren con la capacidad de miARN específicos. En general, los inhibidores son ácidos nucleicos o ácidos nucleicos modificados en la naturaleza, lo que incluye oligonucleótidos que incluyen ARN, ARN modificado, ADN, ADN modificado, ácidos nucleicos bloqueados (LNA) o cualquier combinación de los anteriores.

Las modificaciones incluyen modificaciones 2' (lo que incluye modificaciones 2'-O alquilo y modificaciones 2' F) y modificaciones internucleósidos (por ejemplo, modificaciones de fosforotioato) que pueden afectar el suministro, la estabilidad, la especificidad, la compartimentación intracelular o la potencia. Además, los inhibidores de miARN pueden incluir conjugados que pueden afectar el suministro, la compartimentación intracelular, la estabilidad y/o la potencia. Los inhibidores pueden adoptar una variedad de configuraciones, lo que incluye diseños monocatenarios, bicatenarios (ARN/ARN o dúplex ARN/ADN) y de horquillas, en general, los inhibidores de microARN incluyen contener una o más secuencias o porciones de secuencias que son complementarias o parcialmente complementarias con la hebra (o hebras) madura(s) del miARN a ser tomado como objetivo. Además, el inhibidor de miARN también puede incluir secuencias adicionales ubicadas en 5' y 3' con respecto a la secuencia que es el complemento inverso del miARN maduro. Las secuencias adicionales pueden ser los complementos inversos de las secuencias que están adyacentes al miARN maduro en el pri-miARN del cual se deriva el miARN maduro, o las secuencias adicionales pueden ser secuencias arbitrarias (que tienen una mezcla de A, G, C o U). En modalidades, una o ambas secuencias adicionales son secuencias arbitrarias capaces de formar horquillas. Por tanto, en modalidades, la secuencia que es el complemento inverso del miARN está flanqueada en el lado 5' y en el lado 3' por estructuras en horquilla. Los inhibidores de micro-ARN, cuando son bicatenarios, pueden incluir errores de emparejamiento entre los nucleótidos en hebras opuestas. Además, los inhibidores de micro-ARN pueden enlazarse a restos conjugados para facilitar la captación del inhibidor a una célula. Por ejemplo, un inhibidor de micro-ARN puede estar unido a 5-(bis(4-metoxifenil)(fenil)metoxi)-3 hidroxipentilcarbamato) de colesterilo que permite la captación pasiva de un inhibidor de micro-ARN a una célula. Los inhibidores de microARN, lo que incluye inhibidores de miARN en horquilla, se describen en detalle en Vermeulen y otros, RNA 13: 723- 730 (2007) y en los documentos WO2007/095387 y WO 2008/036825, cada uno de los cuales

se incorpora en la presente descripción como referencia en su totalidad. Un experto en la técnica puede seleccionar una secuencia de la base de datos para un miARN deseado y diseñar un inhibidor útil para los métodos descritos en la presente descripción.

Los grupos de enlace o restos de enlace bifuncionales tales como los conocidos en la técnica son compatibles con los compuestos proporcionados en la presente descripción. Los grupos de enlace son útiles para la unión de grupos químicos funcionales, grupos conjugados, grupos indicadores y otros grupos a sitios selectivos en un compuesto parental tal como, por ejemplo, un AC. En general, un resto de enlace bifuncional incluye un resto de hidrocarbilo que tiene dos grupos funcionales. Uno de los grupos funcionales se selecciona para unirse a una molécula parental o compuesto de interés, y el otro se selecciona para unirse esencialmente a cualquier grupo seleccionado tal como un grupo químico funcional o un grupo conjugado. Puede usarse cualquiera de los enlazadores descritos aquí. En modalidades, el enlazador incluye una estructura de cadena o un oligómero de unidades repetitivas tales como unidades de etilenglicol o aminoácidos. Los ejemplos de grupos funcionales que se usan rutinariamente en un resto de enlace bifuncional incluyen, pero no se limitan a, electrófilos para reaccionar con grupos nucleofílicos y nucleófilos para reaccionar con grupos electrofílicos. En modalidades, los restos de enlace bifuncionales incluyen amino, hidroxilo, ácido carboxílico, tiol, insaturaciones (por ejemplo, dobles o triples enlaces), y similares. Algunos ejemplos no limitantes de restos de enlace bifuncionales incluyen ácido 8-amino-3,6-dioxaoctanoico (ADO), 4-(N-maleimidometil) ciclohexano-1-carboxilato de succinimidilo (SMCC) y ácido 6-aminohexanoico (AHEX o AHA). Otros grupos de enlace incluyen, pero no se limitan a, alquilo C1-C10 sustituido, alqueno C2-C10 sustituido o no sustituido, o alquino C2-C10 sustituido o no sustituido, en donde una lista no limitante de grupos sustituyentes incluye hidroxilo, amino, alcoxi, carboxi, bencilo, fenilo, nitro, tiol, tioalcoxi, halógeno, alquilo, arilo, alqueno y alquino.

En modalidades, el AC incluye la modificación de nucleótidos diseñada para no soportar la actividad de la ARNasa H. Se conocen modificaciones nucleotídicas de compuestos antisentido que no soportan la actividad de la ARNasa H e incluyen, pero no se limitan a, modificaciones 2'-O-metoxi etil/fosforotioato (MOE). Ventajosamente, el AC con modificaciones MOE tiene mayor afinidad por el ARN objetivo y aumenta la estabilidad ante las nucleasas.

Oligonucleótidos inmunoestimuladores

En modalidades, el resto terapéutico es un oligonucleótido inmunoestimulador. Los oligonucleótidos inmunoestimuladores (ISS; monocatenarios o bicatenarios) son capaces de inducir una respuesta inmunitaria cuando se administran a un paciente, que puede ser un mamífero u otro paciente. Los

ISS incluyen, por ejemplo, ciertos palíndromos que conducen a estructuras secundarias en horquilla (ver Yamamoto S., y otros (1992) J. Immunol. 148: 4,072-4,076), o motivos CpG, así como también otras características conocidas de los ISS (tales como dominios multi-G, ver el documento WO 96/11266).

5 La respuesta inmunitaria puede ser una respuesta inmunitaria innata o adaptativa. El sistema inmunitario se divide en un sistema inmunitario más innato y en un sistema inmunitario adaptativo adquirido de los vertebrados, este último se divide además en componentes celulares y humorales. En modalidades particulares, la respuesta inmunitaria puede ser mucosal.

Los ácidos nucleicos inmunoestimuladores se consideran no específicos de secuencia, cuando no se requiere que se unan específicamente a y reduzcan la expresión de un polinucleótido objetivo para provocar una respuesta inmunitaria. Por tanto, ciertos ácidos nucleicos inmunoestimuladores pueden incluir una secuencia correspondiente a una región de un gen o ARNm de origen natural, pero aún pueden considerarse ácidos nucleicos inmunoestimuladores que no son específicos de secuencia.

En modalidades, el ácido nucleico u oligonucleótido inmunoestimulador incluye al menos un dinucleótido CpG. El oligonucleótido o el dinucleótido CpG pueden estar metilados o no metilados. En otra modalidad, el ácido nucleico inmunoestimulador incluye al menos un dinucleótido CpG que tiene una citosina metilada. En modalidades, el ácido nucleico incluye un único dinucleótido CpG, en donde la citosina en dicho dinucleótido CpG está metilada. En una modalidad específica, el ácido nucleico incluye la secuencia 5' TAACGTTGAGGG'CAT 3' (SEQ ID NO: 369). En una modalidad alternativa, el ácido nucleico incluye al menos dos dinucleótidos CpG, en donde al menos una citosina en los dinucleótidos CpG está metilada. En una modalidad adicional, cada citosina en los dinucleótidos CpG presentes en la secuencia está metilada. En otra modalidad, el ácido nucleico incluye una pluralidad de dinucleótidos CpG, en donde al menos uno de dichos dinucleótidos CpG incluye una citosina metilada.

25 Secuencias de ácidos nucleicos específicas adicionales de oligonucleótidos (ODN) adecuadas para su uso en las composiciones y métodos se describen en Raney y otros, Journal of Pharmacology and Experimental Therapeutics, 298:1,185-1,192 (2001). En ciertas modalidades, los ODN usados en las composiciones y métodos tienen una cadena principal de fosfodiéster ("PO") o una cadena principal de fosforotioato ("PS"), y/o al menos un residuo de citosina metilada en un motivo CpG.

30 Oligonucleótidos señuelo

En modalidades, el resto terapéutico es un oligonucleótido señuelo. Debido a que los factores de transcripción reconocen sus secuencias de unión relativamente cortas, incluso en ausencia del ADN

genómico colindante, los oligonucleótidos cortos que portan la secuencia de unión consenso de un factor de transcripción específico pueden usarse como herramientas para manipular la expresión génica en células vivas. Esta estrategia implica el suministro intracelular de dichos "oligonucleótidos señuelo", que después se reconocen y se unen por el factor objetivo. La ocupación del sitio de unión al ADN del factor de transcripción por parte del señuelo hace que el factor de transcripción sea incapaz de unirse subsecuentemente a las regiones promotoras de los genes objetivo. Los señuelos pueden usarse como agentes terapéuticos, ya sea para inhibir la expresión de genes que se activan mediante un factor de transcripción, o para regular positivamente los genes que se suprimen mediante la unión de un factor de transcripción. Ejemplos de la utilización de oligonucleótidos señuelo pueden encontrarse en Mann y otros, J. Clin. Invest, 2000, 106: 1071-1075, que se incorpora expresamente como referencia en la presente descripción, en su totalidad.

Adaptador U1

En algunas modalidades, el resto terapéutico es un adaptador U1. Los adaptadores U1 inhiben sitios poliA y son oligonucleótidos bifuncionales con un dominio objetivo complementario a un sitio en el exón terminal del gen objetivo y un 'dominio U1' que se une al componente de ARN nuclear más pequeño U1 del snRNP U1 (Goracznik, y otros, 2008, Nature Biotechnology, 27(3), 257-263, que se incorpora expresamente como referencia en la presente descripción, en su totalidad). El snRNP U1 es un complejo de ribonucleoproteína que funciona principalmente para dirigir las primeras etapas en la formación de espliceosomas mediante la unión al límite intrón-exón del pre-ARNm (Brown y Simpson, 1998, Annu Rev Plant Physiol Plant Mol Biol 49:77-95). Los nucleótidos 2-11 del extremo 5' del par de bases de ARNnp U1 se unen con los 5'ss del pre-ARN. En una modalidad, los oligonucleótidos son adaptadores U1. En una modalidad, el adaptador U1 puede administrarse en combinación con al menos otro agente iARN.

Maquinaria de edición de genes (CRISPR)

En modalidades, los compuestos descritos en la presente descripción incluyen uno o más CPP (o cCPP) conjugados a la maquinaria de edición de genes CRISPR. Como se usa en la presente, "maquinaria de edición de genes CRISPR" se refiere a las proteínas, ácidos nucleicos o sus combinaciones, que pueden usarse para editar un genoma. Los ejemplos no limitantes de maquinaria de edición de genes incluyen ARNg, nucleasas, inhibidores de nucleasas y combinaciones y complejos de estos. Los siguientes documentos de patente describen la maquinaria de edición de genes CRISPR: Patente de Estados Unidos núm. 8,697,359, patente de Estados Unidos núm. 8,771,945, patente de Estados Unidos núm. 8,795,965, patente de Estados Unidos núm. 8,865,406,

patente de Estados Unidos núm. 8,871,445, patente de Estados Unidos núm. 8,889,356, patente de Estados Unidos núm. 8,895,308, patente de Estados Unidos núm. 8,906,616, patente de Estados Unidos núm. 8,932,814, patente de Estados Unidos núm. 8,945,839, patente de Estados Unidos núm. 8,993,233, patente de Estados Unidos núm. 8,999,641, solicitud de patente de Estados Unidos núm. 14/704,551, y la solicitud de patente de Estados Unidos núm. 13/842,859. Cada uno de los documentos de patente mencionados anteriormente se incorpora como referencia en la presente descripción en su totalidad.

En modalidades, un enlazador conjuga el cCPP a la maquinaria de edición de genes CRISPR. Puede utilizarse cualquier enlazador descrito en esta descripción o que sea conocido por un experto en la técnica.

ARNg

En modalidades, los compuestos incluyen que el CPP (o cCPP) se conjugue a un ARNg. Un ARNg se dirige a un loci genómico en una célula procariota o eucariota.

En modalidades, el ARNg es un ARN guía de molécula simple (ARNgs). Un ARNg incluye una secuencia espaciadora y una secuencia de andamio. Una secuencia espaciadora es una secuencia de ácido nucleico corta usada para dirigir una nucleasa (por ejemplo, una nucleasa Cas9) a una región de nucleótidos específica de interés (por ejemplo, una secuencia de ADN genómico a escindir). En modalidades, el espaciador puede tener aproximadamente 17-24 bases de longitud, tal como aproximadamente 20 bases de longitud. En modalidades, el espaciador puede tener aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24, aproximadamente 25, aproximadamente 26, aproximadamente 27, aproximadamente 28, aproximadamente 29 o aproximadamente 30 bases de longitud. En modalidades, el espaciador puede tener al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 21, al menos 22, al menos 23, al menos 24, al menos 25, al menos 26, al menos 27, al menos 28, al menos 29 o al menos 30 bases de longitud. En modalidades, el espaciador puede tener aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24, aproximadamente 25, aproximadamente 26, aproximadamente 27, aproximadamente 28, aproximadamente 29 o aproximadamente 30 bases de longitud. En modalidades, la secuencia espaciadora tiene entre aproximadamente 40 % a aproximadamente 80 % de contenido de GC.

En modalidades, el espaciador se dirige a un sitio que precede inmediatamente a un motivo adyacente al protoespaciador (PAM) 5'. La secuencia PAM puede seleccionarse basado en la nucleasa deseada. Por ejemplo, la secuencia PAM puede ser una cualquiera de las secuencias PAM mostradas en la Tabla 3 más abajo, en donde N se refiere a cualquier ácido nucleico, R se refiere a A o G, Y se refiere a C o T, W se refiere a A o T, y V se refiere a A o C o G.

Tabla 3: Nucleasas y secuencias PAM ilustrativas

Secuencia PAM (5' a 3')	Nucleasa	Aislada de
NGG	SpCas9	<i>Streptococcus pyogenes</i>
NGRRT o NGRRN	SaCas9	<i>Staphylococcus aureus</i>
NNNNGATT	NmeCas9	<i>Neisseria meningitidis</i>
NNNNRYAC	CjCas9	<i>Campylobacter jejuni</i>
NNAGAAW	StCas9	<i>Termófilos de estreptocococo</i>
TTTV	LbCpf1	<i>Lachnospiraceae bacterium</i>
TTTV	AsCpf1	<i>Acidaminococcus sp.</i>

En modalidades, un espaciador puede dirigirse a una secuencia de un gen de mamífero, tal como un gen humano. En modalidades, el espaciador puede dirigirse a un gen mutante. En modalidades, el espaciador puede dirigirse a una secuencia codificante. En modalidades, el espaciador puede dirigirse a una secuencia exónica. En modalidades, el espaciador puede dirigirse a un sitio de poliadenilación (PS). En modalidades, el espaciador puede dirigirse a un elemento de secuencia de un PS. En modalidades, el espaciador puede dirigirse a una señal de poliadenilación (PAS), una secuencia interviniente (IS), un sitio de escisión (CS), un elemento aguas abajo (DES), o una porción o sus combinaciones. En modalidades, un espaciador puede dirigirse a un elemento de corte y empalme (SE) o a un elemento regulador del corte y empalme en cis (SRE).

La secuencia de andamio es la secuencia dentro del ARNs que es responsable de la unión a nucleasas (por ejemplo, Cas9). La secuencia de andamio no incluye la secuencia de direccionamiento/espaciador. En modalidades, el andamio puede tener aproximadamente 1 a aproximadamente 10, aproximadamente 10 a aproximadamente 20, aproximadamente 20 a aproximadamente 30, aproximadamente 30 a aproximadamente 40, aproximadamente 40 a aproximadamente 50, aproximadamente 50 a aproximadamente 60, aproximadamente 60 a aproximadamente 70, aproximadamente 70 a aproximadamente 80, aproximadamente 80 a aproximadamente 90, aproximadamente 90 a aproximadamente 100, aproximadamente 100 a aproximadamente 110, aproximadamente 110 a aproximadamente 120, o aproximadamente 120 a aproximadamente 130 nucleótidos de longitud. En modalidades, el andamio puede tener

aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4,
aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8,
aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12,
aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16,
5 aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20,
aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24,
aproximadamente 25, aproximadamente 26, aproximadamente 27, aproximadamente 28,
aproximadamente 29, aproximadamente 30, aproximadamente 31, aproximadamente 32,
aproximadamente 33, aproximadamente 34, aproximadamente 35, aproximadamente 36,
10 aproximadamente 37, aproximadamente 38, aproximadamente 39, aproximadamente 40,
aproximadamente 41, aproximadamente 42, aproximadamente 43, aproximadamente 44,
aproximadamente 45, aproximadamente 46, aproximadamente 47, aproximadamente 48,
aproximadamente 49, aproximadamente 50, aproximadamente 51, aproximadamente 52,
aproximadamente 53, aproximadamente 54, aproximadamente 55, aproximadamente 56,
15 aproximadamente 57, aproximadamente 58, aproximadamente 59, aproximadamente
60,aproximadamente 60, aproximadamente 61, aproximadamente 62, aproximadamente 63,
aproximadamente 64, aproximadamente 65, aproximadamente 66, aproximadamente 67,
aproximadamente 68, aproximadamente 69, aproximadamente 70, aproximadamente 71,
aproximadamente 72, aproximadamente 73, aproximadamente 74, aproximadamente 75,
20 aproximadamente 76, aproximadamente 77, aproximadamente 78, aproximadamente 79,
aproximadamente 80, aproximadamente 81, aproximadamente 82, aproximadamente 83,
aproximadamente 84, aproximadamente 85, aproximadamente 86, aproximadamente 87,
aproximadamente 88, aproximadamente 89, aproximadamente 90, aproximadamente 91,
aproximadamente 92, aproximadamente 93, aproximadamente 94, aproximadamente 95,
25 aproximadamente 96, aproximadamente 97, aproximadamente 98, aproximadamente 99,
aproximadamente 100, aproximadamente 101, aproximadamente 102, aproximadamente 103,
aproximadamente 104, aproximadamente 105, aproximadamente 106, aproximadamente 107,
aproximadamente 108, aproximadamente 109, aproximadamente 110, aproximadamente 111,
aproximadamente 112, aproximadamente 113, aproximadamente 114, aproximadamente 115,
30 aproximadamente 116, aproximadamente 117, aproximadamente 118, aproximadamente 119,
aproximadamente 120, aproximadamente 121, aproximadamente 122, aproximadamente 123,
aproximadamente 124, o aproximadamente 125 nucleótidos de longitud. En modalidades, el

andamio puede tener al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 110, al menos 120 o al menos 125 nucleótidos de longitud.

En modalidades, el ARNg es un ARN guía de doble molécula, por ejemplo, ARNcr y ARNtracr. En modalidades, el ARNg puede incluir además una cola poli(A).

En modalidades, un compuesto que incluye un CPP se conjuga con un ácido nucleico que incluye un ARNg. En modalidades, el ácido nucleico incluye aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19 o aproximadamente 20 ARNg. En modalidades, los ARNg reconocen el mismo objetivo. En modalidades, los ARNg reconocen objetivos diferentes. En modalidades, el ácido nucleico que incluye un ARNg incluye una secuencia que codifica un promotor, en donde el promotor dirige la expresión del ARNg.

Nucleasa

En modalidades, los compuestos incluyen un péptido penetrante de células conjugado a una nucleasa. En modalidades, la nucleasa es una nucleasa tipo II, tipo V-A, tipo V-B, tipo VC, tipo V-U, tipo VI-B. En modalidades, la nucleasa es una nucleasa efectora similar a activador de la transcripción (TALEN), una meganucleasa o una nucleasa de dedos de zinc. En modalidades, la nucleasa es una nucleasa Cas9, Cas12a (Cpf1), Cas12b, Cas12c, similar a Tnp-B, Cas13a (C2c2), Cas13b o Cas14. Por ejemplo, en algunas modalidades, la nucleasa es una nucleasa Cas9 o una nucleasa Cpf1.

En modalidades, la nucleasa es una forma modificada o variante de una nucleasa Cas9, Cas12a (Cpf1), Cas12b, Cas12c, similar a Tnp-B, Cas13a (C2c2), Cas13b o Cas14. En modalidades, la nucleasa es una forma modificada o variante de una nucleasa TAL, una meganucleasa o una nucleasa de dedos de zinc. Una nucleasa “modificada” o “variante” es una que, por ejemplo, se trunca, se fusiona con otra proteína (tal como otra nucleasa), se inactiva catalíticamente, etc. En modalidades, la nucleasa puede tener al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 98 %, al menos aproximadamente 99 % o aproximadamente 100 % de identidad de secuencia con una nucleasa Cas9, Cas12a (Cpf1), Cas12b, Cas12c, similar a Tnp-B, Cas13a (C2c2), Cas13b, Cas14 o

una TALEN, meganucleasa o nucleasa de dedos de zinc. En modalidades, la nucleasa es una nucleasa Cas9 derivada de *S. pyogenes* (SpCas9). En modalidades, una nucleasa tiene al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 % o al menos aproximadamente 99 % de identidad de secuencia con una nucleasa Cas9 derivada de *S. pyogenes* (SpCas9). En modalidades, la nucleasa es una Cas9 derivada de *S. aureus* (SaCas9). En modalidades, la nucleasa tiene al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 % o al menos aproximadamente 99 % de identidad de secuencia con una Cas9 derivada de *S. aureus* (SaCas9). En modalidades, la Cpf1 es una enzima Cpf1 de *Acidaminococcus* (especie BV3L6, núm. de acceso U2UMQ6 de UniProt). En modalidades, la nucleasa tiene al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 % o al menos aproximadamente 99 % de identidad de secuencia con una enzima Cpf1 de *Acidaminococcus* (especie BV3L6, núm. de acceso U2UMQ6 de UniProt). En modalidades, la Cpf1 es una enzima Cpf1 de *Lachnospiraceae* (especie ND2006, núm. de acceso A0A182DWE3 de UniProt). En modalidades, la nucleasa tiene al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente 97 %, al menos aproximadamente 98 % o al menos aproximadamente 99 % de identidad de secuencia con una enzima Cpf1 de *Lachnospiraceae*. En modalidades, una secuencia que codifica la nucleasa se optimiza en codones para su expresión en células de mamífero. En modalidades, la secuencia que codifica la nucleasa se optimiza en codones para su expresión en células humanas o células de ratón. En modalidades, un compuesto que incluye un CPP se conjuga con una nucleasa. En modalidades, la nucleasa es una proteína soluble. En modalidades, un compuesto que incluye un CPP se conjuga con un ácido nucleico que codifica una nucleasa. En modalidades, el ácido nucleico que codifica una nucleasa incluye una secuencia que codifica un promotor, en donde el promotor dirige la expresión de la nucleasa.

Combinaciones de ARNg y nucleasas

En modalidades, los compuestos incluyen uno o más CPP (o cCPP) conjugados a un ARNg y una nucleasa. En modalidades, el uno o más CPP (o cCPP) se conjugan a un ácido nucleico que codifica un ARNg y/o una nucleasa. En modalidades, el ácido nucleico que codifica una nucleasa y un ARNg incluye una secuencia que codifica un promotor, en donde el promotor dirige la expresión de la nucleasa y el ARNg. En modalidades, el ácido nucleico que codifica una nucleasa y un ARNg

incluye dos promotores, en donde un primer promotor controla la expresión de la nucleasa y un segundo promotor controla la expresión del ARNg. En modalidades, el ácido nucleico que codifica un ARNg y una nucleasa codifica de aproximadamente 1 a aproximadamente 20 ARNg, o de aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18 o aproximadamente 19 y hasta aproximadamente 20 ARNg. En modalidades, los ARNg reconocen objetivos diferentes. En modalidades, los ARNg reconocen el mismo objetivo.

En modalidades, los compuestos incluyen un péptido penetrante de células (o cCPP) conjugado a una ribonucleoproteína (RNP) que incluye un ARNg y una nucleasa.

En modalidades, una composición que incluye: (a) un CPP conjugado a un ARNg y (b) una nucleasa, se suministra a una célula. En modalidades, una composición que incluye: (a) un CPP conjugado a una nucleasa y (b) un ARNg, se suministra a una célula.

En modalidades, una composición que incluye: (a) un primer CPP conjugado a un ARNg y (b) un segundo CPP conjugado a una nucleasa, se suministra a una célula. En modalidades, el primer CPP y el segundo CPP son el mismo. En modalidades, el primer CPP y el segundo CPP son diferentes.

Elemento genético de interés

En modalidades, los compuestos descritos en la presente descripción incluyen un péptido penetrante de células conjugado a un elemento genético de interés. En modalidades, un elemento genético de interés reemplaza una secuencia de ADN genómico escindida por una nucleasa. Los ejemplos no limitantes de elementos genéticos de interés incluyen genes, un polimorfismo de un solo nucleótido, un promotor o terminadores.

Inhibidores de nucleasas

En modalidades, los compuestos descritos en la presente descripción incluyen un péptido penetrante de células conjugado a un inhibidor de una nucleasa (por ejemplo, Cas9). Una limitación de la edición de genes es la potencial edición fuera de objetivo. El suministro de un inhibidor de nucleasas limitará la edición fuera de objetivo. En modalidades, el inhibidor de nucleasas es un polipéptido, un polinucleótido o una molécula pequeña. Inhibidores de nucleasas ilustrativos se describen en la publicación de los Estados Unidos núm. 2020/087354, la publicación internacional núm. 2018/085288, la publicación de los Estados Unidos núm. 2018/0382741, la publicación internacional

núm. 2019/089761, la publicación internacional núm. 2020/068304, la publicación internacional núm. 2020/041384 y la publicación internacional núm. 2019/076651, cada una de las cuales se incorpora como referencia en la presente descripción en su totalidad.

Polipéptidos terapéuticos

- 5 En modalidades, el resto terapéutico incluye un polipéptido. En modalidades, el resto terapéutico incluye una proteína o un fragmento de esta. En modalidades, el resto terapéutico incluye una proteína de unión al ARN o un fragmento de unión al ARN de esta. En modalidades, el resto terapéutico incluye una enzima. En modalidades, el resto terapéutico incluye una enzima de escisión del ARN o un fragmento activo de esta.

10 Grupos conjugados

- En modalidades, los AC se modifican mediante la unión covalente de uno o más grupos conjugados. En general, los grupos conjugados modifican una o más propiedades del AC unido, lo que incluye, pero no se limita a, farmacodinámica, farmacocinética, unión, absorción, distribución celular, captación celular, carga y aclaramiento. Los grupos conjugados se usan rutinariamente en las técnicas químicas y se enlazan directamente o mediante un resto de enlace o grupo de enlace opcional a un compuesto parental tal como un AC. Los grupos conjugados incluyen, sin limitación, intercaladores, moléculas indicadoras, poliaminas, poliamidas, polietilenglicoles, tioéteres, poliéteres, colesteroles, tiocolesteroles, restos de ácido cólico, folato, lípidos, fosfolípidos, biotina, fenazina, fenantridina, antraquinona, adamantano, acridina, fluoresceínas, rodaminas, cumarinas y colorantes. En modalidades, el grupo conjugado es un polietilenglicol (PEG), y el PEG se conjuga con ya sea el AC o el CPP.
- 15 20

- Los grupos conjugados incluyen restos lipídicos tales como un resto de colesterol (Letsinger y otros, Proc. Natl. Acad. Sci. USA, 1989, 86, 6,553); ácido cólico (Manoharan y otros, Bioorg. Med. Chem. Lett., 1994, 4, 1,053); un tioéter, por ejemplo, hexil-S-tritiltiol (Manoharan y otros, Ann. N.Y. Acad. Sci., 1992, 660, 306; Manoharan y otros, Bioorg. Med. Chem. Let., 1993, 3, 2,765); un tiocolesterol (Oberhauser y otros, Nucl. Acids Res., 1992, 20, 533); una cadena alifática, por ejemplo, residuos de dodecandiol o undecilo (Saison-Behmoaras y otros, EMBO J., 1991, 10, 111; Kabanov y otros, FEBS Lett., 1990, 259, 327; Svinarchuk y otros, Biochimie, 1993, 75, 49); un fosfolípido, por ejemplo, di-hexadecil-rac-glicerol o trietilamonio-1,2-di-O-hexadecil-rac-glicero-3-H-fosfonato (Manoharan y otros, Tetrahedron Lett., 1995, 36, 3,651; Shea y otros, Nucl. Acids Res., 1990, 18, 3,777); una poliamina o una cadena de polietilenglicol (Manoharan y otros, Nucleosides & Nucleotides, 1995, 14, 969); ácido acético de adamantano (Manoharan y otros, Tetrahedron Lett.,
- 25 30

1995, 36, 3,651); un resto de palmitilo (Mishra y otros, Biochim. Biofis. Acta, 1995, 1,264, 229); o un resto de octadecilamina o hexilamino-carbonil-oxicoolesterol (Crooke y otros, J. Pharmacol. Exp. Ther., 1996,277,923).

Los grupos de enlace o restos de enlace bifuncionales tales como los conocidos en la técnica son compatibles con los compuestos proporcionados en la presente descripción. Los grupos de enlace son útiles para la unión de grupos químicos funcionales, grupos conjugados, grupos indicadores y otros grupos a sitios selectivos en un compuesto parental tal como, por ejemplo, un AC. En general, un resto de enlace bifuncional comprende un resto de hidrocarbilo que tiene dos grupos funcionales. Uno de los grupos funcionales se selecciona para unirse a una molécula parental o compuesto de interés, y el otro se selecciona para unirse esencialmente a cualquier grupo seleccionado tal como un grupo químico funcional o un grupo conjugado. Puede usarse cualquiera de los enlazadores descritos aquí. En modalidades, el enlazador comprende una estructura de cadena o un oligómero de unidades repetitivas tales como unidades de etilenglicol o aminoácidos. Los ejemplos de grupos funcionales que se usan rutinariamente en un resto de enlace bifuncional incluyen, pero no se limitan a, electrófilos para reaccionar con grupos nucleofílicos y nucleófilos para reaccionar con grupos electrofílicos. En modalidades, los restos de enlace bifuncionales incluyen amino, hidroxilo, ácido carboxílico, tiol, insaturaciones (por ejemplo, dobles o triples enlaces), y similares. Algunos ejemplos no limitantes de restos de enlace bifuncionales incluyen ácido 8-amino-3,6-dioxaoctanoico (ADO), 4-(N-maleimidometil) ciclohexano-1-carboxilato de succinimidilo (SMCC) y ácido 6-aminohexanoico (AHEX o AHA). Otros grupos de enlace incluyen, pero no se limitan a, alquilo C1-C10 sustituido, alqueno C2-C10 sustituido o no sustituido, o alquino C2-C10 sustituido o no sustituido, en donde una lista no limitante de grupos sustituyentes incluye hidroxilo, amino, alcoxi, carboxi, bencilo, fenilo, nitro, tiol, tioalcoxi, halógeno, alquilo, arilo, alqueno y alquino. En modalidades, el AC puede unirse a una repetición de 10 dipéptidos arginina-serina. Los AC unidos a 10 repeticiones del dipéptido arginina-serina para el reclutamiento artificial de factores potenciadores del corte y empalme se han aplicado in vitro para inducir la inclusión de exones de BRCA1 y SMN2 mutados que de cualquier otra manera se omitirían. Ver Cartegni y Krainer 2003, incorporado como referencia en la presente descripción.

Vehículos de escape endosómico (EEV)

Un vehículo de escape endosómico (EEV) puede usarse para transportar una carga a través de una membrana celular, por ejemplo, para suministrar la carga al citosol o al núcleo de una célula. La carga puede incluir un resto terapéutico (TM). El EEV puede comprender un péptido penetrante de

células (CPP), por ejemplo, un péptido penetrante de células cíclico (cCPP). En modalidades, el EEV comprende un cCPP, que se conjuga con un péptido exocíclico (EP). El EP puede denominarse indistintamente como péptido modulador (MP). El EP puede comprender una secuencia de una señal de localización nuclear (NLS). El EP puede acoplarse a la carga. El EP puede acoplarse al cCPP. El EP puede acoplarse a la carga y al cCPP. El acoplamiento entre el EP, la carga, el cCPP, o sus combinaciones, puede ser covalente o no covalente. El EP puede unirse a través de un enlace peptídico al extremo N-terminal del cCPP. El EP puede unirse a través de un enlace peptídico al extremo C-terminal del cCPP. El EP puede unirse al cCPP a través de una cadena lateral de un aminoácido en el cCPP. El EP puede unirse al cCPP a través de una cadena lateral de una lisina que puede conjugarse con la cadena lateral de una glutamina en el cCPP. El EP puede conjugarse al extremo 5' o 3' de una carga de oligonucleótidos. El EP puede acoplarse a un enlazador. El péptido exocíclico puede conjugarse a un grupo amino del enlazador. El EP puede acoplarse a un enlazador mediante el extremo C-terminal de un EP y un cCPP a través de una cadena lateral en el cCPP y/o EP. Por ejemplo, un EP puede comprender una lisina terminal que después puede acoplarse a un cCPP que contiene una glutamina a través de un enlace amida. Cuando el EP contiene una lisina terminal, y la cadena lateral de la lisina puede usarse para unirse al cCPP, el extremo C o N-terminal puede unirse a un enlazador en la carga.

Péptidos exocíclicos

El péptido exocíclico (EP) puede comprender de 2 a 10 residuos de aminoácidos, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o 10 residuos de aminoácidos, lo que incluye todos los intervalos y valores intermedios. El EP puede comprender de 6 a 9 residuos de aminoácidos. El EP puede comprender de 4 a 8 residuos de aminoácidos.

Cada aminoácido en el péptido exocíclico puede ser un aminoácido natural o no natural. El término “aminoácido no natural” se refiere a un compuesto orgánico que es un congénere de un aminoácido natural porque tiene una estructura similar a un aminoácido natural de modo que imita la estructura y la reactividad de un aminoácido natural. El aminoácido no natural puede ser un aminoácido modificado y/o análogo de aminoácidos, que no es uno de los 20 aminoácidos comunes de origen natural o los aminoácidos naturales raros selenocisteína o pirrolisina. Los aminoácidos no naturales también pueden ser el isómero D de los aminoácidos naturales. Ejemplos de aminoácidos adecuados incluyen, pero no se limitan a, alanina, alosoleucina, arginina, citrulina, asparagina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, naftilalanina, fenilalanina, prolina, ácido piroglutámico, serina, treonina, triptófano, tirosina, valina,

un derivado de estos, o sus combinaciones. Estos, y otros aminoácidos, se enumeran en la Tabla 4 junto con sus abreviaturas usadas en la presente descripción. Por ejemplo, los aminoácidos pueden ser A, G, P, K, R, V, F, H, Nal o citrulina.

El EP puede comprender al menos un residuo de aminoácido cargado positivamente, por ejemplo, al menos un residuo de lisina y/o al menos un residuo de ácido amina que comprende una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta. El EP puede comprender 1 o 2 residuos de aminoácidos que comprenden una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta. El residuo de aminoácido que comprende una cadena lateral que comprende un grupo guanidina puede ser un residuo de arginina. Las formas protonadas pueden significar una sal de esta a lo largo de la descripción.

El EP puede comprender al menos dos, al menos tres o al menos cuatro o más residuos de lisina. El EP puede comprender 2, 3 o 4 residuos de lisina. El grupo amino en la cadena lateral de cada residuo de lisina puede sustituirse con un grupo de protección, lo que incluye, por ejemplo, trifluoroacetilo (-COCF₃), aliloxicarbonilo (Alloc), 1-(4,4-dimetil-2,6-dioxociclohexiliden)etilo (Dde) o (4,4-dimetil-2,6-dioxociclohex-1-iliden-3)-metilbutilo (ivDde). El grupo amino en la cadena lateral de cada residuo de lisina puede sustituirse con un grupo trifluoroacetilo (-COCF₃). El grupo de protección puede incluirse para permitir la conjugación de amida. El grupo de protección puede eliminarse después de que el EP se conjuga con un cCPP.

El EP puede comprender al menos 2 residuos de aminoácidos con una cadena lateral hidrófoba. El residuo de aminoácido con una cadena lateral hidrófoba puede seleccionarse de valina, prolina, alanina, leucina, isoleucina y metionina. El residuo de aminoácido con una cadena lateral hidrófoba puede ser valina o prolina.

El EP puede comprender al menos un residuo de aminoácido cargado positivamente, por ejemplo, al menos un residuo de lisina y/o al menos un residuo de arginina. El EP puede comprender al menos dos, al menos tres o al menos cuatro o más residuos de lisina y/o residuos de arginina.

El EP puede comprender KK, KR, RR, HH, HK, HR, RH, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKH, KHK, HKK, HRR, HRH, HHR, HBH, HHH, HHHH (SEQ ID NO:1), KHKK (SEQ ID NO:2), KKHK (SEQ ID NO:3), KKKH (SEQ ID NO:4), KHKH (SEQ ID NO:5), HKHK (SEQ ID NO:6), KKKK (SEQ ID NO:7), KKRK (SEQ ID NO:8), KRKK (SEQ ID NO:9), KRRK (SEQ ID NO:10), RKKR (SEQ ID NO:11), RRRR (SEQ ID NO:12), KGKK (SEQ ID NO:13), KKGK (SEQ ID NO:14), HBHBH (SEQ ID NO:15), HBKBH (SEQ ID NO:16), RRRRR (SEQ ID NO:17), KKKKK (SEQ ID NO:18), KKKRK (SEQ ID NO:19), RKKKK (SEQ ID NO:20), KRKKK

(SEQ ID NO:21), KKRKK (SEQ ID NO:22), KKKKR (SEQ ID NO:23), KBKKB (SEQ ID NO:24), RKKKKG (SEQ ID NO:25), KRKKKG (SEQ ID NO:26), KKRKKG (SEQ ID NO:27), KKKKRG (SEQ ID NO:28), RKKKKB (SEQ ID NO:29), KRKKKB (SEQ ID NO:30), KKRKKB (SEQ ID NO:31), KKKKRB (SEQ ID NO:32), KKKRKV (SEQ ID NO:33), RRRRRR (SEQ ID NO:34),
5 HHHHHH (SEQ ID NO:35), RHRHRH (SEQ ID NO:36), HRHRHR (SEQ ID NO:37), KRKRKR (SEQ ID NO:38), RKRKRK (SEQ ID NO:39), RBRBRB (SEQ ID NO:40), KBKBKB (SEQ ID NO:41), PKKKRKV (SEQ ID NO:42), PGKKRKV (SEQ ID NO:43), PKGKRKV (SEQ ID NO:44), PKKGRKV (SEQ ID NO:45), PKKKGKV (SEQ ID NO:46), PKKKRGV (SEQ ID NO:47) o PKKKRKG (SEQ ID NO:48), en donde B es beta-alanina. Los aminoácidos en el EP pueden tener
10 una estereoquímica D o L.

El EP puede comprender KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK (SEQ ID NO:7), KKRK (SEQ ID NO:8), KRKK (SEQ ID NO:9), KRRK (SEQ ID NO:10), RKKR (SEQ ID NO:11), RRRR (SEQ ID NO:12), KGKK (SEQ ID NO:13), KKGK (SEQ ID NO:14), KKKKK (SEQ ID NO:18), KKKRK (SEQ ID NO:19), KBKKB (SEQ ID NO:24), KKKRKV (SEQ
15 ID NO:33), PKKKRKV (SEQ ID NO:42), PGKKRKV (SEQ ID NO:43), PKGKRKV (SEQ ID NO:44), PKKGRKV (SEQ ID NO:45), PKKKGKV (SEQ ID NO:46), PKKKRGV (SEQ ID NO:47) o PKKKRKG (SEQ ID NO:48). El EP puede comprender PKKKRKV (SEQ ID NO:42), RR, RRR, RHR, RBR, RBRBR (SEQ ID NO:49), RBHBR (SEQ ID NO:50) o HBRBH (SEQ ID NO:51), en donde B es beta-alanina. Los aminoácidos en el EP pueden tener una estereoquímica D o L.

20 El EP puede consistir en KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK (SEQ ID NO:7), KKRK (SEQ ID NO:8), KRKK (SEQ ID NO:9), KRRK (SEQ ID NO:10), RKKR (SEQ ID NO:11), RRRR (SEQ ID NO:12), KGKK (SEQ ID NO:13), KKGK (SEQ ID NO:14), KKKKK (SEQ ID NO:18), KKKRK (SEQ ID NO:19), KBKKB (SEQ ID NO:24), KKKRKV (SEQ ID NO:33), PKKKRKV (SEQ ID NO:42), PGKKRKV (SEQ ID NO:Z43), PKGKRKV (SEQ ID
25 NO:Z44), PKKGRKV (SEQ ID NO:Z45), PKKKGKV (SEQ ID NO:46), PKKKRGV (SEQ ID NO:47) o PKKKRKG (SEQ ID NO:48). El EP puede consistir en PKKKRKV (SEQ ID NO:42), RR, RRR, RHR, RBR, RBRBR (SEQ ID NO:49), RBHBR (SEQ ID NO:50) o HBRBH (SEQ ID NO:51), en donde B es beta-alanina. Los aminoácidos en el EP pueden tener una estereoquímica D o L.

30 El EP puede comprender una secuencia de aminoácidos identificada en la técnica como una secuencia de localización nuclear (NLS). El EP puede consistir en una secuencia de aminoácidos identificada en la técnica como una secuencia de localización nuclear (NLS). El EP puede

comprender una NLS que comprende la secuencia de aminoácidos PKKKRKV (SEQ ID NO: 42). El EP puede consistir en una NLS que comprende la secuencia de aminoácidos PKKKRKV (SEQ ID NO: 42). El EP puede comprender una NLS que comprende una secuencia de aminoácidos seleccionada de NLSKRPAAIKKAGQAKKKK (SEQ ID NO:52), PAAKRVKLD (SEQ ID NO:53), RQRRNELKRSF (SEQ ID NO:54), RMRKFKNKGKDTAELRRRRVEVSVELR (SEQ ID NO:55), KAKKDEQILKRRNV (SEQ ID NO:56), VSRKRPRP (SEQ ID NO:57), PPKKARED (SEQ ID NO:58), PPKKKKPL (SEQ ID NO:59), SALIKKKKKMAP (SEQ ID NO:60), DRLRR (SEQ ID NO:61), PKQKKRK (SEQ ID NO:62), RKLKKKIKKL (SEQ ID NO:63), REKKKFLKRR (SEQ ID NO:64), KRKGDEVDGVDEVAKKKSKK (SEQ ID NO:65) y RKCLQAGMNLEARKTKK (SEQ ID NO:66). El EP puede consistir en una NLS que comprende una secuencia de aminoácidos seleccionada de NLSKRPAAIKKAGQAKKKK (SEQ ID NO:52), PAAKRVKLD (SEQ ID NO:53), RQRRNELKRSF (SEQ ID NO:54), RMRKFKNKGKDTAELRRRRVEVSVELR (SEQ ID NO:55), KAKKDEQILKRRNV (SEQ ID NO:56), VSRKRPRP (SEQ ID NO:57), PPKKARED (SEQ ID NO:58), PPKKKKPL (SEQ ID NO:59), SALIKKKKKMAP (SEQ ID NO:60), DRLRR (SEQ ID NO:61), PKQKKRK (SEQ ID NO:62), RKLKKKIKKL (SEQ ID NO:63), REKKKFLKRR (SEQ ID NO:64), KRKGDEVDGVDEVAKKKSKK (SEQ ID NO:65) y RKCLQAGMNLEARKTKK (SEQ ID NO:66).

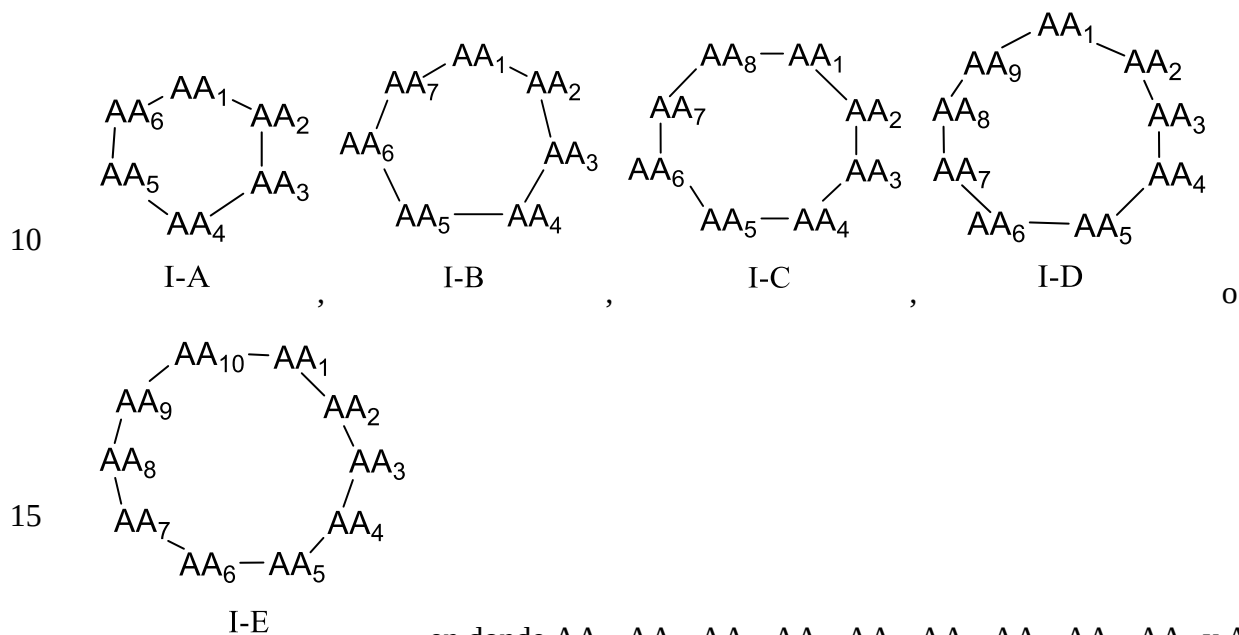
Todas las secuencias exocíclicas pueden contener también un grupo acetilo N-terminal. Por tanto, por ejemplo, el EP puede tener la estructura: Ac-PKKKRKV (SEQ ID NO: 42).

Péptidos penetrantes de células (CPP)

El péptido penetrante de células (CPP) puede comprender de 6 a 20 residuos de aminoácidos. El péptido penetrante de células puede ser un péptido penetrante de células cíclico (cCPP). El cCPP es capaz de penetrar una membrana celular. Un péptido exocíclico (EP) puede conjugarse con el cCPP, y el constructo resultante puede denominarse vehículo de escape endosómico (EEV). El cCPP puede dirigir una carga (por ejemplo, un resto terapéutico (TM) tal como un oligonucleótido, péptido o molécula pequeña) para penetrar la membrana de una célula. El cCPP puede suministrar la carga al citosol de la célula. El cCPP puede suministrar la carga a una ubicación celular donde se encuentra un objetivo (por ejemplo, pre-ARNm). Para conjugar el cCPP a una carga (por ejemplo, péptido, oligonucleótido o molécula pequeña), puede reemplazarse al menos un enlace o par solitario de electrones en el cCPP.

El número total de residuos de aminoácidos en el cCPP está en el intervalo de 6 a 20 residuos de

aminoácidos, por ejemplo, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 residuos de aminoácidos, lo que incluye todos los intervalos y subintervalos intermedios. El cCPP puede comprender 6 a 13 residuos de aminoácidos. El cCPP descrito en la presente descripción puede comprender 6 a 10 aminoácidos. A manera de ejemplo, el cCPP que comprende 6-10 residuos de aminoácidos puede tener una estructura de acuerdo con cualquiera de las fórmulas I-A a I-E:



, en donde AA₁, AA₂, AA₃, AA₄, AA₅, AA₆, AA₇, AA₈, AA₉ y AA₁₀ son residuos de aminoácidos.

El cCPP puede comprender 6 a 8 aminoácidos. El cCPP puede comprender 8 aminoácidos.

20 Cada aminoácido en el cCPP puede ser un aminoácido natural o no natural. El término “aminoácido no natural” se refiere a un compuesto orgánico que es un congénere de un aminoácido natural porque tiene una estructura similar a un aminoácido natural de modo que imita la estructura y la reactividad de un aminoácido natural. El aminoácido no natural puede ser un aminoácido modificado y/o análogo de aminoácidos, que no es uno de los 20 aminoácidos comunes de origen natural o los aminoácidos naturales raros selenocisteína o pirrolisina. Los aminoácidos no naturales también pueden ser un isómero D de un aminoácido natural. Ejemplos de aminoácidos adecuados incluyen, pero no se limitan a, alanina, alosoleucina, arginina, citrulina, asparagina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, naftilalanina, fenilalanina, prolina, ácido piroglutámico, serina, treonina, triptófano, tirosina, valina, un derivado de estos, o sus combinaciones. Estos, y otros aminoácidos, se enumeran en la Tabla 5 junto con sus abreviaturas usadas en la presente descripción.

25

30

Tabla 4. Abreviaturas de aminoácidos

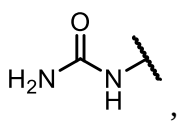
	Aminoácido	Abreviaturas* L-aminoácido	Abreviaturas* D-aminoácido
	Ácido 2-[2-[2-aminoeto]eto]acético	AEAA, miniPEG, PEG2	NA
5	Alanina	Ala (A)	ala (a)
	Alo-isoleucina	Aile	Aile
	Arginina	Arg (R)	arg (r)
	Asparagina	Asn (N)	asn (n)
	ácido aspártico	Asp (D)	asp (d)
	Cisteína	Cys (C)	cys (c)
	Citrulina	Cit	Cit
	Ciclohexilalanina	Cha	cha
10	Ácido 2,3-diaminopropiónico	Dap	dap
	4-fluorofenilalanina	Fpa (Σ)	pfa
	ácido glutámico	Glu (E)	glu (e)
	glutamina	Gln (Q)	gln (q)
	glicina	Gly (G)	gly (g)
	histidina	His (H)	his (h)
	Homoprolina (también conocida como ácido pipecólico)	Pip (Θ)	pip (θ)
15	isoleucina	Ile (I)	ile (i)
	leucina	Leu (L)	leu (l)
	lisina	Lys (K)	lys (k)
	metionina	Met (M)	met (m)
	3-(2-naftil)-alanina	Nal (Φ)	nal (ϕ)
	3-(1-naftil)-alanina	1-Nal	1-nal
	norleucina	Nle (Ω)	nle
20	fenilalanina	Phe (F)	phe (f)
	fenilglicina	Phg (Ψ)	phg
	4-(fosfonodifluorometil)fenilalanina	F ₂ Pmp (Λ)	f ₂ pmp
	prolina	Pro (P)	pro (p)
	sarcosina	Sar (Ξ)	sar
	selenocisteína	Sec (U)	sec (u)
	serina	Ser (S)	ser (s)
25	treonina	Thr (T)	thr (y)
	tirosina	Tyr (Y)	tyr (y)
	triptófano	Trp (W)	trp (w)
	valina	Val (V)	val (v)
	Terc-butil-alanina	Tle	tle
	Penicilamina	Pen	Pen
	Homoarginina	HomoArg	homoarg
30	Nicotinil-lisina	Lys(NIC)	lys(NIC)
	Trifluoroacetil-lisina	Lys(TFA)	lys(TFA)
	Metil-leucina	MeLeu	meLeu
	3-(3-benzotienil)-alanina	Bta	bta

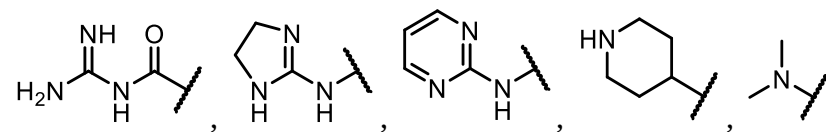
* abreviaturas de una sola letra: las letras en mayúsculas indican la forma de L-aminoácido, la letra en minúscula indica la forma de D-aminoácido.

Como se usa en la presente, “polietilenglicol” y “PEG” se usan indistintamente. “PEG_m” y “PEG_m”, son, o se derivan de, una molécula de la fórmula HO(CO)-(CH₂)_n-(OCH₂CH₂)_m-NH₂, donde n es cualquier número entero de 1 a 5, y m es cualquier número entero de 1 a 23. En modalidades, n es 1 o 2. En modalidades, n es 1. En modalidades, n es 2. En modalidades, n es 1 y m es 2. En modalidades, n es 2 y m es 2. En modalidades, n es 1 y m es 4. En modalidades, n es 2 y m es 4. En modalidades, n es 1 y m es 12. En modalidades, n es 2 y m es 12.

Como se usa en la presente, “miniPEG_m” o “miniPEG_m” son, o se derivan de, una molécula de la fórmula HO(CO)-(CH₂)_n-(OCH₂CH₂)_m-NH₂, donde n es 1, y m es cualquier número entero de 1 a 23. Por ejemplo, “miniPEG₂” o “miniPEG₂” es, o se deriva de, ácido (2-[2-[2-aminoetoxi]etoxi]acético), y “miniPEG₄” o “miniPEG₄” es, o se deriva de, HO(CO)-(CH₂)_n-(OCH₂CH₂)_m-NH₂, donde n es 1 y m es 4.

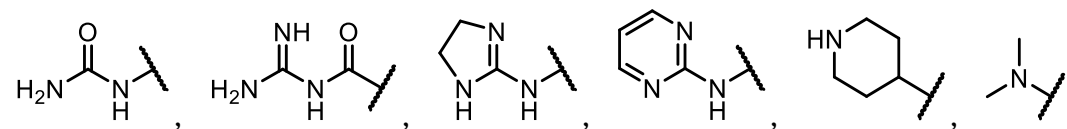
El cCPP puede comprender de 4 a 20 aminoácidos, en donde: (i) al menos un aminoácido tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta; (ii) al menos un

aminoácido no tiene cadena lateral o una cadena lateral que comprende ,

, o una forma protonada de esta; y

(iii) al menos dos aminoácidos tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático.

Al menos dos aminoácidos pueden no tener cadena lateral o una cadena lateral que comprende

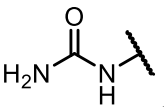
, o una forma

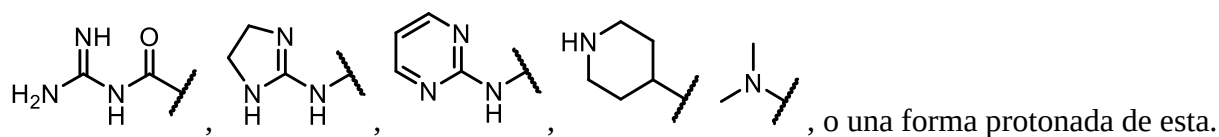
protonada de esta. Como se usa en la presente, cuando no está presente ninguna cadena lateral, el aminoácido tiene dos átomos de hidrógeno en el o los átomos de carbono (por ejemplo, -CH₂-) que enlazan la amina y el ácido carboxílico.

El aminoácido que no tiene cadena lateral puede ser glicina o β-alanina.

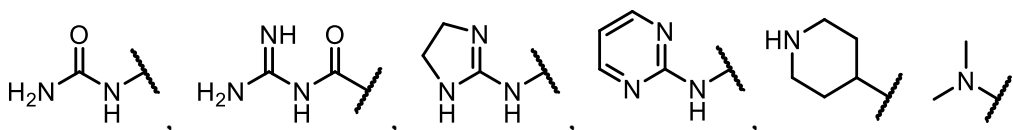
El cCPP puede comprender de 6 a 20 residuos de aminoácidos que forman el cCPP, en donde: (i) al

menos un aminoácido puede ser residuos de glicina, β-alanina, o ácido 4-aminobutírico; (ii) al menos un aminoácido puede tener una cadena lateral que comprende un grupo arilo o heteroarilo; y (iii) al

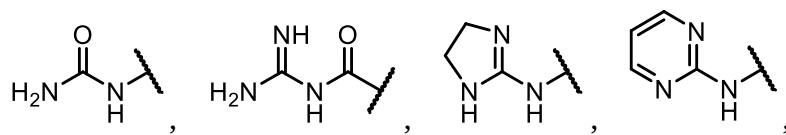
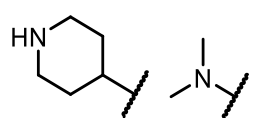
menos un aminoácido tiene una cadena lateral que comprende un grupo guanidina, ,



El cCPP puede comprender de 6 a 20 residuos de aminoácidos que forman el cCPP, en donde: (i) al menos dos aminoácidos pueden ser independientemente residuos de glicina, β-alanina o ácido 4-aminobutírico; (ii) al menos un aminoácido puede tener una cadena lateral que comprende un grupo arilo o heteroarilo; y (iii) al menos un aminoácido tiene una cadena lateral que comprende un grupo

guanidina, , o una forma protonada de esta.

El cCPP puede comprender de 6 a 20 residuos de aminoácidos que forman el cCPP, en donde: (i) al menos tres aminoácidos pueden ser independientemente residuos de glicina, β-alanina o ácido 4-aminobutírico; (ii) al menos un aminoácido puede tener una cadena lateral que comprende un grupo aromático o heteroaromático; y (iii) al menos un aminoácido puede tener una cadena lateral que

comprende un grupo guanidina, , , o una forma protonada de esta.

Residuos de glicina y aminoácidos relacionados

El cCPP puede comprender (i) 1, 2, 3, 4, 5 o 6 residuos de glicina, β-alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 2 residuos de glicina, β-alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3 residuos de glicina, β-alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 4 residuos de glicina, β-alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 5 residuos de glicina, β-alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 6

residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3, 4 o 5 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3 o 4 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones.

5 El cCPP puede comprender (i) 1, 2, 3, 4, 5 o 6 residuos de glicina. El cCPP puede comprender (i) 2 residuos de glicina. El cCPP puede comprender (i) 3 residuos de glicina. El cCPP puede comprender (i) 4 residuos de glicina. El cCPP puede comprender (i) 5 residuos de glicina. El cCPP puede comprender (i) 6 residuos de glicina. El cCPP puede comprender (i) 3, 4 o 5 residuos de glicina. El cCPP puede comprender (i) 3 o 4 residuos de glicina. El cCPP puede comprender (i) 2 o 3 residuos de glicina. El cCPP puede comprender (i) 1 o 2 residuos de glicina.

10 El cCPP puede comprender (i) 3, 4, 5 o 6 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 4 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 5 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 6 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3, 4 o 5 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones. El cCPP puede comprender (i) 3 o 4 residuos de glicina, β -alanina, ácido 4-aminobutírico, o sus combinaciones.

15 El cCPP puede comprender al menos tres residuos de glicina. El cCPP puede comprender (i) 3, 4, 5 o 6 residuos de glicina. El cCPP puede comprender (i) 3 residuos de glicina. El cCPP puede comprender (i) 4 residuos de glicina. El cCPP puede comprender (i) 5 residuos de glicina. El cCPP puede comprender (i) 6 residuos de glicina. El cCPP puede comprender (i) 3, 4 o 5 residuos de glicina. El cCPP puede comprender (i) 3 o 4 residuos de glicina

20 En modalidades, ninguno de los residuos de glicina, β -alanina o ácido 4-aminobutírico en el cCPP son contiguos. Dos o tres residuos de glicina, β -alanina o ácido 4-aminobutírico pueden ser contiguos. Dos residuos de glicina, β -alanina o ácido 4-aminobutírico pueden ser contiguos.

25 En modalidades, ninguno de los residuos de glicina en el cCPP son contiguos. Cada residuo de glicina en el cCPP puede separarse mediante un residuo de aminoácido que no puede ser glicina.

30 Dos o tres residuos de glicina pueden ser contiguos. Dos residuos de glicina pueden ser contiguos.

Cadenas laterales de aminoácidos con un grupo aromático o heteroaromático

El cCPP puede comprender (ii) 2, 3, 4, 5 o 6 residuos de aminoácidos que tienen independientemente

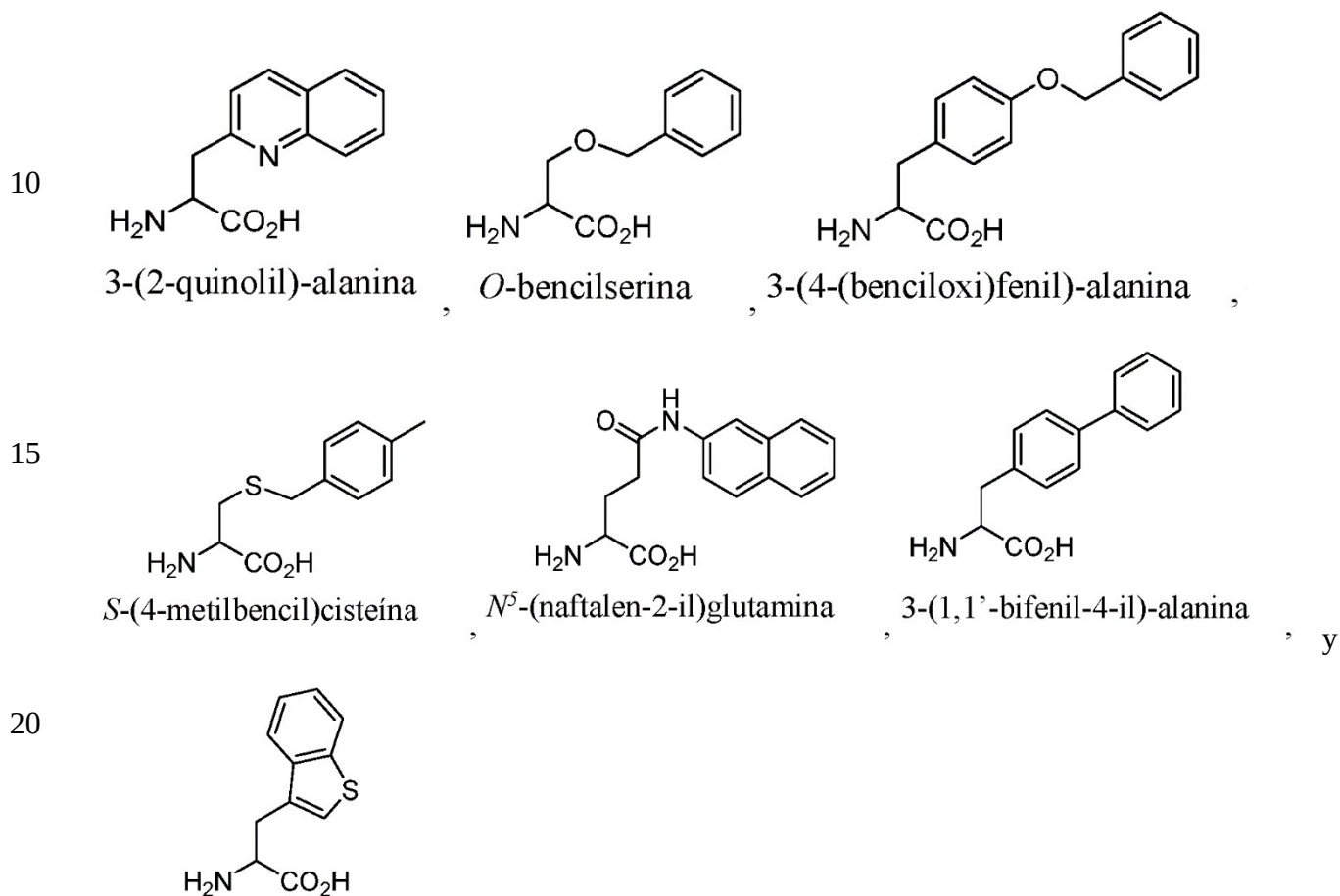
una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 2 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 3 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 4 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 5 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 6 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 2, 3 o 4 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático. El cCPP puede comprender (ii) 2 o 3 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático.

El cCPP puede comprender (ii) 2, 3, 4, 5 o 6 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 2 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 3 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 4 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 5 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 6 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 2, 3 o 4 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El cCPP puede comprender (ii) 2 o 3 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo aromático. El grupo aromático puede ser un arilo de 6 a 14 miembros. El arilo puede ser fenilo, naftilo o antraceno, cada uno de los cuales se sustituye opcionalmente. El arilo puede ser fenilo o naftilo, cada uno de los cuales se sustituye opcionalmente. El grupo heteroaromático puede ser un heteroarilo de 6 a 14 miembros que tiene 1, 2 o 3 heteroátomos seleccionados de N, O y S. El heteroarilo puede ser piridilo, quinolilo o isoquinolilo.

El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático puede ser cada uno independientemente bis(homonaftilalanina), homonaftilalanina,

naftilalanina, fenilglicina, bis(homofenilalanina), homofenilalanina, fenilalanina, triptófano, 3-(3-benzotienil)-alanina, 3-(2-quinolil)-alanina, O-bencilserina, 3-(4-(benciloxi)fenil)-alanina, S-(4-metilbencil)cisteína, N-(naftalen-2-il)glutamina, 3-(1,1'-bifenil-4-il)-alanina, 3-(3-benzotienil)-alanina o tirosina, cada uno de los cuales se sustituye opcionalmente con uno o más sustituyentes.

5 El aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático puede seleccionarse cada uno independientemente de:



25 , en donde la H en el extremo N-terminal y/o la H en el extremo C-terminal se reemplazan por un enlace peptídico.

El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático puede ser cada uno independientemente un residuo de fenilalanina, naftilalanina, fenilglicina, homofenilalanina, homonaftilalanina, bis(homofenilalanina), bis-(homonaftilalanina), triptófano o tirosina, cada uno de los cuales se sustituye opcionalmente con uno o más sustituyentes.

30 El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser cada uno independientemente un residuo de tirosina, fenilalanina, 1-naftilalanina, 2-naftilalanina, triptófano, 3-benzotienilalanina, 4-fenilfenilalanina, 3,4-difluorofenilalanina, 4-

trifluorometilfenilalanina, 2,3,4,5,6-pentafluorofenilalanina, homofenilalanina, β -homofenilalanina, 4-terc-butil-fenilalanina, 4-piridinilalanina, 3-piridinilalanina, 4-metilfenilalanina, 4-fluorofenilalanina, 4-clorofenilalanina, 3-(9-antrilo)-alanina. El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser cada uno independientemente un residuo de fenilalanina, naftilalanina, fenilglicina, homofenilalanina u homonaftilalanina, cada uno de los cuales se sustituye opcionalmente con uno o más sustituyentes. El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser cada uno independientemente un residuo de fenilalanina, naftilalanina, homofenilalanina, homonaftilalanina, bis(homonaftilalanina) o bis(homonaftilalanina), cada uno de los cuales se sustituye opcionalmente con uno o más sustituyentes. El residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser cada uno independientemente un residuo de fenilalanina o naftilalanina, cada uno de los cuales se sustituye opcionalmente con uno o más sustituyentes. Al menos un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser un residuo de fenilalanina. Al menos dos residuos de aminoácidos que tienen una cadena lateral que comprende un grupo aromático pueden ser residuos de fenilalanina. Cada residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático puede ser un residuo de fenilalanina.

En modalidades, ninguno de los aminoácidos que tienen la cadena lateral que comprende el grupo aromático o heteroaromático son contiguos. Dos aminoácidos que tienen la cadena lateral que comprende el grupo aromático o heteroaromático pueden ser contiguos. Dos aminoácidos contiguos pueden tener una estereoquímica opuesta. Los dos aminoácidos contiguos pueden tener la misma estereoquímica. Tres aminoácidos que tienen la cadena lateral que comprende el grupo aromático o heteroaromático pueden ser contiguos. Tres aminoácidos contiguos pueden tener la misma estereoquímica. Tres aminoácidos contiguos pueden tener una estereoquímica alterna.

Los residuos de aminoácidos que comprenden grupos aromáticos o heteroaromáticos pueden ser L-aminoácidos. Los residuos de aminoácidos que comprenden grupos aromáticos o heteroaromáticos pueden ser D-aminoácidos. Los residuos de aminoácidos que comprenden grupos aromáticos o heteroaromáticos pueden ser una mezcla de D y L-aminoácidos.

El sustituyente opcional puede ser cualquier átomo o grupo que no reduzca significativamente (por ejemplo, en más del 50 %) la eficiencia de suministro citosólico del cCPP, por ejemplo, en comparación con una secuencia de cualquier otra manera idéntica que no tiene el sustituyente. El sustituyente opcional puede ser un sustituyente hidrófobo o un sustituyente hidrófilo. El sustituyente opcional puede ser un sustituyente hidrófobo. El sustituyente puede aumentar el área superficial

accesible al solvente (como se define en la presente descripción) del aminoácido hidrófobo. El sustituyente puede ser halógeno, alquilo, alquenilo, alquinilo, cicloalquilo, cicloalquenilo, cicloalquinilo, heterociclilo, arilo, heteroarilo, alcoxi, arilooxi, acilo, alquilcarbamoilo, alquilcarboxamidilo, alcoxicarbonilo, alquiltio o ariltio. El sustituyente puede ser halógeno.

5 Aunque no se desea limitarse por la teoría, se cree que los aminoácidos que tienen un grupo aromático o heteroaromático que tiene valores de hidrofobicidad más altos (es decir, aminoácidos que tienen cadenas laterales que comprenden grupos aromáticos o heteroaromáticos) pueden mejorar la eficiencia del suministro citosólico de un cCPP con relación a los aminoácidos que tienen un valor de hidrofobicidad más bajo. Cada aminoácido hidrófobo puede tener independientemente un valor de hidrofobicidad mayor que el de la glicina. Cada aminoácido hidrófobo puede ser independientemente un aminoácido hidrófobo que tiene un valor de hidrofobicidad mayor que el de la alanina. Cada aminoácido hidrófobo puede tener independientemente un valor de hidrofobicidad mayor o igual a la fenilalanina. La hidrofobicidad puede medirse mediante el uso de escalas de hidrofobicidad conocidas en la técnica. La Tabla 5 enumera los valores de hidrofobicidad para diversos aminoácidos como se informa por Eisenberg y Weiss (Proc. Natl. Acad. Sci. U. S. A. 1984;81(1):140–144), Engleman, y otros (Ann. Rev. of Biophys. Biophys. Chem. 1986;1986(15):321–53), Kite y Doolittle (J. Mol. Biol. 1982;157(1):105–132), Hoop y Woods (Proc. Natl. Acad. Sci. U. S. A. 1981;78(6):3,824–3,828) y Janin (Nature. 1979;277(5696):491–492), la totalidad de cada uno de los cuales se incorpora en la presente descripción como referencia. La hidrofobicidad puede medirse mediante el uso de la escala de hidrofobicidad informada en Engleman, y otros

Tabla 5. Hidrofobicidad del aminoácido

Aminoácido	Grupo	Eisenberg y Weiss	Engleman y otros	Kirie y Doolittle	Aro y Woods	Janin
Ile	No polar	0.73	3.1	4.5	-1.8	0.7
Phe	No polar	0.61	3.7	2.8	-2.5	0.5
Val	No polar	0.54	2.6	4.2	-1.5	0.6
Leu	No polar	0.53	2.8	3.8	-1.8	0.5
Trp	No polar	0.37	1.9	-0.9	-3.4	0.3
Met	No polar	0.26	3.4	1.9	-1.3	0.4
Ala	No polar	0.25	1.6	1.8	-0.5	0.3
Gly	No polar	0.16	1.0	-0.4	0.0	0.3
Cys	No cargado/Polar	0.04	2.0	2.5	-1.0	0.9
Tyr	No cargado/Polar	0.02	-0.7	-1.3	-2.3	-0.4

Pro	No polar	-0.07	-0.2	-1.6	0.0	-0.3
Thr	No cargado/Polar	-0.18	1.2	-0.7	-0.4	-0.2
Ser	No cargado/Polar	-0.26	0.6	-0.8	0.3	-0.1
His	Cargado	-0.40	-3.0	-3.2	-0.5	-0.1
Glu	Cargado	-0.62	-8.2	-3.5	3.0	-0.7
Asn	No cargado/Polar	-0.64	-4.8	-3.5	0.2	-0.5
Gln	No cargado/Polar	-0.69	-4.1	-3.5	0.2	-0.7
Asp	Cargado	-0.72	-9.2	-3.5	3.0	-0.6
Lys	Cargado	-1.10	-8.8	-3.9	3.0	-1.8
Arg	Cargado	-1.80	-12.3	-4.5	3.0	-1.4

El tamaño de los grupos aromáticos o heteroaromáticos puede seleccionarse para mejorar la eficiencia del suministro citosólico del cCPP. Aunque no se desea limitarse por la teoría, se cree que un grupo aromático o heteroaromático más grande en la cadena lateral de un aminoácido puede mejorar la eficiencia del suministro citosólico en comparación con una secuencia de cualquier otra manera idéntica que tiene un aminoácido hidrófobo más pequeño. El tamaño del aminoácido hidrófobo puede medirse en términos de peso molecular del aminoácido hidrófobo, los efectos estéricos del aminoácido hidrófobo, el área superficial accesible al solvente (SASA) de la cadena lateral, o sus combinaciones. El tamaño del aminoácido hidrófobo puede medirse en términos del peso molecular del aminoácido hidrófobo, y el aminoácido hidrófobo más grande tiene una cadena lateral con un peso molecular de al menos aproximadamente 90 g/mol, o al menos aproximadamente 130 g/mol o al menos aproximadamente 141 g/mol. El tamaño del aminoácido puede medirse en términos del SASA de la cadena lateral hidrófoba. El aminoácido hidrófobo puede tener una cadena lateral con un SASA mayor o igual que la alanina, o mayor o igual que la glicina. Los aminoácidos hidrófobos más grandes pueden tener una cadena lateral con un SASA mayor que la alanina, o mayor que la glicina. El aminoácido hidrófobo puede tener un grupo aromático o heteroaromático con un SASA mayor o igual a aproximadamente la del ácido piperidin-2-carboxílico, mayor o igual a aproximadamente la del triptófano, mayor o igual a aproximadamente la de la fenilalanina, o mayor o igual a aproximadamente la de la naftilalanina. Un primer aminoácido hidrófobo (AA_{H1}) puede tener una cadena lateral con un SASA de al menos aproximadamente 200 Å², al menos aproximadamente 210 Å², al menos aproximadamente 220 Å², al menos aproximadamente 240 Å², al menos aproximadamente 250 Å², al menos aproximadamente 260 Å², al menos aproximadamente

270 Å², al menos aproximadamente 280 Å², al menos aproximadamente 290 Å², al menos aproximadamente 300 Å², al menos aproximadamente 310 Å², al menos aproximadamente 320 Å² o al menos aproximadamente 330 Å². Un segundo aminoácido hidrófobo (AA_{H2}) puede tener una cadena lateral con un SASA de al menos aproximadamente 200 Å², al menos aproximadamente 210 Å², al menos aproximadamente 220 Å², al menos aproximadamente 240 Å², al menos aproximadamente 250 Å², al menos aproximadamente 260 Å², al menos aproximadamente 270 Å², al menos aproximadamente 280 Å², al menos aproximadamente 290 Å², al menos aproximadamente 300 Å², al menos aproximadamente 310 Å², al menos aproximadamente 320 Å² o al menos aproximadamente 330 Å². Las cadenas laterales de AA_{H1} y AA_{H2} pueden tener un SASA combinado de al menos aproximadamente 350 Å², al menos aproximadamente 360 Å², al menos aproximadamente 370 Å², al menos aproximadamente 380 Å², al menos aproximadamente 390 Å², al menos aproximadamente 400 Å², al menos aproximadamente 410 Å², al menos aproximadamente 420 Å², al menos aproximadamente 430 Å², al menos aproximadamente 440 Å², al menos aproximadamente 450 Å², al menos aproximadamente 460 Å², al menos aproximadamente 470 Å², al menos aproximadamente 480 Å², al menos aproximadamente 490 Å², más de aproximadamente 500 Å², al menos aproximadamente 510 Å², al menos aproximadamente 520 Å², al menos aproximadamente 530 Å², al menos aproximadamente 540 Å², al menos aproximadamente 550 Å², al menos aproximadamente 560 Å², al menos aproximadamente 570 Å², al menos aproximadamente 580 Å², al menos aproximadamente 590 Å², al menos aproximadamente 600 Å², al menos aproximadamente 610 Å², al menos aproximadamente 620 Å², al menos aproximadamente 630 Å², al menos aproximadamente 640 Å², mayor que aproximadamente 650 Å², al menos aproximadamente 660 Å², al menos aproximadamente 670 Å², al menos aproximadamente 680 Å², al menos aproximadamente 690 Å² o al menos aproximadamente 700 Å². AA_{H2} puede ser un residuo de aminoácido hidrófobo con una cadena lateral que tiene un SASA que es menor o igual que el SASA de la cadena lateral hidrófoba de AA_{H1}. A manera de ejemplo, y no de limitación, un cCPP que tiene un motivo Nal-Arg puede exhibir una eficiencia de suministro citosólico mejorada en comparación con un cCPP que tiene un motivo Phe-Arg de cualquier otra manera idéntico; un cCPP que tiene un motivo Phe-Nal-Arg puede exhibir una eficiencia de suministro citosólico mejorada en comparación con un cCPP que tiene un motivo Nal-Phe-Arg de cualquier otra manera idéntico; y un motivo phe-Nal-Arg puede exhibir una eficiencia de suministro citosólico mejorada en comparación con un cCPP que tiene un motivo nal-Phe-Arg de cualquier otra manera idéntico.

Como se usa en la presente, “área superficial hidrófoba” o “SASA” se refiere al área superficial

(informada como Ångstroms cuadrados; Å²) de la cadena lateral de un aminoácido que es accesible a un solvente, SASA puede calcularse mediante el uso del algoritmo de la “bola rodante” desarrollado por Shrake y Rupley (*J Mol Biol.* 79(2): 351–71), que se incorpora en la presente descripción como referencia en su totalidad para todos los propósitos. Este algoritmo usa una “esfera” de solvente de un radio particular para sondear la superficie de la molécula. Un valor típico de la esfera es 1.4 Å, que se aproxima al radio de una molécula de agua.

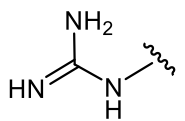
Los valores de SASA para ciertas cadenas laterales se muestran más abajo en la Tabla 6. Los valores de SASA descritos en la presente descripción se basan en los valores teóricos enumerados en la Tabla 6 más abajo, como se informó por Tien, y otros (PLOS ONE 8(11): e80635, disponible en doi.org/10.1371/journal.pone.0080635), que se incorpora en la presente descripción como referencia en su totalidad para todos los propósitos.

Tabla 6. Valores SASA de aminoácidos

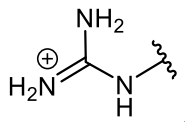
Residuo	Teórico	Empírico	Miller y otros (1987)	Rose y otros (1985)
Alanina	129.0	121.0	113.0	118.1
Arginina	274.0	265.0	241.0	256.0
Asparagina	195.0	187.0	158.0	165.5
Aspartato	193.0	187.0	151.0	158.7
Cisteína	167.0	148.0	140.0	146.1
Glutamato	223.0	214.0	183.0	186.2
Glutamina	225.0	214.0	189.0	193.2
Glicina	104.0	97.0	85.0	88.1
Histidina	224.0	216.0	194.0	202.5
Isoleucina	197.0	195.0	182.0	181.0
Leucina	201.0	191.0	180.0	193.1
Lisina	236.0	230.0	211.0	225.8
Metionina	224.0	203.0	204.0	203.4
Fenilalanina	240.0	228.0	218.0	222.8
Prolina	159.0	154.0	143.0	146.8
Serina	155.0	143.0	122.0	129.8
Treonina	172.0	163.0	146.0	152.5
Triptófano	285.0	264.0	259.0	266.3
Tirosina	263.0	255.0	229.0	236.8
Valina	174.0	165.0	160.0	164.5

Residuos de aminoácidos que tienen una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina o una forma protonada de esta

Como se usa en la presente, guanidina se refiere a la estructura:



Como se usa en la presente, una forma protonada de la guanidina se refiere a la estructura:

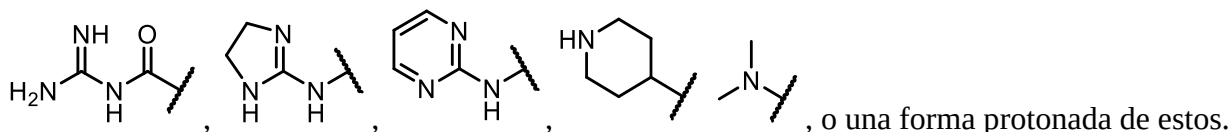


Los grupos de reemplazo de la guanidina se refieren a grupos funcionales en la cadena lateral de aminoácidos que se cargarán positivamente en o por encima del pH fisiológico o aquellos que pueden reproducir el enlace de hidrógeno que dona y acepta la actividad de los grupos guanidinio.

Los grupos de reemplazo de la guanidina facilitan la penetración celular y el suministro de agentes terapéuticos al tiempo que reducen la toxicidad asociada con los grupos guanidina o las formas protonadas de esta. El cCPP puede comprender al menos un aminoácido que tiene una cadena lateral que comprende un grupo de reemplazo de la guanidina o guanidinio. El cCPP puede comprender al menos dos aminoácidos que tienen una cadena lateral que comprende un grupo de reemplazo de la guanidina o guanidinio. El cCPP puede comprender al menos tres aminoácidos que tienen una cadena lateral que comprende un grupo de reemplazo de la guanidina o guanidinio.

El grupo guanidina o guanidinio puede ser un isómero de guanidina o guanidinio. El grupo de reemplazo de la guanidina o guanidinio puede ser menos básico que la guanidina.

Como se usa en la presente, un grupo de reemplazo de la guanidina se refiere a

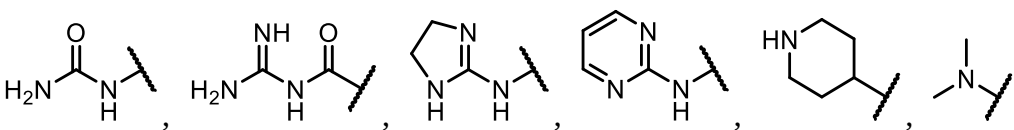


La descripción se refiere a un cCPP que comprende de 4 a 20 residuos de aminoácidos, en donde: (i) al menos un aminoácido tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta; (ii) al menos un residuo de aminoácido no tiene cadena lateral o una cadena lateral

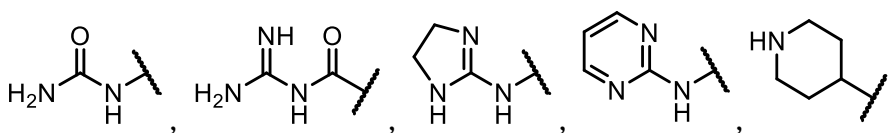
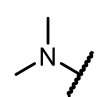
que comprende

una forma protonada de esta; y (iii) al menos dos residuos de aminoácidos tienen independientemente una cadena lateral que comprende un grupo aromático o heteroaromático.

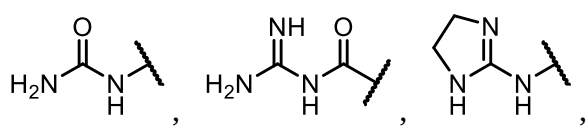
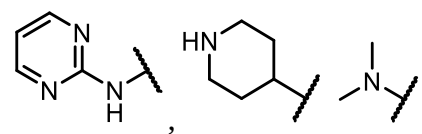
Al menos dos residuos de aminoácidos pueden no tener cadena lateral o tener una cadena lateral que

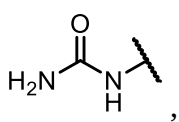
comprende , o una forma protonada de esta. Como se usa en la presente, cuando no está presente una cadena lateral, el residuo de aminoácido tiene dos átomos de hidrógeno en el o los átomos de carbono (por ejemplo, -CH₂-) que enlazan la amina y el ácido carboxílico.

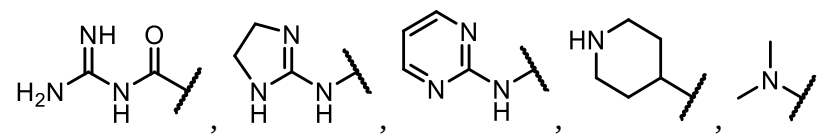
El cCPP puede comprender al menos un aminoácido que tiene una cadena lateral que comprende

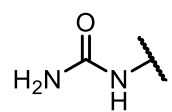
uno de los siguientes restos: , , o una forma protonada de esta.

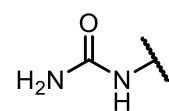
El cCPP puede comprender al menos dos aminoácidos, cada uno de los cuales tiene

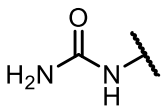
independientemente uno de los siguientes restos , , o una forma protonada de estos. Al menos dos aminoácidos pueden

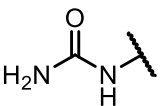
tener una cadena lateral que comprende el mismo resto seleccionado de: ,

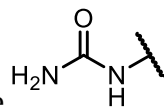
, o una forma protonada de esta. Al

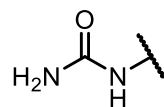
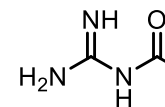
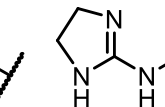
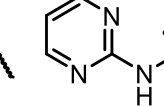
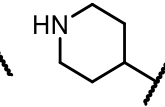
menos un aminoácido puede tener una cadena lateral que comprende , o una forma protonada de esta. Al menos dos aminoácidos pueden tener una cadena lateral que comprende

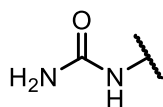
, o una forma protonada de esta. Uno, dos, tres o cuatro aminoácidos pueden tener una

cadena lateral que comprende , o una forma protonada de esta. Un aminoácido puede

5 tener una cadena lateral que comprende , o una forma protonada de esta. Dos

aminoácidos pueden tener una cadena lateral que comprende , o una forma protonada

10 de esta. , , , , , o una forma protonada de estos, puede unirse al extremo terminal de la cadena lateral de aminoácidos.

 puede unirse al extremo terminal de la cadena lateral de aminoácidos.

15 El cCPP puede comprender (iii) 2, 3, 4, 5 o 6 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 2 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 3 residuos de

20 aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 4 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 5 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 6 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 2, 3, 4 o 5 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) 2, 3 o 4

30 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede

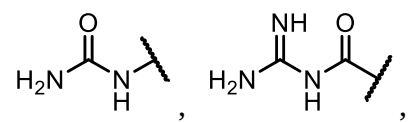
comprender (iii) 2 o 3 residuos de aminoácidos que tienen independientemente una cadena lateral que comprende un grupo guanidina, un grupo de reemplazo de la guanidina, o una forma protonada de esta. El cCPP puede comprender (iii) al menos un residuo de aminoácido que tiene una cadena lateral que comprende un grupo guanidina o una forma protonada de esta. El cCPP puede comprender (iii) dos residuos de aminoácidos que tienen una cadena lateral que comprende un grupo guanidina o una forma protonada de esta. El cCPP puede comprender (iii) tres residuos de aminoácidos que tienen una cadena lateral que comprende un grupo guanidina o una forma protonada de esta.

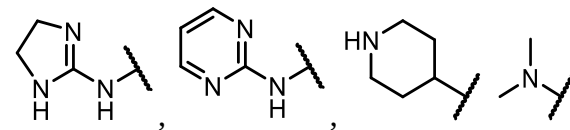
Los residuos de aminoácidos que pueden tener independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, que no son contiguos. Dos residuos de aminoácidos que pueden tener independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser contiguos. Tres residuos de aminoácidos que pueden tener independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser contiguos. Cuatro residuos de aminoácidos que pueden tener independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser contiguos. Los residuos de aminoácidos contiguos pueden tener la misma estereoquímica. Los aminoácidos contiguos pueden tener una estereoquímica alterna.

Los residuos de aminoácidos que tienen independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser L-aminoácidos. Los residuos de aminoácidos que tienen independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser aminoácidos D. Los residuos de aminoácidos que tienen independientemente la cadena lateral que comprende el grupo guanidina, el grupo de reemplazo de la guanidina, o la forma protonada de esta, pueden ser una mezcla de L- o D-aminoácidos.

Cada residuo de aminoácido que tiene la cadena lateral que comprende el grupo guanidina, o la forma protonada de esta, puede ser independientemente un residuo de arginina, homoarginina, ácido 2-amino-3-propiónico, ácido 2-amino-4-guanidinobutírico o una forma protonada de estos. Cada residuo de aminoácido que tiene la cadena lateral que comprende el grupo guanidina, o la forma protonada de esta, puede ser independientemente un residuo de arginina o una forma protonada de esta.

Cada aminoácido que tiene la cadena lateral que comprende un grupo de reemplazo de la guanidina,

o una forma protonada de esta, puede ser independientemente ,

5 , o una forma protonada de estos.

Sin estar limitado por la teoría, se plantea la hipótesis de que los grupos de reemplazo de la guanidina tienen una basicidad reducida con relación a la arginina, y en algunos casos no están cargados a pH fisiológico (por ejemplo, un -N(H)C(O)), y son capaces de mantener las interacciones de unión a hidrógeno bidentado con fosfolípidos en la membrana plasmática que se cree que facilitan la asociación efectiva de la membrana y la subsecuente internalización. Se cree también que la eliminación de la carga positiva reduce la toxicidad del cCPP.

Los expertos en la técnica apreciarán que los extremos N y/o C terminales de los aminoácidos hidrófobos aromáticos no naturales anteriores, tras la incorporación en los péptidos descritos en la presente descripción, forman enlaces amida.

El cCPP puede comprender un primer aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático y un segundo aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático, en donde un extremo N-terminal de una primera glicina forma un enlace peptídico con el primer aminoácido que tiene la cadena lateral que comprende el grupo aromático o heteroaromático, y un extremo C-terminal de la primera glicina forma un enlace peptídico con el segundo aminoácido que tiene la cadena lateral que comprende el grupo aromático o heteroaromático. Aunque por convención, el término “primer aminoácido” con frecuencia se refiere al aminoácido N-terminal de una secuencia peptídica, como se usa en la presente, “primer aminoácido” se usa para distinguir el aminoácido de referencia de otro aminoácido (por ejemplo, un “segundo aminoácido”) en el cCPP de manera que el término “primer aminoácido” puede o puede referirse a un aminoácido ubicado en el extremo N-terminal de la secuencia peptídica.

El cCPP puede comprender un extremo N-terminal de una segunda glicina que forma un enlace peptídico con un aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático, y un extremo C-terminal de la segunda glicina forma un enlace peptídico con un aminoácido que tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta.

El cCPP puede comprender un primer aminoácido que tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta, y un segundo aminoácido que tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta, en donde un extremo N-terminal de una tercera glicina forma un enlace peptídico con un primer aminoácido que tiene una
5 cadena lateral que comprende un grupo guanidina, o una forma protonada de esta, y un extremo C-terminal de la tercera glicina forma un enlace peptídico con un segundo aminoácido que tiene una cadena lateral que comprende un grupo guanidina, o una forma protonada de esta.

El cCPP puede comprender un residuo de asparagina, ácido aspártico, glutamina, ácido glutámico u homoglutamina. El cCPP puede comprender un residuo de asparagina. El cCPP puede comprender
10 un residuo de glutamina.

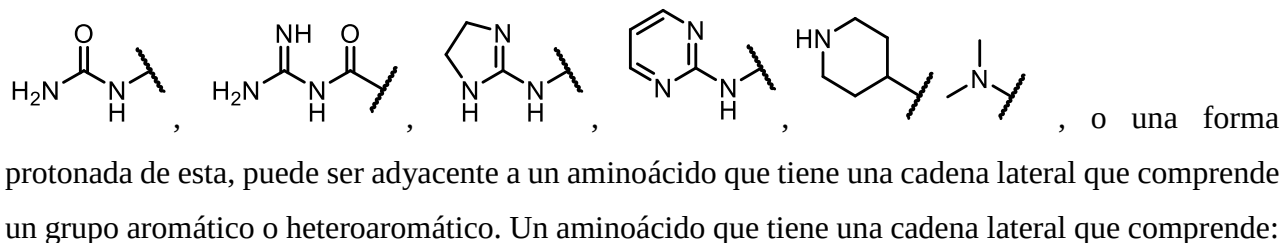
El cCPP puede comprender un residuo de tirosina, fenilalanina, 1-naftilalanina, 2-naftilalanina, triptófano, 3-benzotienilalanina, 4-fenilfenilalanina, 3,4-difluorofenilalanina, 4-trifluorometilfenilalanina, 2,3,4,5,6-pentafluorofenilalanina, homofenilalanina, β -homofenilalanina, 4-terc-butil-fenilalanina, 4-piridinilalanina, 3-piridinilalanina, 4-metilfenilalanina, 4-
15 fluorofenilalanina, 4-clorofenilalanina, 3-(9-antril)-alanina.

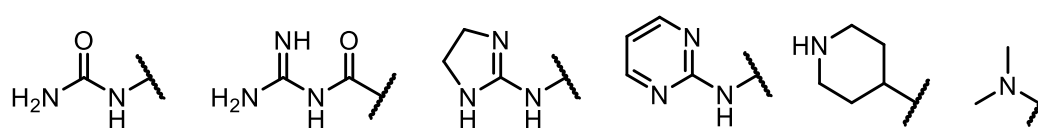
Aunque no se desea limitarse por la teoría, se cree que la quiralidad de los aminoácidos en los cCPP puede afectar la eficiencia de la captación citosólica. El cCPP puede comprender al menos un D aminoácido. El cCPP puede comprender uno a quince D aminoácidos. El cCPP puede comprender uno a diez D aminoácidos. El cCPP puede comprender 1, 2, 3 o 4 D aminoácidos. El cCPP puede
20 comprender 2, 3, 4, 5, 6, 7 u 8 aminoácidos contiguos que tienen quiralidad alterna D y L. El cCPP puede comprender tres aminoácidos contiguos que tienen la misma quiralidad. El cCPP puede comprender dos aminoácidos contiguos que tienen la misma quiralidad. Al menos dos de los aminoácidos pueden tener la quiralidad opuesta. Los al menos dos aminoácidos que tienen la quiralidad opuesta pueden ser adyacentes entre sí. Al menos tres aminoácidos pueden tener una
25 estereoquímica alterna entre sí. Los al menos tres aminoácidos que tienen la quiralidad alterna entre sí pueden ser adyacentes entre sí. Al menos cuatro aminoácidos tienen una estereoquímica alterna entre sí. Los al menos cuatro aminoácidos que tienen la quiralidad alterna entre sí pueden ser adyacentes entre sí. Al menos dos de los aminoácidos pueden tener la misma quiralidad. Al menos dos aminoácidos que tienen la misma quiralidad pueden ser adyacentes entre sí. Al menos dos
30 aminoácidos tienen la misma quiralidad y al menos dos aminoácidos tienen la quiralidad opuesta. Los al menos dos aminoácidos que tienen la quiralidad opuesta pueden ser adyacentes a los al menos dos aminoácidos que tienen la misma quiralidad. En consecuencia, los aminoácidos adyacentes en

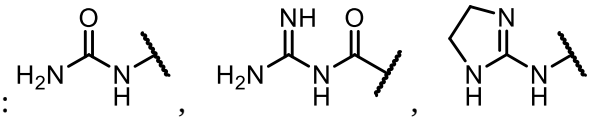
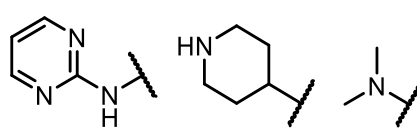
el cCPP pueden tener cualquiera de las siguientes secuencias: D-L; L-D; D-L-L-D; L-D-D-L; L-D-L-L-D; D-L-D-D-L; D-L-L-D-L; o L-D-D-L-D. Los residuos de aminoácidos que forman el cCPP pueden ser todos L-aminoácidos. Los residuos de aminoácidos que forman el cCPP pueden ser todos D-aminoácidos.

- 5 Al menos dos de los aminoácidos pueden tener una quiralidad diferente. Al menos dos aminoácidos que tienen una quiralidad diferente pueden ser adyacentes entre sí. Al menos tres aminoácidos pueden tener una quiralidad diferente con relación a un aminoácido adyacente. Al menos cuatro aminoácidos pueden tener una quiralidad diferente con relación a un aminoácido adyacente. Al menos dos aminoácidos tienen la misma quiralidad y al menos dos aminoácidos tienen una quiralidad diferente. Uno o más residuos de aminoácidos que forman el cCPP pueden ser aquirales. El cCPP puede comprender un motivo de 3, 4 o 5 aminoácidos, en donde dos aminoácidos que tienen la misma quiralidad pueden estar separados por un aminoácido aquiral. Los cCPP pueden comprender las siguientes secuencias: D-X-D; D-X-D-X; D-X-D-X-D; L-X-L; L-X-L-X; o L-X-L-X-L, en donde X es un aminoácido aquiral. El aminoácido aquiral puede ser glicina.

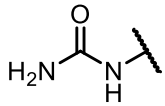
- 15 Un aminoácido que tiene una cadena lateral que comprende:

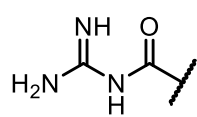
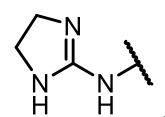
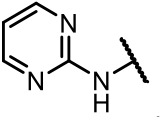
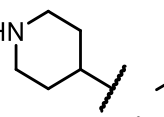
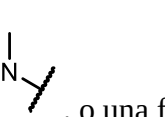


- 20 
 , o una forma protonada de esta, puede ser adyacente a al menos un aminoácido que tiene una cadena lateral que comprende una guanidina o una forma protonada de esta. Un aminoácido que tiene una cadena lateral que comprende una guanidina o una forma protonada de esta, puede ser adyacente a un aminoácido que tiene una cadena lateral que comprende un grupo aromático o heteroaromático. Dos aminoácidos

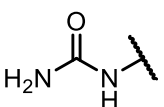
- 25 que tienen una cadena lateral que comprende: 
 , o formas protonadas de estos, pueden ser adyacentes entre sí. Dos
- 30 
 , o formas protonadas de estos, pueden ser adyacentes entre sí. Dos

aminoácidos que tienen una cadena lateral que comprende una guanidina o una forma protonada de esta, son adyacentes entre sí. Los cCPP pueden comprender al menos dos aminoácidos contiguos que tienen una cadena lateral que comprende un grupo aromático o heteroaromático y al menos dos

aminoácidos no adyacentes que tienen una cadena lateral que comprende: ,

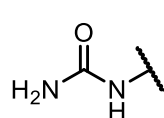
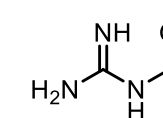
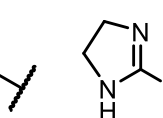
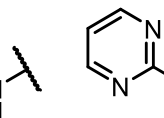
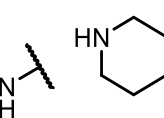
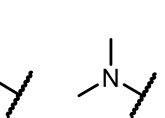
, , , , , o una forma protonada de esta. Los

cCPP pueden comprender al menos dos aminoácidos contiguos que tienen una cadena lateral que comprende un grupo aromático o heteroaromático y al menos dos aminoácidos no adyacentes que

tienen una cadena lateral que comprende , o una forma protonada de esta. Los

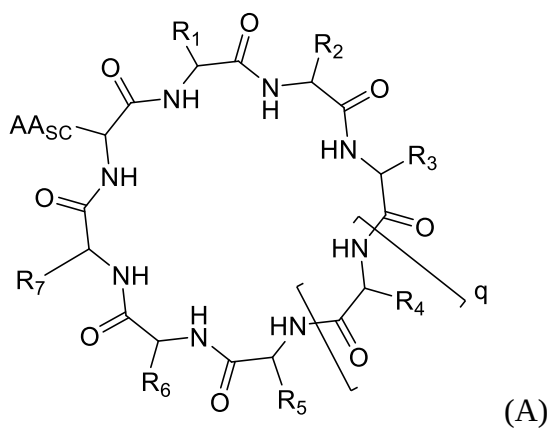
aminoácidos adyacentes pueden tener la misma quiralidad. Los aminoácidos adyacentes pueden tener la quiralidad opuesta. Otras combinaciones de aminoácidos pueden tener cualquier disposición de D y L aminoácidos, por ejemplo, cualquiera de las secuencias descritas en el párrafo anterior.

Al menos dos aminoácidos que tienen una cadena lateral que comprende:

, , , , , , o una forma

protonada de esta, se alternan con al menos dos aminoácidos que tienen una cadena lateral que comprende un grupo guanidina o una forma protonada de esta.

El cCPP puede comprender la estructura de fórmula (A):



o una forma protonada de esta,

en donde:

R₁, R₂ y R₃ son cada uno independientemente H o una cadena lateral aromática o heteroaromática de un aminoácido;

al menos uno de R₁, R₂ y R₃ es una cadena lateral aromática o heteroaromática de un aminoácido;

5 R₄, R₅, R₆, R₇ son independientemente H o una cadena lateral de aminoácido;

al menos uno de R₄, R₅, R₆, R₇ es la cadena lateral del ácido 3-guanidino-2-aminopropiónico, ácido 4-guanidino-2-aminobutanoico, arginina, homoarginina, N-metilarginina, N,N-dimetilarginina, ácido 2,3-diaminopropiónico, ácido 2,4-diaminobutanoico, lisina, N-metilisina, N,N-dimetilisina, N-etilisina, N,N,N-trimetilisina, 4-guanidinofenilalanina, citrulina, N,N-dimetilisina, β-homoarginina, 3-(1-piperidinil)alanina;

AA_{SC} es una cadena lateral de aminoácido; y

q es 1, 2, 3 o 4.

En modalidades, al menos uno de R₄, R₅, R₆, R₇ es independientemente una cadena lateral no aromática no cargada de un aminoácido. En modalidades, al menos uno de R₄, R₅, R₆, R₇ es
15 independientemente H o una cadena lateral de la citrulina.

En modalidades, se proporcionan compuestos que incluyen un péptido cíclico que tiene de 6 a 12 aminoácidos, en donde al menos dos aminoácidos del péptido cíclico son aminoácidos cargados, al menos dos aminoácidos del péptido cíclico son aminoácidos aromáticos hidrófobos, y al menos dos aminoácidos del péptido cíclico son aminoácidos no aromáticos no cargados. En modalidades, al
20 menos dos aminoácidos cargados del péptido cíclico son arginina. En modalidades, al menos dos aminoácidos aromáticos hidrófobos del péptido cíclico son fenilalanina o naftilalanina. En modalidades, al menos dos aminoácidos no aromáticos no cargados del péptido cíclico son citrulina o glicina.

En modalidades, el péptido cíclico de fórmula (A) no se selecciona de un péptido cíclico que tiene
25 una secuencia de SEQ ID NO: 89-117.

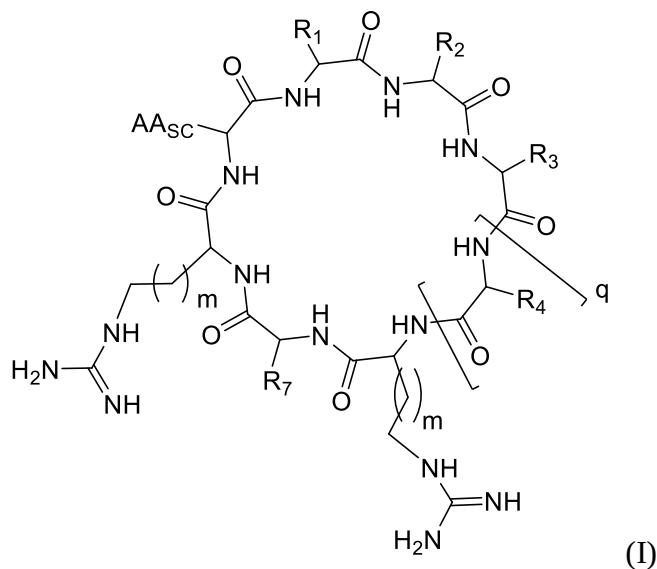
En modalidades, el péptido cíclico de fórmula (A) se selecciona de un péptido cíclico que tiene una secuencia de SEQ ID NO: 89-117.

Secuencias de CPP y SEQ ID NO					
FΦRRRQ	89	RRFRΦRQ	99	FΦRRRRQK	109
FΦRRRC	90	FRRRRΦQ	100	FΦRRRRQC	110
FΦRRRU	91	rRFRΦRQ	101	fΦRrRrRQ	111
RRRΦFQ	92	RRΦFRRQ	102	FΦRRRRRQ	112
RRRRΦF	93	CRRRRFWQ	103	RRRRΦFDΩC	113

FΦRRRR	94	FfΦRrRrQ	104	FΦRRR	114
FφrRrRq	95	FFΦRRRRQ	105	FWRRR	115
FφrRrRQ	96	RFRFRΦRQ	106	RRRΦF	116
FΦRRRRQ	97	URRRRrFWQ	107	RRRWF	117
fΦRrRrQ	98	CRRRRrFWQ	108		

Φ = L-naftilalanina; φ = D-naftilalanina; Ω = L-norleucina

El cCPP puede comprender la estructura de fórmula (I):



o una forma protonada de esta,

en donde:

R₁, R₂ y R₃ pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

al menos uno de R₁, R₂ y R₃ es una cadena lateral aromática o heteroaromática de un aminoácido;

R₄ y R₇ son independientemente H o una cadena lateral de aminoácido;

AA_{SC} es una cadena lateral de aminoácido;

q es 1, 2, 3 o 4; y

cada m es independientemente un número entero de 0, 1, 2 o 3.

R₁, R₂ y R₃ pueden ser cada uno independientemente H, -alquileo-arilo o -alquileo-heteroarilo.

R₁, R₂ y R₃ pueden ser cada uno independientemente H, -alquileo C₁₋₃-arilo o -alquileo C₁₋₃-heteroarilo. R₁, R₂ y R₃ pueden ser cada uno independientemente H o -alquileo-arilo. R₁, R₂ y R₃

pueden ser cada uno independientemente H o -alquileo C₁₋₃-arilo. El alquileo C₁₋₃ puede ser metileno. El arilo puede ser un arilo de 6 a 14 miembros. El heteroarilo puede ser un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O y S. El arilo puede

seleccionarse de fenilo, naftilo o antracenilo. El arilo puede ser fenilo o naftilo. El arilo puede ser fenilo. El heteroarilo puede ser piridilo, quinolilo e isoquinolilo. R₁, R₂ y R₃ pueden ser cada uno independientemente H, -alquileo C₁₋₃-Ph o -alquileo C₁₋₃-naftilo. R₁, R₂ y R₃ pueden ser cada uno independientemente H, -CH₂Ph o -CH₂naftilo. R₁, R₂ y R₃ pueden ser cada uno independientemente H o -CH₂Ph.

R₁, R₂ y R₃ pueden ser cada uno independientemente la cadena lateral de la tirosina, fenilalanina, 1-naftilalanina, 2-naftilalanina, triptófano, 3-benzotienilalanina, 4-fenilfenilalanina, 3,4-difluorofenilalanina, 4-trifluorometilfenilalanina, 2,3,4,5,6-pentafluorofenilalanina, homofenilalanina, β-homofenilalanina, 4-terc-butil-fenilalanina, 4-piridinilalanina, 3-piridinilalanina, 4-metilfenilalanina, 4-fluorofenilalanina, 4-clorofenilalanina, 3-(9-antril)-alanina.

R₁ puede ser la cadena lateral de la tirosina. R₁ puede ser la cadena lateral de la fenilalanina. R₁ puede ser la cadena lateral de la 1-naftilalanina. R₁ puede ser la cadena lateral de la 2-naftilalanina. R₁ puede ser la cadena lateral del triptófano. R₁ puede ser la cadena lateral de la 3-benzotienilalanina. R₁ puede ser la cadena lateral de la 4-fenilfenilalanina. R₁ puede ser la cadena lateral de la 3,4-difluorofenilalanina. R₁ puede ser la cadena lateral de la 4-trifluorometilfenilalanina. R₁ puede ser la cadena lateral de la 2,3,4,5,6-pentafluorofenilalanina. R₁ puede ser la cadena lateral de la homofenilalanina. R₁ puede ser la cadena lateral de la β-homofenilalanina. R₁ puede ser la cadena lateral de la 4-terc-butil-fenilalanina. R₁ puede ser la cadena lateral de la 4-piridinilalanina. R₁ puede ser la cadena lateral de la 3-piridinilalanina. R₁ puede ser la cadena lateral de la 4-metilfenilalanina. R₁ puede ser la cadena lateral de la 4-fluorofenilalanina. R₁ puede ser la cadena lateral de la 4-clorofenilalanina. R₁ puede ser la cadena lateral de la 3-(9-antril)-alanina.

R₂ puede ser la cadena lateral de la tirosina. R₂ puede ser la cadena lateral de la fenilalanina. R₂ puede ser la cadena lateral de la 1-naftilalanina. R₂ puede ser la cadena lateral de la 2-naftilalanina. R₂ puede ser la cadena lateral del triptófano. R₂ puede ser la cadena lateral de la 3-benzotienilalanina. R₂ puede ser la cadena lateral de la 4-fenilfenilalanina. R₂ puede ser la cadena lateral de la 3,4-difluorofenilalanina. R₂ puede ser la cadena lateral de la 4-trifluorometilfenilalanina. R₂ puede ser la cadena lateral de la 2,3,4,5,6-pentafluorofenilalanina. R₂ puede ser la cadena lateral de la homofenilalanina. R₂ puede ser la cadena lateral de la β-homofenilalanina. R₂ puede ser la cadena lateral de la 4-terc-butil-fenilalanina. R₂ puede ser la cadena lateral de la 4-piridinilalanina. R₂ puede ser la cadena lateral de la 3-piridinilalanina. R₂ puede ser la cadena lateral de la 4-metilfenilalanina. R₂ puede ser la cadena lateral de la 4-fluorofenilalanina. R₂ puede ser la cadena lateral de la 4-clorofenilalanina. R₂ puede ser la cadena lateral de la 3-(9-antril)-alanina.

R₃ puede ser la cadena lateral de la tirosina. R₃ puede ser la cadena lateral de la fenilalanina. R₃ puede ser la cadena lateral de la 1-naftilalanina. R₃ puede ser la cadena lateral de la 2-naftilalanina. R₃ puede ser la cadena lateral del triptófano. R₃ puede ser la cadena lateral de la 3-benzotienilalanina. R₃ puede ser la cadena lateral de la 4-fenilfenilalanina. R₃ puede ser la cadena lateral de la 3,4-difluorofenilalanina. R₃ puede ser la cadena lateral de la 4-trifluorometilfenilalanina. R₃ puede ser la cadena lateral de la 2,3,4,5,6-pentafluorofenilalanina. R₃ puede ser la cadena lateral de la homofenilalanina. R₃ puede ser la cadena lateral de la β-homofenilalanina. R₃ puede ser la cadena lateral de la 4-terc-butil-fenilalanina. R₃ puede ser la cadena lateral de la 4-piridinilalanina. R₃ puede ser la cadena lateral de la 3-piridinilalanina. R₃ puede ser la cadena lateral de la 4-metilfenilalanina. R₃ puede ser la cadena lateral de la 4-fluorofenilalanina. R₃ puede ser la cadena lateral de la 4-clorofenilalanina. R₃ puede ser la cadena lateral de la 3-(9-antril)-alanina.

R₄ puede ser H, -alquilenarilo, -alquilenoheteroarilo. R₄ puede ser H, -alquilenarilo, o -alquilenoheteroarilo. R₄ puede ser H o -alquilenarilo. R₄ puede ser H o -alquilenarilo. El alquilenarilo puede ser un metileno. El arilo puede ser un arilo de 6 a 14 miembros. El heteroarilo puede ser un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O y S. El arilo puede seleccionarse de fenilo, naftilo o antracenilo. El arilo puede ser fenilo o naftilo. El arilo puede ser fenilo. El heteroarilo puede ser piridilo, quinolilo e isoquinolilo. R₄ puede ser H, -alquilenarilo o -alquilenarilo. R₄ puede ser H o la cadena lateral de un aminoácido en la Tabla 4 o la Tabla 6. R₄ puede ser H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático. R₄ puede ser H, -CH₂Ph o -CH₂naftilo. R₄ puede ser H o -CH₂Ph.

R₅ puede ser H, -alquilenarilo, -alquilenoheteroarilo. R₅ puede ser H, -alquilenarilo, o -alquilenoheteroarilo. R₅ puede ser H o -alquilenarilo. R₅ puede ser H o -alquilenarilo. El alquilenarilo puede ser un metileno. El arilo puede ser un arilo de 6 a 14 miembros. El heteroarilo puede ser un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O y S. El arilo puede seleccionarse de fenilo, naftilo o antracenilo. El arilo puede ser fenilo o naftilo. El arilo puede ser fenilo. El heteroarilo puede ser piridilo, quinolilo e isoquinolilo. R₅ puede ser H, -alquilenarilo o -alquilenarilo. R₅ puede ser H o la cadena lateral de un aminoácido en la Tabla 4 o la Tabla 6. R₅ puede ser H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático. R₅ puede ser H, -CH₂Ph o -CH₂naftilo. R₅ puede ser H o -CH₂Ph.

R₆ puede ser H, -alquilenarilo, -alquilenoheteroarilo. R₆ puede ser H, -alquilenarilo, o -alquilenoheteroarilo. R₆ puede ser H o -alquilenarilo. R₆ puede ser H o -alquilenarilo. El alquilenarilo puede ser un metileno. El arilo puede ser un arilo de 6 a 14 miembros. El heteroarilo

puede ser un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O y S. El arilo puede seleccionarse de fenilo, naftilo o antracenilo. El arilo puede ser fenilo o naftilo. El arilo puede ser fenilo. El heteroarilo puede ser piridilo, quinolilo e isoquinolilo. R₆ puede ser H, -alquilenos C₁₋₃-Ph o -alquilenos C₁₋₃-naftilo. R₆ puede ser H o la cadena lateral de un aminoácido en la Tabla 4 o la Tabla 6. R₆ puede ser H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático. R₆ puede ser H, -CH₂Ph o -CH₂naftilo. R₆ puede ser H o -CH₂Ph. R₇ puede ser H, -alquilenos-arilo, -alquilenos-heteroarilo. R₇ puede ser H, -alquilenos C₁₋₃-arilo, o -alquilenos C₁₋₃-heteroarilo. R₇ puede ser H o -alquilenos-arilo. R₇ puede ser H o -alquilenos C₁₋₃-arilo. El alquilenos C₁₋₃ puede ser un metileno. El arilo puede ser un arilo de 6 a 14 miembros. El heteroarilo puede ser un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O y S. El arilo puede seleccionarse de fenilo, naftilo o antracenilo. El arilo puede ser fenilo o naftilo. El arilo puede ser fenilo. El heteroarilo puede ser piridilo, quinolilo e isoquinolilo. R₇ puede ser H, -alquilenos C₁₋₃-Ph o -alquilenos C₁₋₃-naftilo. R₇ puede ser H o la cadena lateral de un aminoácido en la Tabla 4 o la Tabla 6. R₇ puede ser H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático. R₇ puede ser H, -CH₂Ph o -CH₂naftilo. R₇ puede ser H o -CH₂Ph. Uno, dos o tres de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser -CH₂Ph. Uno de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ puede ser -CH₂Ph. Dos de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser -CH₂Ph. Tres de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser -CH₂Ph. Al menos uno de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ puede ser -CH₂Ph. No más de cuatro de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser -CH₂Ph. Uno, dos o tres de R₁, R₂, R₃ y R₄ son -CH₂Ph. Uno de R₁, R₂, R₃ y R₄ es -CH₂Ph. Dos de R₁, R₂, R₃ y R₄ son -CH₂Ph. Tres de R₁, R₂, R₃ y R₄ son -CH₂Ph. Al menos uno de R₁, R₂, R₃ y R₄ es -CH₂Ph. Uno, dos o tres de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser H. Uno de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ puede ser H. Dos de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ son H. Tres de R₁, R₂, R₃, R₅, R₆ y R₇ pueden ser H. Al menos uno de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser H. No más de tres de R₁, R₂, R₃, R₄, R₅, R₆ y R₇ pueden ser -CH₂Ph. Uno, dos o tres de R₁, R₂, R₃ y R₄ son H. Uno de R₁, R₂, R₃ y R₄ es H. Dos de R₁, R₂, R₃ y R₄ son H. Tres de R₁, R₂, R₃ y R₄ son H. Al menos uno de R₁, R₂, R₃ y R₄ es H. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral del ácido 3-guanidino-2-aminopropiónico. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral del ácido 4-guanidino-2-aminobutanoico. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la arginina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la homoarginina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N-metilarginina. Al menos uno de R₄, R₅, R₆ y R₇ puede

ser la cadena lateral de la N,N-dimetilarginina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral del ácido 2,3-diaminopropiónico. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral del ácido 2,4-diaminobutanoico, lisina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N-metilisina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N,N-dimetilisina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N-etilisina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N,N,N-trimetilisina, 4-guanidinofenilalanina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la citrulina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la N,N-dimetilisina, β-homoarginina. Al menos uno de R₄, R₅, R₆ y R₇ puede ser la cadena lateral de la 3-(1-piperidinil)alanina.

Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 3-guanidino-2-aminopropiónico. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 4-guanidino-2-aminobutanoico. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la arginina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la homoarginina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-metilarginina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N-dimetilarginina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 2,3-diaminopropiónico. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 2,4-diaminobutanoico, lisina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-metilisina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N-dimetilisina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-etilisina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N,N-trimetilisina, 4-guanidinofenilalanina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la citrulina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N-dimetilisina, β-homoarginina. Al menos dos de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la 3-(1-piperidinil)alanina.

Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 3-guanidino-2-aminopropiónico. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 4-guanidino-2-aminobutanoico. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la arginina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la homoarginina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-metilarginina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N-dimetilarginina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 2,3-diaminopropiónico. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral del ácido 2,4-diaminobutanoico, lisina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-metilisina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la

N,N-dimetilisina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N-etilisina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N,N-trimetilisina, 4-guanidinofenilalanina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la citrulina. Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la N,N-dimetilisina, β-homoarginina.

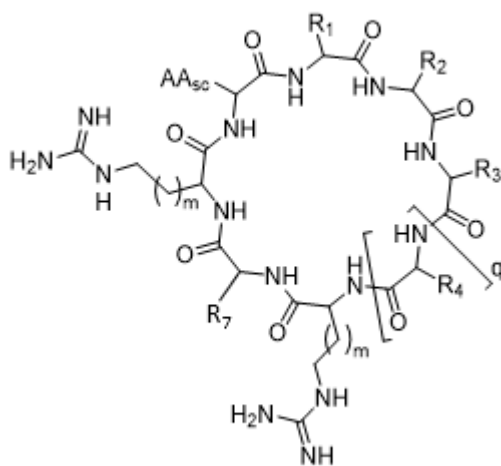
Al menos tres de R₄, R₅, R₆ y R₇ pueden ser la cadena lateral de la 3-(1-piperidinil)alanina.

AA_{SC} puede ser una cadena lateral de un residuo de asparagina, glutamina u homoglutamina. AA_{SC} puede ser una cadena lateral de un residuo de glutamina. El cCPP puede comprender además un enlazador conjugado con el AA_{SC}, por ejemplo, el residuo de asparagina, glutamina u homoglutamina. Por tanto, el cCPP puede comprender además un enlazador conjugado con el residuo de asparagina, glutamina u homoglutamina. El cCPP puede comprender además un enlazador conjugado al residuo de glutamina.

q puede ser 1, 2 o 3. q puede ser 1 o 2. q puede ser 1. q puede ser 2. q puede ser 3. q puede ser 4.

m puede ser 1-3. m puede ser 1 o 2. m puede ser 0. m puede ser 1. m puede ser 2. m puede ser 3.

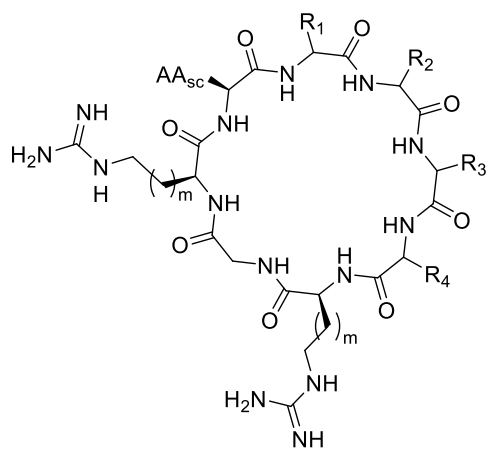
El cCPP de fórmula (A) puede comprender la estructura de fórmula (I)



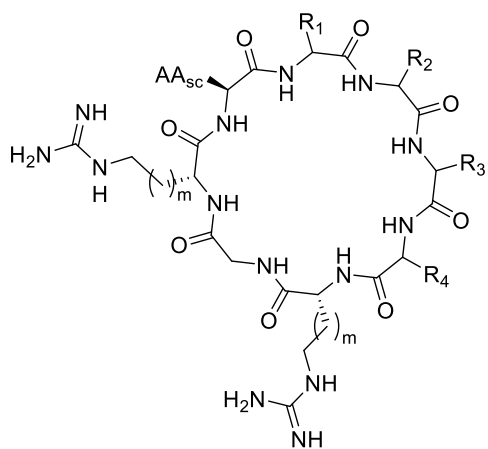
(I)

o una forma protonada de esta, en donde AA_{SC}, R₁, R₂, R₃, R₄, R₇, m y q son como se define en la presente descripción.

El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-a) o fórmula (I-b):



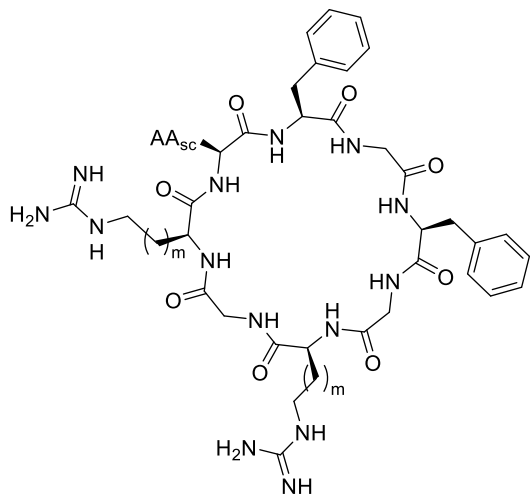
(I-a),



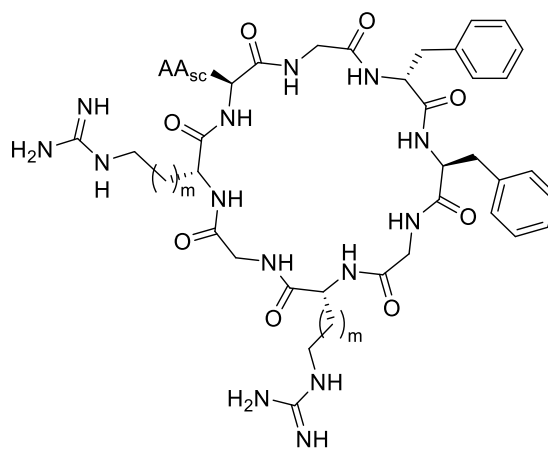
(I-b),

o una forma protonada de estas, en donde AA_{sc}, R₁, R₂, R₃, R₄, y m son como se define en la presente descripción.

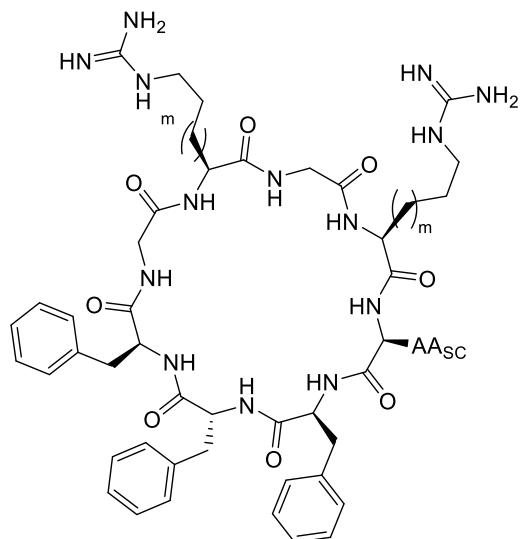
El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-1), (I-2), (I-3) o (I-4):



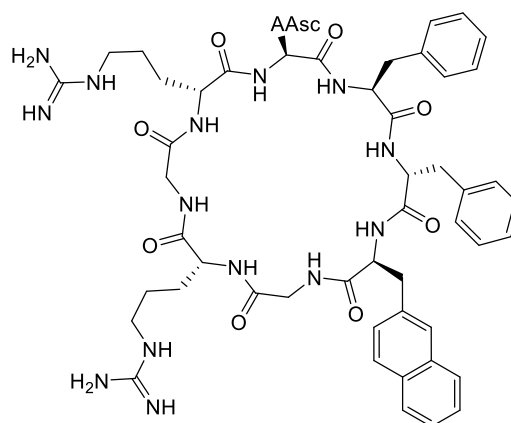
(I-1),



(I-2),



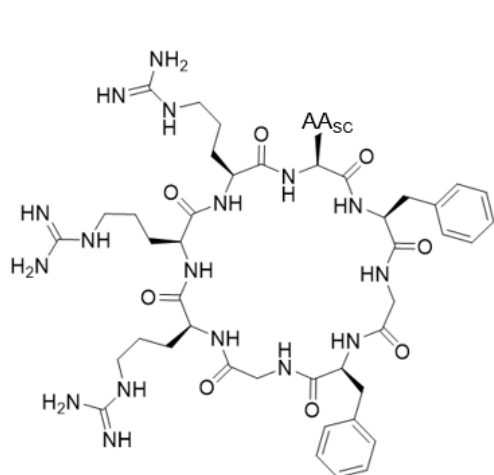
(I-3),



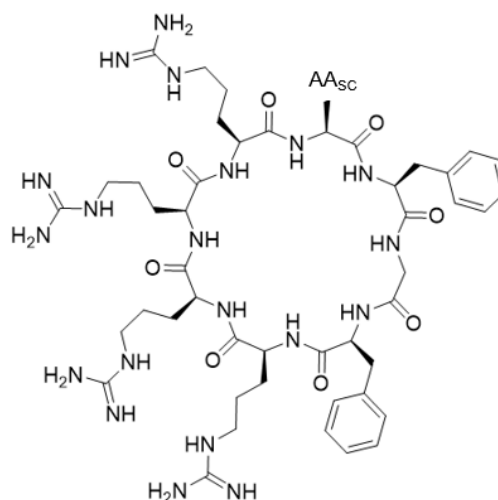
(I-4)

o una forma protonada de estas, en donde AA_{SC} y m son como se define en la presente descripción.

El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-5) o (I-6):



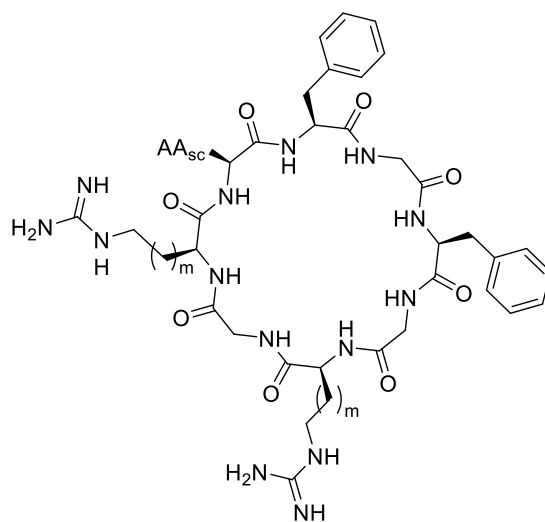
(I-5), o



(I-6) o

una forma protonada de estas, en donde AASC es como se define en la presente descripción.

El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-1):

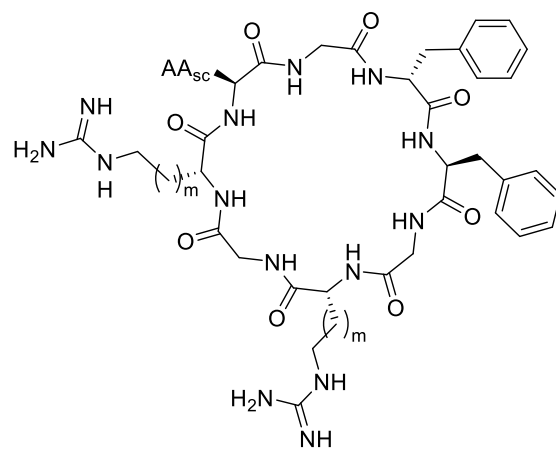


(I-1),

o una forma protonada de esta,

en donde AA_{SC} y m son como se define en la presente descripción.

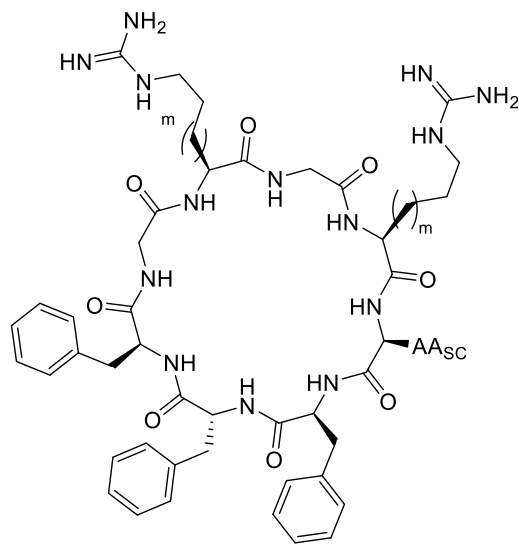
El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-2):



o una forma protonada de esta,

en donde AA_{sc} y m son como se define en la presente descripción.

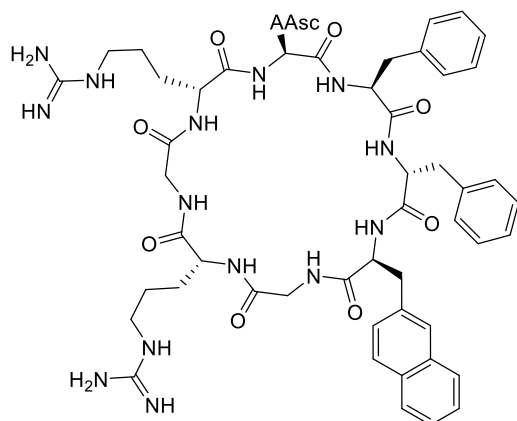
El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-3):



o una forma protonada de esta,

en donde AA_{sc} y m son como se define en la presente descripción.

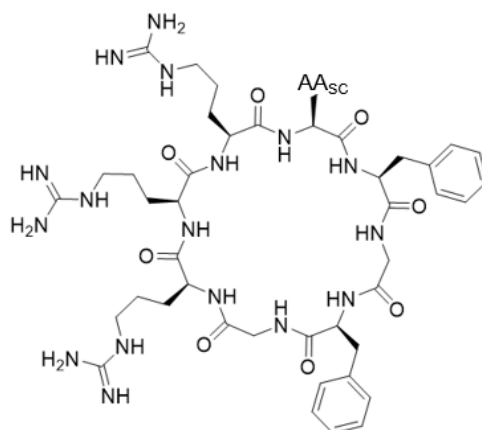
El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-4):



(I-4), o una forma protonada de esta,

en donde AA_{SC} y m son como se define en la presente descripción.

El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-5):

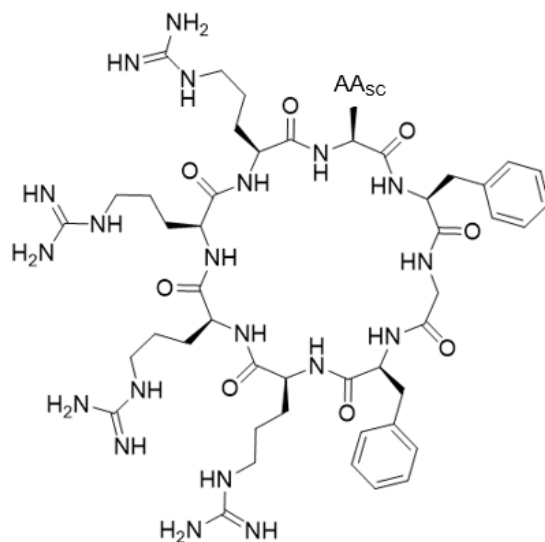


(I-5),

o una forma protonada de esta,

en donde AA_{SC} y m son como se define en la presente descripción.

El cCPP de fórmula (A) puede comprender la estructura de fórmula (I-6):

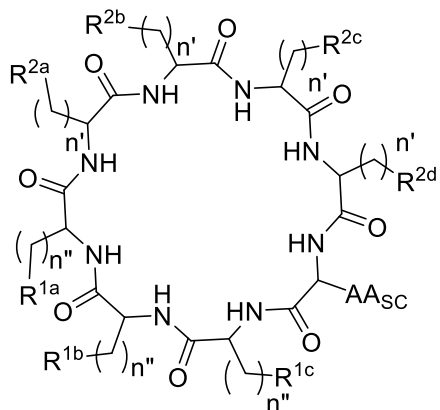


(I-6),

o una forma protonada de esta, en donde AA_{sc} y m son como se define en la presente descripción.

El cCPP puede comprender una de las siguientes secuencias: FGFGRR (SEQ ID NO:68); GfFGrGr (SEQ ID NO:69), FfΦGRGR (SEQ ID NO:70); FfFGRGR (SEQ ID NO:71); o FfΦGrGr (SEQ ID NO:72). El cCPP puede tener una de las siguientes secuencias: FGfΦ (SEQ ID NO:73); GfFGrGrQ (SEQ ID NO:74), FfΦGRGRQ (SEQ ID NO:75); FfFGRGRQ (SEQ ID NO:76); o FfΦGrGrQ (SEQ ID NO:77).

La descripción también se refiere a un cCPP que tiene la estructura de fórmula (II):



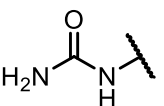
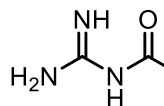
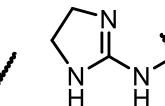
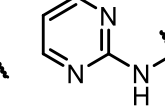
(II)

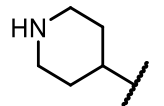
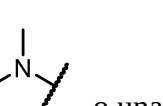
en donde:

AA_{sc} es una cadena lateral de aminoácido;

R^{1a}, R^{1b} y R^{1c} son cada uno independientemente un arilo de 6 a 14 miembros o un heteroarilo de 6 a 14 miembros;

R^{2a}, R^{2b}, R^{2c} y R^{2d} son independientemente una cadena lateral de aminoácido;

al menos uno de R^{2a} , R^{2b} , R^{2c} y R^{2d} es , , , ,

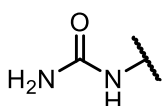
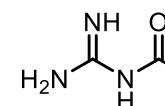
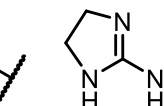
5 , , o una forma protonada de estos;

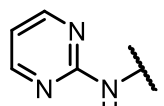
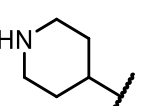
al menos uno de R^{2a} , R^{2b} , R^{2c} y R^{2d} es guanidina o una forma protonada de esta;

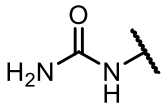
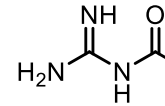
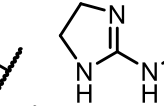
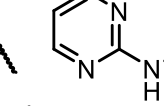
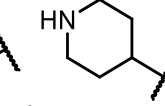
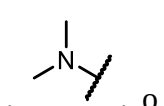
cada n'' es independientemente un número entero 0, 1, 2, 3, 4 o 5;

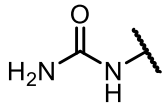
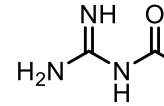
cada n' es independientemente un número entero de 0, 1, 2 o 3; y

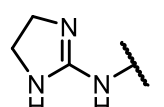
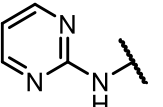
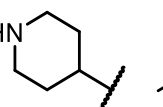
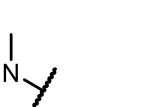
si n' es 0 entonces R^{2a} , R^{2b} , R^{2c} o R^{2d} está ausente.

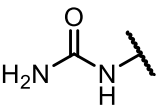
10 Al menos dos de R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser , , ,

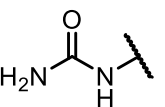
15 , , , o una forma protonada de estos. Dos o tres de R^{2a} , R^{2b} , R^{2c} y R^{2d}

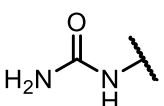
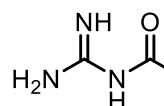
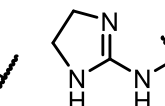
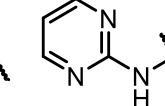
pueden ser , , , , , , o una

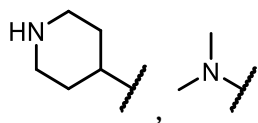
20 forma protonada de estos. Uno de R^{2a} , R^{2b} , R^{2c} y R^{2d} puede ser , ,

, , , , o una forma protonada de estos. Al menos uno de

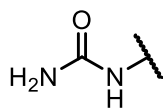
25 R^{2a} , R^{2b} , R^{2c} y R^{2d} puede ser , o una forma protonada de este, y los restantes de R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser guanidina o una forma protonada de esta. Al menos dos de R^{2a} , R^{2b} , R^{2c} y

R^{2d} pueden ser , o una forma protonada de este, y los restantes de R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser guanidina, o una forma protonada de esta.

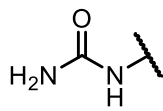
30 Todo los R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser , , , ,



, o una forma protonada de estos. Al menos de R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser

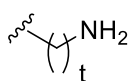
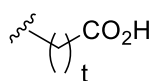


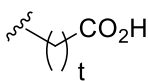
5 , o una forma protonada de este, y los restantes de R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser guaninida o una forma protonada de esta. Al menos dos grupos R^{2a} , R^{2b} , R^{2c} y R^{2d} pueden ser



10 , o una forma protonada de este, y los restantes de R^{2a} , R^{2b} , R^{2c} y R^{2d} son guanidina, o una forma protonada de esta.

Cada uno de R^{2a} , R^{2b} , R^{2c} y R^{2d} puede ser independientemente ácido 2,3-diaminopropiónico, ácido 2,4-diaminobutírico, las cadenas laterales de la ornitina, lisina, metilisina, dimetilisina, trimetilisina, homolisina, serina, homoserina, treonina, alo-treonina, histidina, 1-metilhistidina, ácido 2-aminobutanodioico, ácido aspártico, ácido glutámico o ácido homo-glutámico.

15 AA_{SC} puede ser  o  , en donde t puede ser un número entero de 0 a 5. AA_{SC}

puede ser  , en donde t puede ser un número entero de 0 a 5. t puede ser de 1 a 5. t es 2 o 3. t puede ser 2. t puede ser 3.

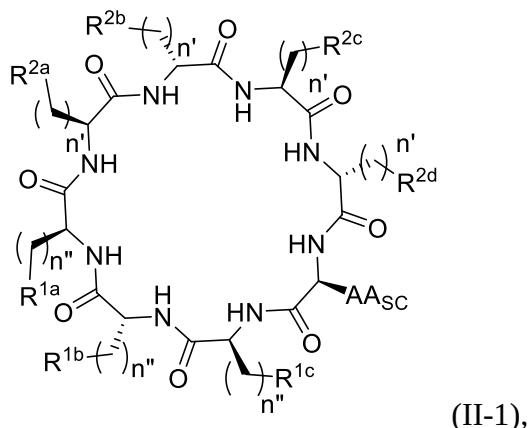
20 R^{1a} , R^{1b} y R^{1c} pueden ser cada uno independientemente un arilo de 6 a 14 miembros. R^{1a} , R^{1b} y R^{1c} pueden ser cada uno independientemente un heteroarilo de 6 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, O o S. R^{1a} , R^{1b} y R^{1c} pueden seleccionarse cada uno independientemente de fenilo, naftilo, antracenilo, piridilo, quinolilo o isoquinolilo. R^{1a} , R^{1b} y R^{1c} pueden seleccionarse cada uno independientemente de fenilo, naftilo o antracenilo. R^{1a} , R^{1b} y R^{1c} pueden ser cada uno independientemente fenilo o naftilo. R^{1a} , R^{1b} y R^{1c} pueden seleccionarse cada uno independientemente de piridilo, quinolilo o isoquinolilo.

25 Cada n' puede ser independientemente 1 o 2. Cada n' puede ser 1. Cada n' puede ser 2. Al menos un n' puede ser 0. Al menos un n' puede ser 1. Al menos un n' puede ser 2. Al menos un n' puede ser 3. Al menos un n' puede ser 4. Al menos un n' puede ser 5.

30 Cada n'' puede ser independientemente un número entero de 1 a 3. Cada n'' puede ser independientemente 2 o 3. Cada n'' puede ser 2. Cada n'' puede ser 3. Al menos un n'' puede ser 0. Al menos un n'' puede ser 1. Al menos un n'' puede ser 2. Al menos un n'' puede ser 3.

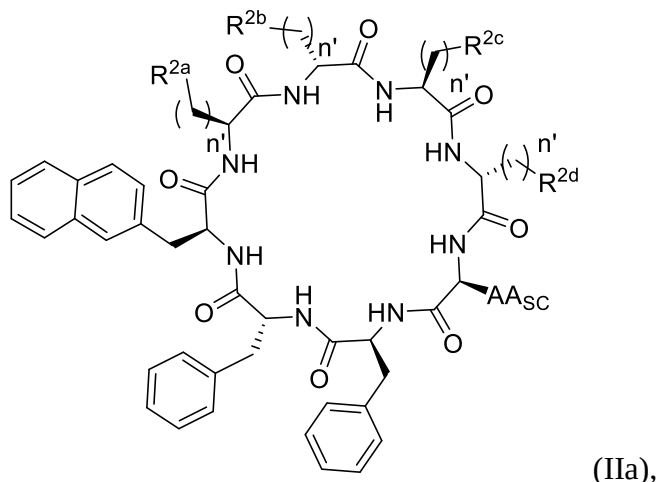
Cada n'' puede ser independientemente 1 o 2 y cada n' puede ser independientemente 2 o 3. Cada n'' puede ser 1 y cada n' puede ser independientemente 2 o 3. Cada n'' puede ser 1 y cada n' puede ser 2. Cada n'' es 1 y cada n' es 3.

El cCPP de fórmula (II) puede tener la estructura de fórmula (II-1):



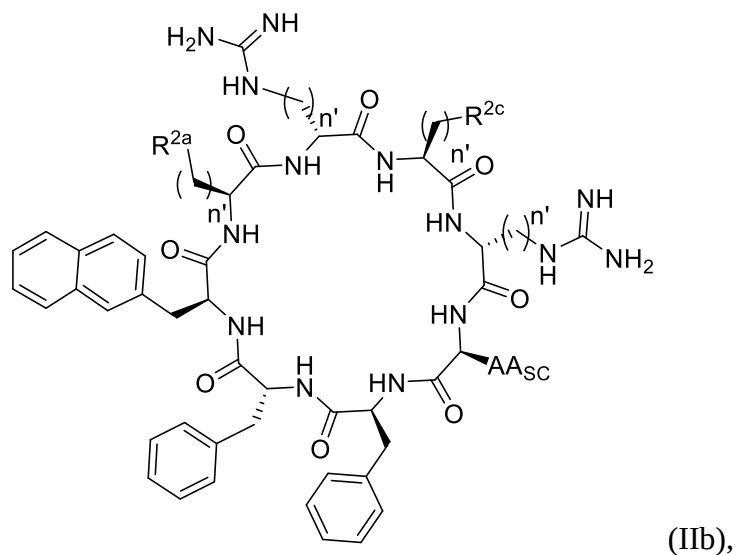
en donde R^{1a} , R^{1b} , R^{1c} , R^{2a} , R^{2b} , R^{2c} , R^{2d} , AA_{Sc} , n' y n'' son como se define en la presente descripción.

El cCPP de fórmula (II) puede tener la estructura de fórmula (IIa):



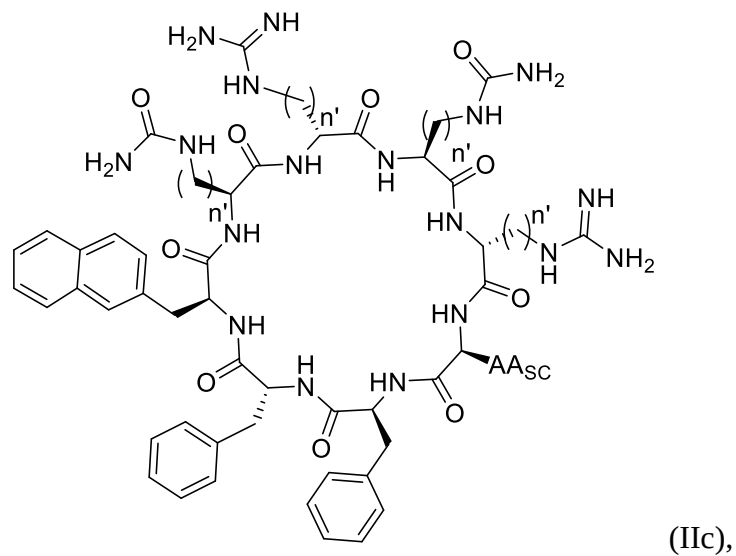
en donde R^{1a} , R^{1b} , R^{1c} , R^{2a} , R^{2b} , R^{2c} , R^{2d} , AA_{Sc} y n' son como se define en la presente descripción.

El cCPP de fórmula (II) puede tener la estructura de fórmula (IIb):



en donde R^{2a} , R^{2b} , AA_{SC} y n' son como se define en la presente descripción.

El cCPP puede tener la estructura de fórmula (IIc):



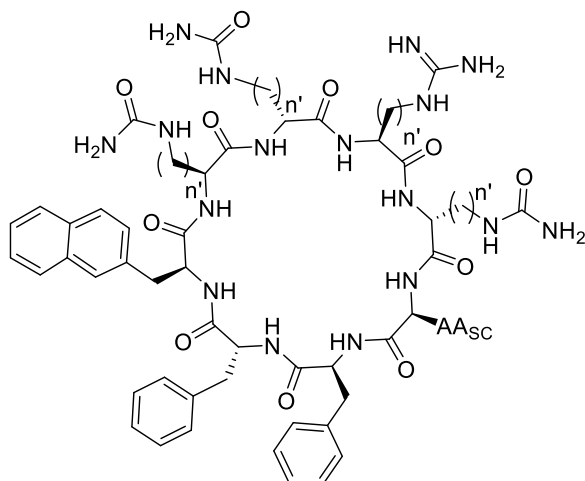
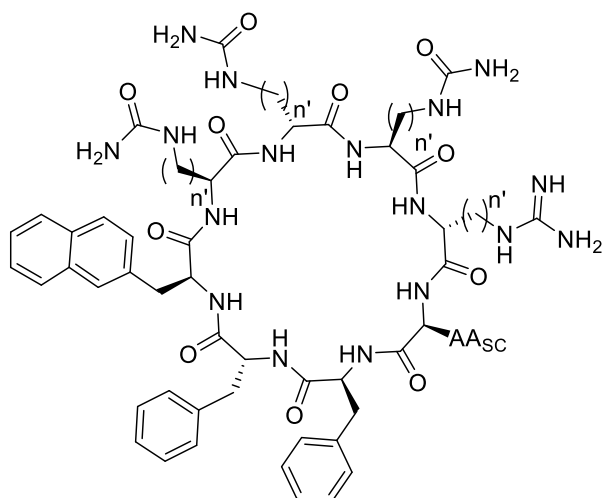
o una forma protonada de esta,

en donde:

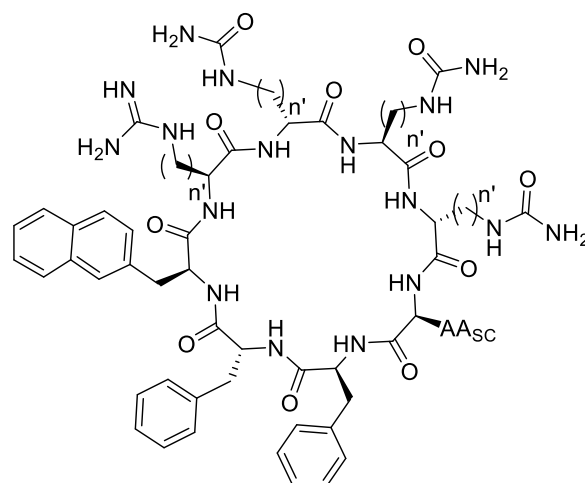
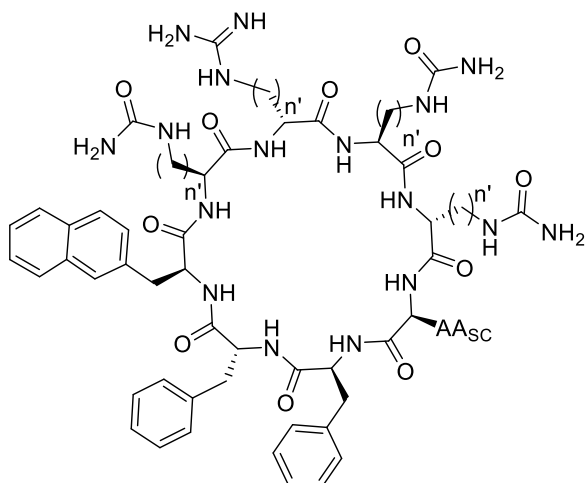
AA_{SC} y n' son como se define en la presente descripción.

El cCPP de fórmula (IIa) tiene una de las siguientes estructuras:

5



10



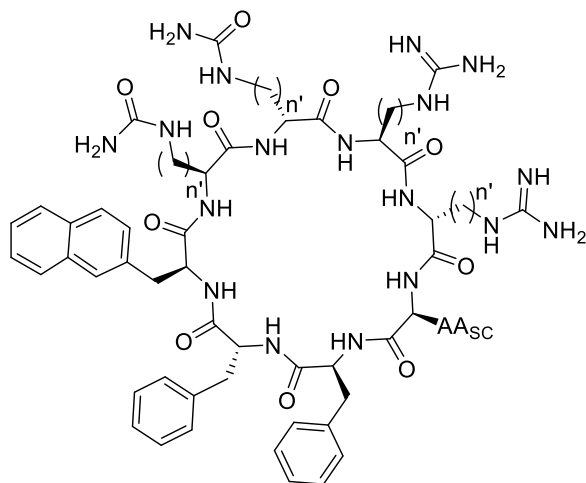
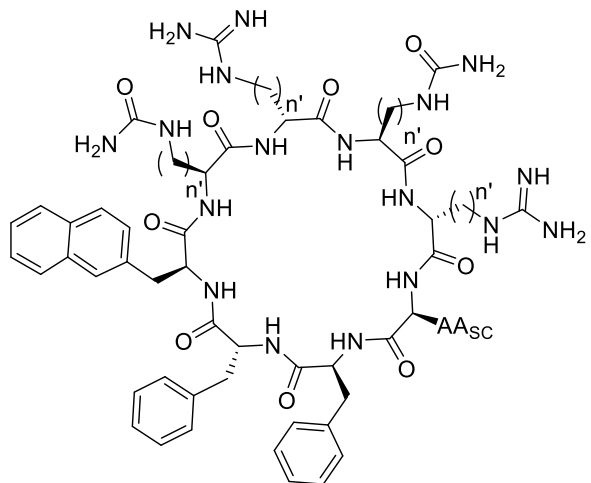
15

20

en donde AA_{SC} y n son como se define en la presente descripción.

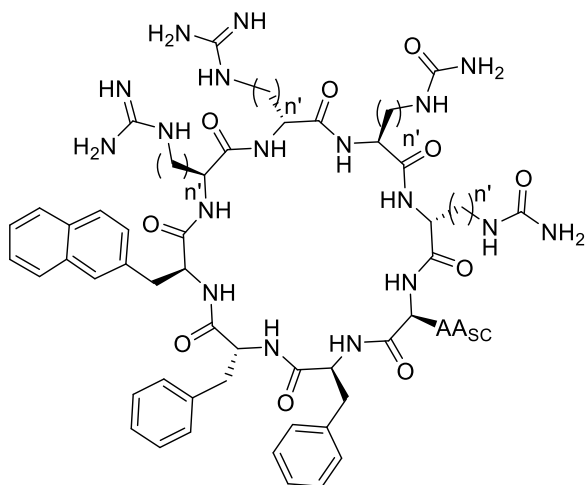
El cCPP de fórmula (IIa) tiene una de las siguientes estructuras:

25

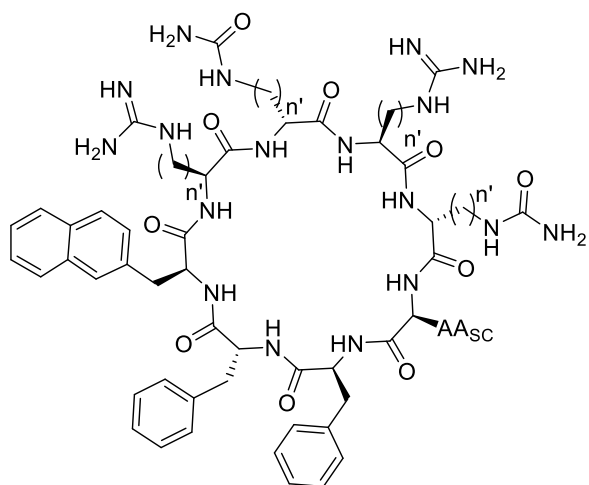


30

5



, 0



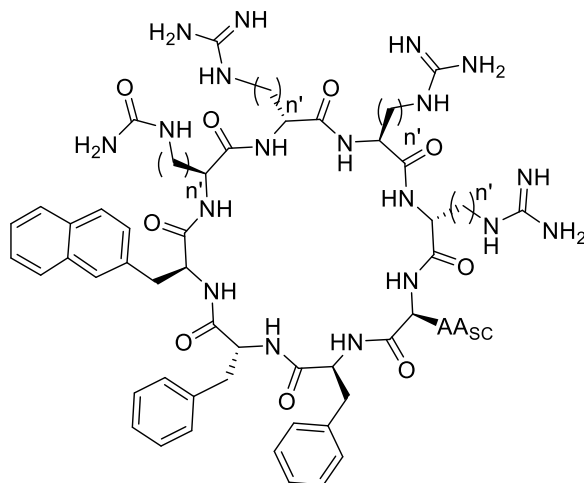
,

10

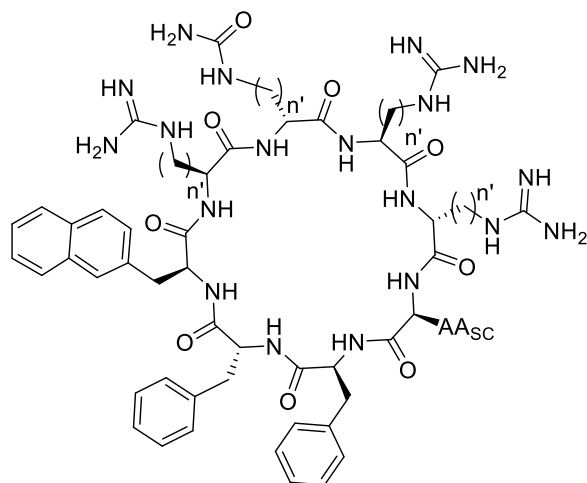
en donde AA_{SC} y n son como se define en la presente descripción

El cCPP de fórmula (IIa) tiene una de las siguientes estructuras:

15



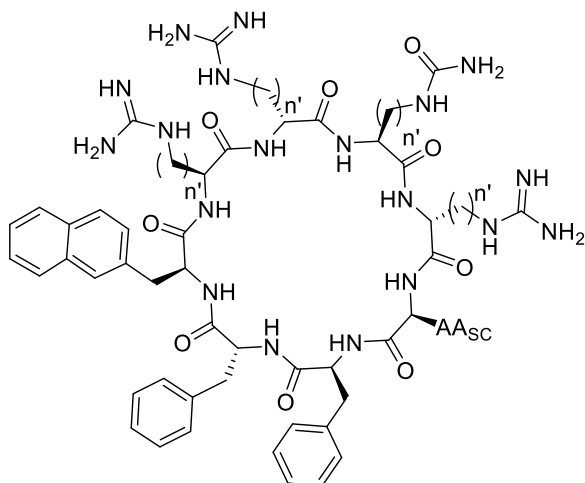
,



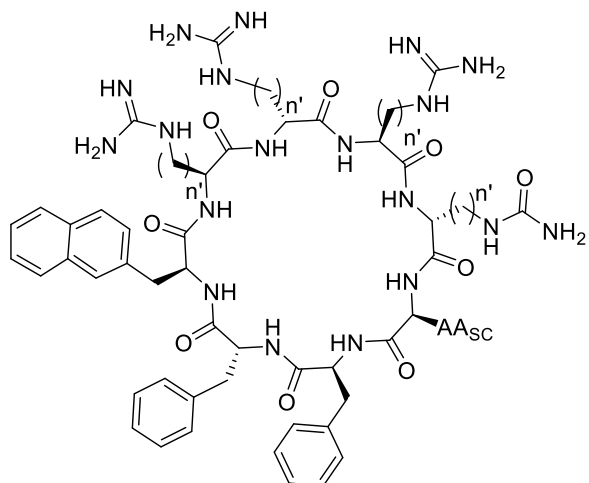
,

20

25



, 0

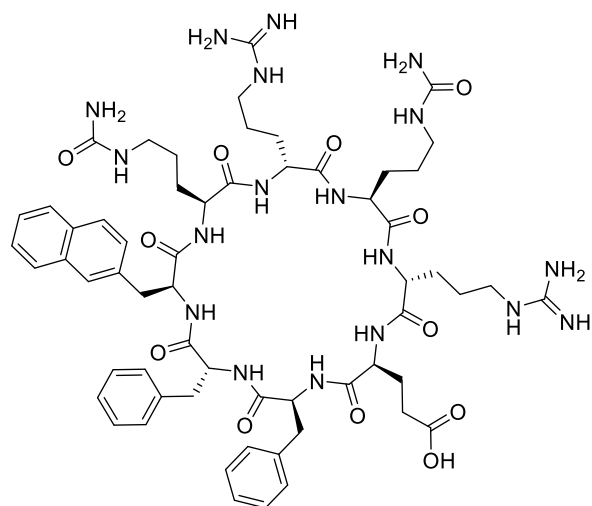


,

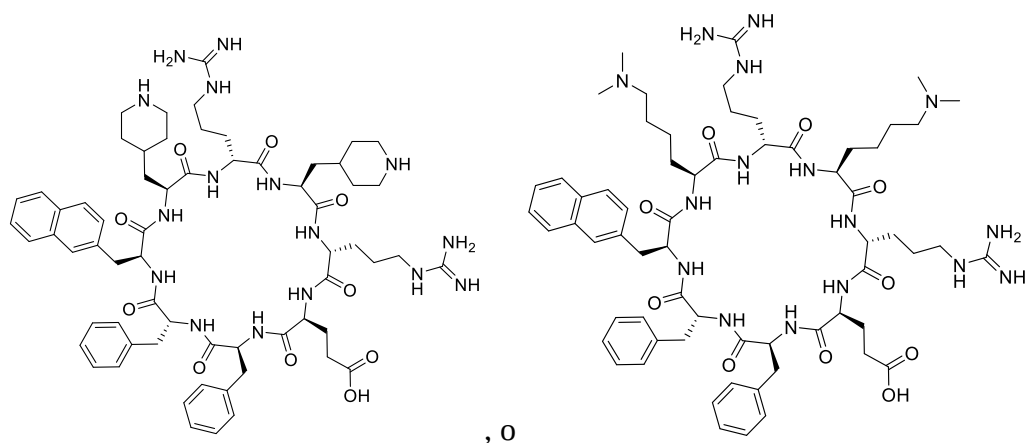
30

en donde AA_{SC} y n son como se define en la presente descripción.

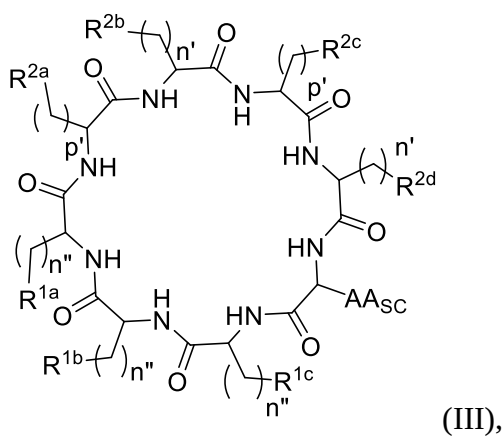
El cCPP de fórmula (II) puede tener la estructura:



El cCPP de fórmula (II) puede tener la estructura:



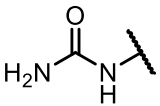
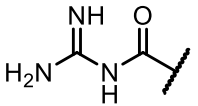
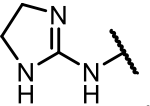
El cCPP puede tener la estructura de fórmula (III):

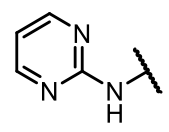
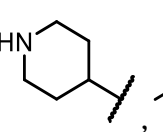
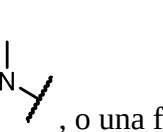


en donde:

AA_{SC} es una cadena lateral de aminoácido;

R^{1a}, R^{1b} y R^{1c} son cada uno independientemente un arilo de 6 a 14 miembros o un heteroarilo de 6 a 14 miembros;

R^{2a} y R^{2c} son cada uno independientemente H, , , ,

, , , o una forma protonada de estos;

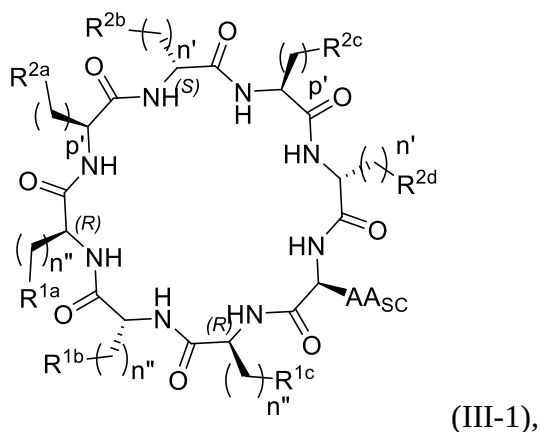
R^{2b} y R^{2d} son cada uno independientemente guanidina o una forma protonada de esta;

cada n'' es independientemente un número entero de 1 a 3;

cada n' es independientemente un número entero de 1 a 5; y

cada p' es independientemente un número entero de 0 a 5.

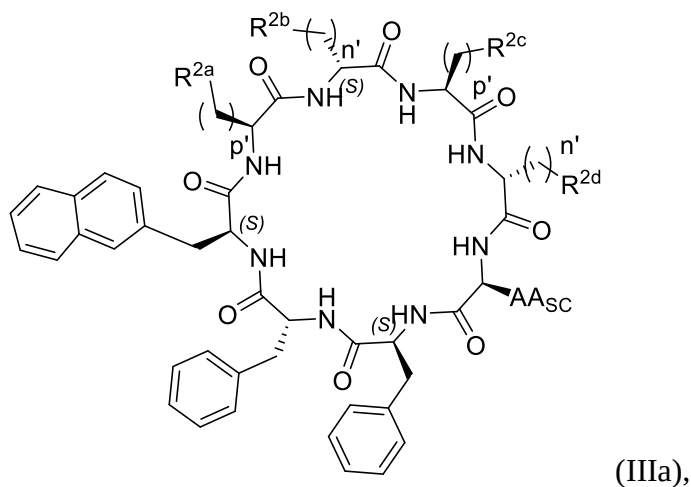
El cCPP de fórmula (III) puede tener la estructura de fórmula (III-1):



en donde:

AA_{SC} , R^{1a} , R^{1b} , R^{1c} , R^{2a} , R^{2c} , R^{2b} , R^{2d} , n' , n'' y p' son como se define en la presente descripción.

El cCPP de fórmula (III) puede tener la estructura de fórmula (IIIa):



en donde:

AA_{SC}, R^{2a}, R^{2c}, R^{2b}, R^{2d} n', n'' y p' son como se define en la presente descripción.

En las fórmulas (III), (III-1) y (IIIa), R^a y R^c pueden ser H. R^a y R^c pueden ser H y R^b y R^d pueden ser cada uno independientemente guanidina o una forma protonada de esta. R^a puede ser H. R^b puede ser H. p' puede ser 0. R^a y R^c pueden ser H y cada p' puede ser 0.

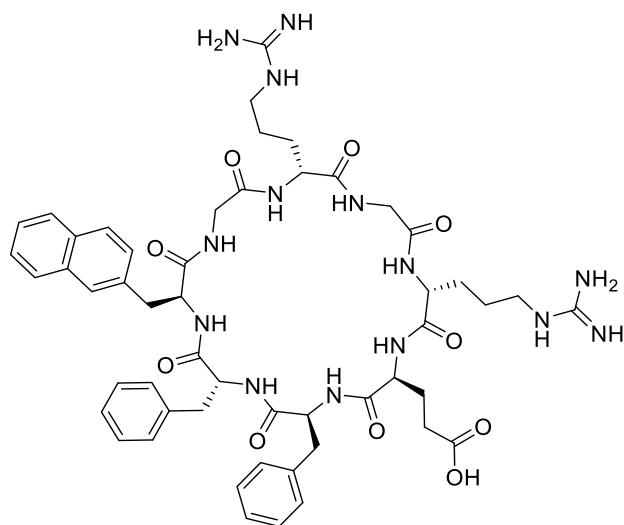
5 En las fórmulas (III), (III-1) y (IIIa), R^a y R^c pueden ser H, R^b y R^d pueden ser cada uno independientemente guanidina o una forma protonada de esta, n'' puede ser 2 o 3, y cada p' puede ser 0.

p' puede ser 0. p' puede ser 1. p' puede ser 2. p' puede ser 3. p' puede ser 4. p' puede ser 5.

El cCPP puede tener la estructura:

10

15



20 El cCPP de fórmula (A) puede seleccionarse de:

25

Secuencia de CPP	SEQ ID NO:
(FfΦRrRrQ)	78
(FfΦCit-r-Cit-rQ)	79
(FfΦGrGrQ)	80
(FfFGRGRQ)	81
(FGFGRGRQ)	82
(GfFGrGrQ)	83
(FGFGRRRQ)	84
(FGFRRRRQ)	85

El cCPP de fórmula (A) puede seleccionarse de:

30

Secuencia de CPP	SEQ ID NO:
FΦRRRRQ	86
fΦRrRrQ	87
FfΦRrRrQ	78
FfΦCit-r-Cit-rQ	79

5

FfΦGrGrQ	80
FfΦRGRGQ	88
FfFGRGRQ	81
FGFGRGRQ	82
GfFGrGrQ	83
FGFGRRRQ	84
FGFRRRRQ	85

En modalidades, el cCPP se selecciona de:

10

15

Secuencias de CPP y SEQ ID NO					
FΦRRRQ	89	RRFRΦRQ	99	FΦRRRRQK	109
FΦRRRC	90	FRRRRΦQ	100	FΦRRRRQC	110
FΦRRRU	91	rRFRΦRQ	101	fΦRrRrRQ	111
RRRΦFQ	92	RRΦFRRQ	102	FΦRRRRRQ	112
RRRRΦF	93	CRRRRFWQ	103	RRRRΦFDΩC	113
FΦRRRR	94	FfΦRrRrQ	104	FΦRRR	114
FφrRrRq	95	FFΦRRRRQ	105	FWRRR	115
FφrRrRQ	96	RFRFRΦRQ	106	RRRΦF	116
FΦRRRRQ	97	URRRRFWQ	107	RRRWF	117
fΦRrRrQ	98	CRRRRFWQ	108		

Donde Φ = L-naftilalanina; φ = D-naftilalanina; Ω = L-norleucina

En modalidades, el cCPP no se selecciona de:

20

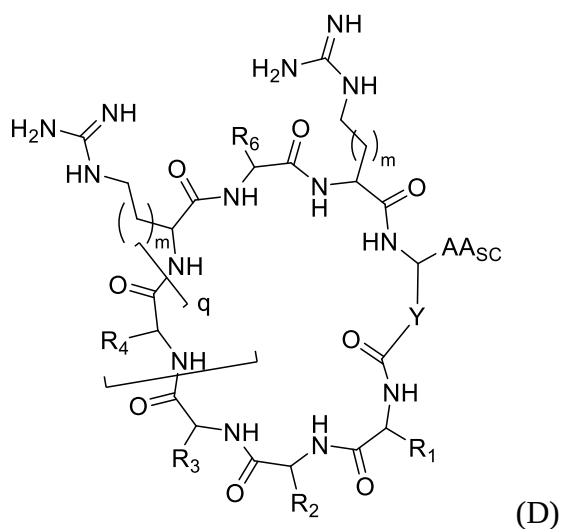
25

Secuencias de CPP y SEQ ID NO					
FΦRRRQ	89	RRFRΦRQ	99	FΦRRRRQK	109
FΦRRRC	90	FRRRRΦQ	100	FΦRRRRQC	110
FΦRRRU	91	rRFRΦRQ	101	fΦRrRrRQ	111
RRRΦFQ	92	RRΦFRRQ	102	FΦRRRRRQ	112
RRRRΦF	93	CRRRRFWQ	103	RRRRΦFDΩC	113
FΦRRRR	94	FfΦRrRrQ	104	FΦRRR	114
FφrRrRq	95	FFΦRRRRQ	105	FWRRR	115
FφrRrRQ	96	RFRFRΦRQ	106	RRRΦF	116
FΦRRRRQ	97	URRRRFWQ	107	RRRWF	117
fΦRrRrQ	98	CRRRRFWQ	108		

Donde Φ = L-naftilalanina; φ = D-naftilalanina; Ω = L-norleucina

El cCPP puede comprender la estructura de fórmula (D)

30

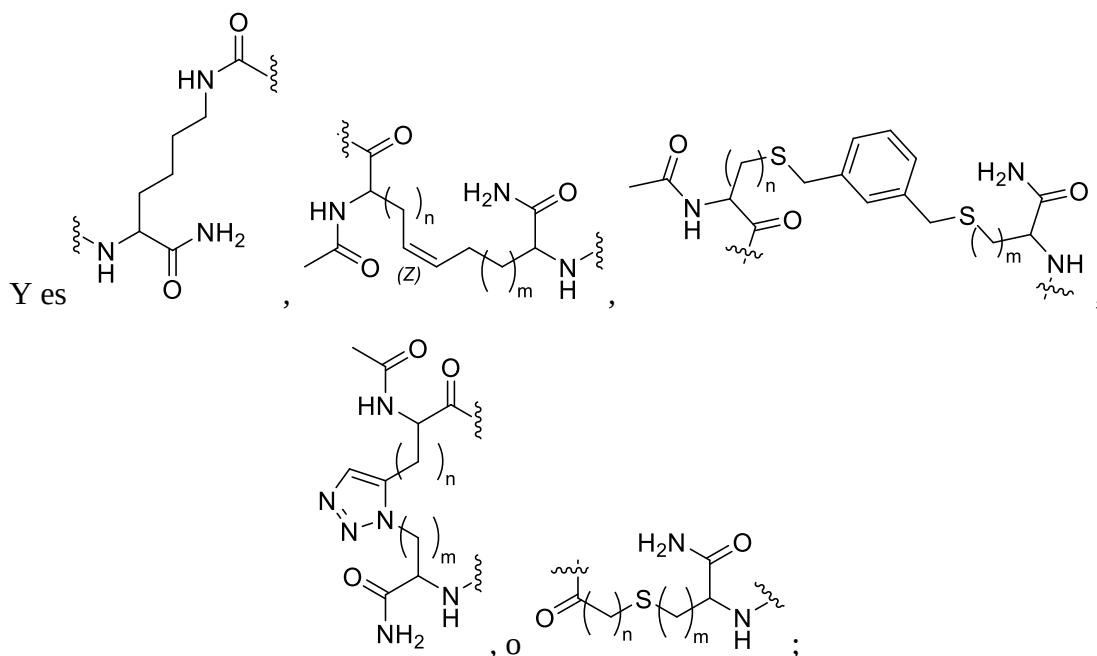


o una forma protonada de esta, en donde:

R_1 , R_2 y R_3 pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

al menos uno de R_1 , R_2 y R_3 es una cadena lateral aromática o heteroaromática de un aminoácido;

R_4 y R_6 son independientemente H o una cadena lateral de aminoácido;



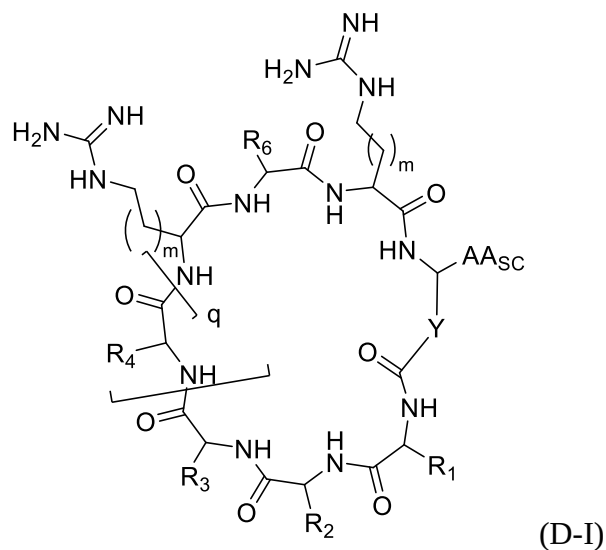
AA_{SC} es una cadena lateral de aminoácido;

q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3, y

cada n es independientemente un número entero 0, 1, 2 o 3.

El cCPP de fórmula (D) puede tener la estructura de fórmula (D-I):



o una forma protonada de esta,

en donde:

R_1 , R_2 y R_3 pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

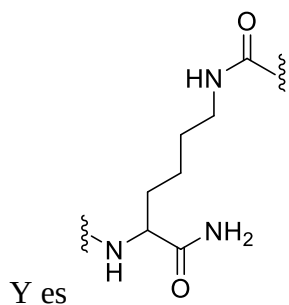
al menos uno de R_1 , R_2 y R_3 es una cadena lateral aromática o heteroaromática de un aminoácido;

R_4 y R_6 son independientemente H o una cadena lateral de aminoácido;

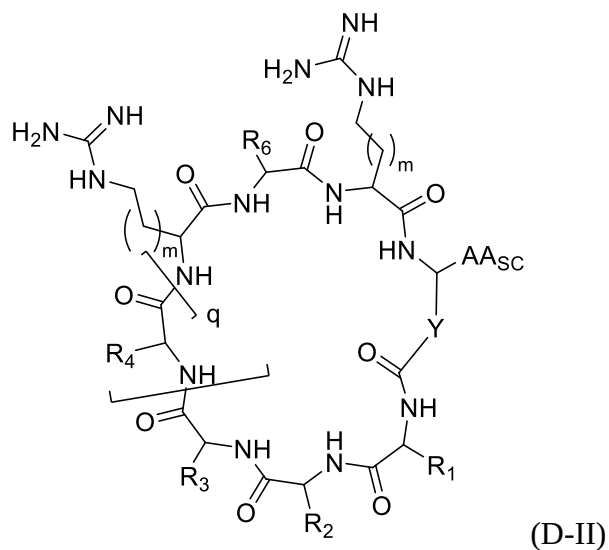
AA_{SC} es una cadena lateral de aminoácido;

q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3, y



El cCPP de fórmula (D) puede tener la estructura de fórmula (D-II):



o una forma protonada de esta,

en donde:

R_1 , R_2 y R_3 pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

al menos uno de R_1 , R_2 y R_3 es una cadena lateral aromática o heteroaromática de un aminoácido;

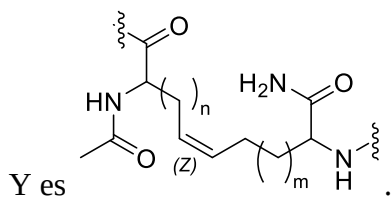
R_4 y R_6 son independientemente H o una cadena lateral de aminoácido;

AA_{SC} es una cadena lateral de aminoácido;

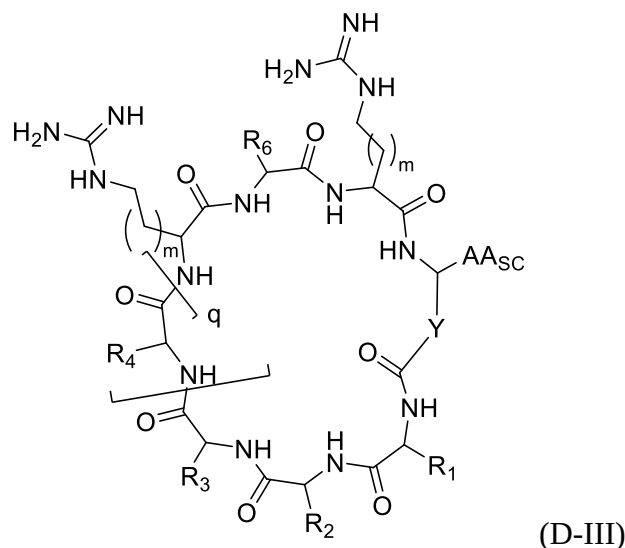
q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3,

cada m es independientemente un número entero 0, 1, 2 o 3, y



El cCPP de fórmula (D) puede tener la estructura de fórmula (D-III):



o una forma protonada de esta,

en donde:

R₁, R₂ y R₃ pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

al menos uno de R₁, R₂ y R₃ es una cadena lateral aromática o heteroaromática de un aminoácido;

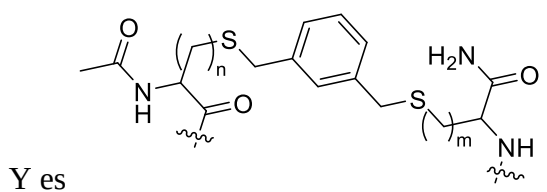
R₄ y R₆ son independientemente H o una cadena lateral de aminoácido;

AA_{SC} es una cadena lateral de aminoácido;

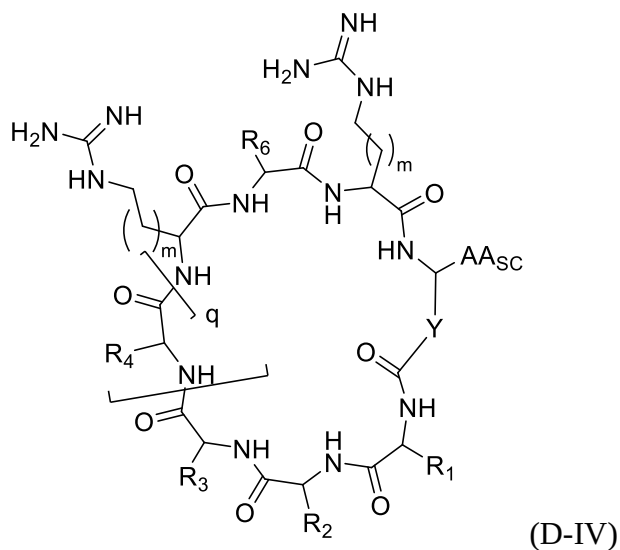
q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3,

cada n es independientemente un número entero 0, 1, 2 o 3, y



El cCPP de fórmula (D) puede tener la estructura de fórmula (D-IV):



o una forma protonada de esta,

en donde:

R₁, R₂ y R₃ pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

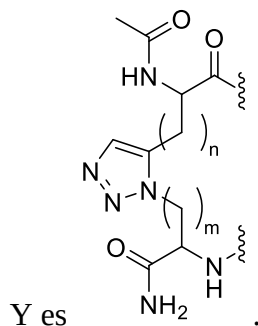
al menos uno de R₁, R₂ y R₃ es una cadena lateral aromática o heteroaromática de un aminoácido;

R₄ y R₆ son independientemente H o una cadena lateral de aminoácido;

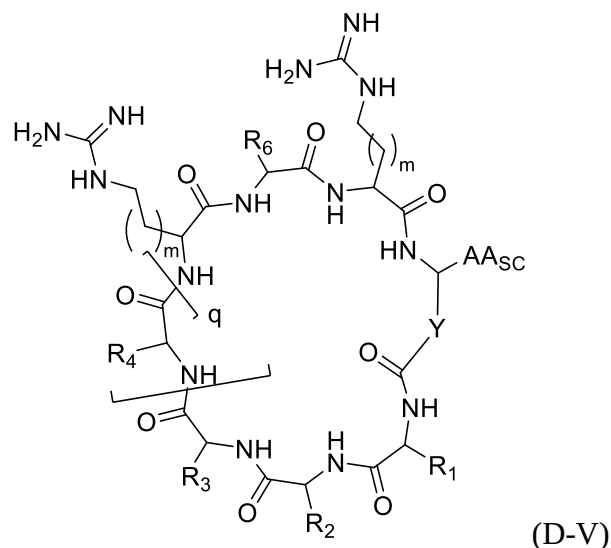
AA_{SC} es una cadena lateral de aminoácido;

q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3, y



El cCPP de fórmula (D) puede tener la estructura de fórmula (D-V):



o una forma protonada de esta,

en donde:

R_1 , R_2 y R_3 pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

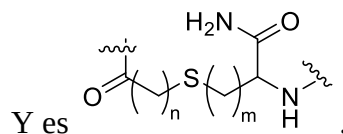
al menos uno de R_1 , R_2 y R_3 es una cadena lateral aromática o heteroaromática de un aminoácido;

R_4 y R_6 son independientemente H o una cadena lateral de aminoácido;

AA_{SC} es una cadena lateral de aminoácido;

q es 1, 2, 3 o 4;

cada m es independientemente un número entero 0, 1, 2 o 3, y



El AA_{SC} puede conjugarse con un enlazador.

Enlazador

El cCPP de la descripción puede conjugarse con un enlazador. El enlazador puede enlazar una carga al cCPP. El enlazador puede unirse a la cadena lateral de un aminoácido del cCPP, y la carga puede unirse a una posición adecuada en el enlazador.

El enlazador puede ser cualquier resto apropiado que pueda conjugar un cCPP a uno o más restos adicionales, por ejemplo, un péptido exocíclico (EP) y/o una carga. Antes de la conjugación con el cCPP y uno o más restos adicionales, el enlazador tiene dos o más grupos funcionales, cada uno de los cuales es independientemente capaz de formar un enlace covalente al cCPP y uno o más restos adicionales. Si la carga es un oligonucleótido, el enlazador puede unirse covalentemente al extremo

5' de la carga o al extremo 3' de la carga. El enlazador puede unirse covalentemente al extremo 5' de la carga. El enlazador puede unirse covalentemente al extremo 3' de la carga. Si la carga es un péptido, el enlazador puede unirse covalentemente al extremo N-terminal o al extremo C-terminal de la carga. El enlazador puede unirse covalentemente a la cadena principal del oligonucleótido o carga peptídica. El enlazador puede ser cualquier resto apropiado que conjuga un cCPP descrito en la presente descripción a una carga tal como un oligonucleótido, péptido o molécula pequeña.

El enlazador puede comprender un enlazador hidrocarbonado.

El enlazador puede comprender un sitio de escisión. El sitio de escisión puede ser un disulfuro o un sitio de escisión por caspasa (por ejemplo, Val-Cit-PABC).

10 El enlazador puede comprender: (i) uno o más D o L aminoácidos, cada uno de los cuales se sustituye opcionalmente; (ii) alquilenos opcionalmente sustituidos; (iii) alquilenos opcionalmente sustituidos; (iv) alquilenos opcionalmente sustituidos; (v) carbociclilos opcionalmente sustituidos; (vi) heterociclilos opcionalmente sustituidos; (vii) una o más subunidades $-(R^1-J-R^2)z''-$, en donde cada uno de R^1 y R^2 , en cada caso, se seleccionan independientemente de alquilenos, alquilenos, alquilenos, carbociclilo y heterociclilo, cada J es independientemente C, NR^3 , $-NR^3C(O)-$, S y O, en donde R^3 se selecciona independientemente de H, alquilo, alqueno, alqueno, carbociclilo y heterociclilo, cada uno de los cuales se sustituye opcionalmente, y z'' es un número entero de 1 a 50; (viii) $-(R^1-J)z''-$ o $-(J-R^1)z''-$, en donde cada uno de R^1 , en cada caso, es independientemente alquilenos, alquilenos, alquilenos, carbociclilo o heterociclilo, cada J es independientemente C, NR^3 , $-NR^3C(O)-$, S u O, en donde R^3 es H, alquilo, alqueno, alqueno, carbociclilo o heterociclilo, cada uno de los cuales se sustituye opcionalmente, y z'' es un número entero de 1 a 50; o (ix) el enlazador puede comprender uno o más de (i) hasta (x).

El enlazador puede comprender uno o más D o L aminoácidos y/o $-(R^1-J-R^2)z''-$, en donde cada uno de R^1 y R^2 , en cada caso, son independientemente alquilenos, cada J es independientemente C, NR^3 , $-NR^3C(O)-$, S y O, en donde R^4 se selecciona independientemente de H y alquilo, y z'' es un número entero de 1 a 50; o sus combinaciones.

El enlazador puede comprender un $-(OCH_2CH_2)_{z'}-$ (por ejemplo, como espaciador), en donde z' es un número entero de 1 a 23, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 o 23. “ $-(OCH_2CH_2)_{z'}$ ” también puede denominarse polietilenglicol (PEG).

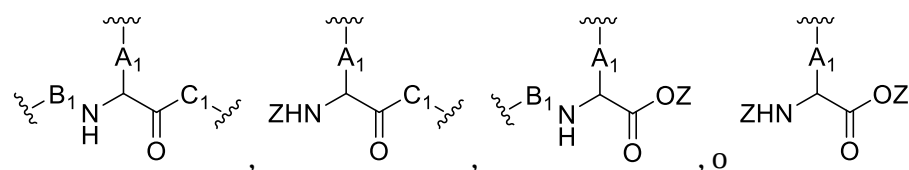
30 El enlazador puede comprender uno o más aminoácidos. El enlazador puede comprender un péptido. El enlazador puede comprender un $-(OCH_2CH_2)_{z'}-$, en donde z' es un número entero de 1 a 23, y un péptido. El péptido puede comprender de 2 a 10 aminoácidos. El enlazador puede comprender

además un grupo funcional (FG) capaz de reaccionar a través de química clic. El FG puede ser una azida o alquino, y se forma una triazol cuando la carga se conjuga al enlazador.

El enlazador puede comprender (i) un residuo de β alanina y un residuo de lisina; (ii) $-(J-R^1)z''$; o (iii) sus combinaciones. Cada R^1 puede ser independientemente alquileo, alquenileno, alquinileno, carbociclilo o heterociclilo, cada J es independientemente C, NR^3 , $-NR^3C(O)-$, S u O, en donde R^3 es H, alquilo, alquenilo, alquinilo, carbociclilo o heterociclilo, cada uno de los cuales se sustituye opcionalmente, y z'' puede ser un número entero de 1 a 50. Cada R^1 puede ser alquileo y cada J puede ser O.

El enlazador puede comprender (i) residuos de β -alanina, glicina, lisina, ácido 4-aminobutírico, ácido 5-aminopentanoico, ácido 6-aminohexanoico o sus combinaciones; y (ii) $-(R^1-J)z''$ o $-(J-R^1)z''$. Cada R^1 puede ser independientemente alquileo, alquenileno, alquinileno, carbociclilo o heterociclilo, cada J es independientemente C, NR^3 , $-NR^3C(O)-$, S u O, en donde R^3 es H, alquilo, alquenilo, alquinilo, carbociclilo o heterociclilo, cada uno de los cuales se sustituye opcionalmente, y z'' puede ser un número entero de 1 a 50. Cada R^1 puede ser alquileo y cada J puede ser O. El enlazador puede comprender glicina, beta-alanina, ácido 4-aminobutírico, ácido 5-aminopentanoico, ácido 6-aminohexanoico o sus combinaciones.

El enlazador puede ser un enlazador trivalente. El enlazador puede tener la estructura:



, en donde A_1 , B_1 y C_1 , puede ser independientemente un enlazador hidrocarbonado (por ejemplo, $NRH-(CH_2)_n-COOH$), un enlazador de PEG (por ejemplo, $NRH-(CH_2O)_n-COOH$, en donde R es H, metil o etilo) o uno o más residuos de aminoácidos, y Z es independientemente un grupo de protección. El enlazador también puede incorporar un sitio de escisión, lo que incluye un sitio de escisión de disulfuro [$NH_2-(CH_2O)_n-S-S-(CH_2O)_n-COOH$], o un sitio de escisión por caspasa (Val-Cit-PABC).

El hidrocarburo puede ser un residuo de glicina o beta-alanina.

El enlazador puede ser bivalente y enlazar el cCPP a una carga. El enlazador puede ser bivalente y enlazar el cCPP a un péptido exocíclico (EP).

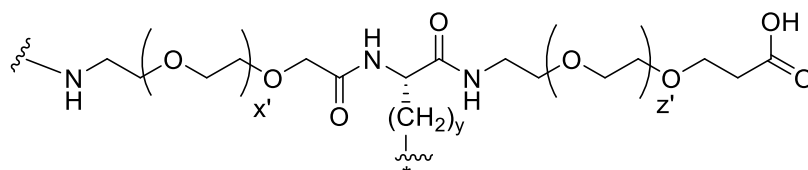
El enlazador puede ser trivalente y enlazar el cCPP a una carga y a un EP.

El enlazador puede ser un alquileo C_1-C_{50} bivalente o trivalente, en donde 1-25 grupos metileno se reemplazan opcional e independientemente por $-N(H)-$, $-N(\text{alquilo } C_1-C_4)-$, $-N(\text{cicloalquilo})-$, $-O-$, $-$

en donde: x' es un número entero de 1-23; y es un número entero de 1-5; z' es un número entero de 1-23; * es el punto de unión al AA_{SC}, y AA_{SC} es una cadena lateral de un residuo de aminoácido del cCPP; y M es un grupo de unión definido en la presente descripción.

El enlazador puede tener la estructura:

5



10 en donde: x' es un número entero de 1-23; y es un número entero de 1-5; y z' es un número entero de 1-23; * es el punto de unión al AA_{SC}, y AA_{SC} es una cadena lateral de un residuo de aminoácido del cCPP.

x puede ser un número entero de 1-10, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, lo que incluye todos los intervalos y subintervalos intermedios.

15 x' puede ser un número entero de 1-23, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 o 23, lo que incluye todos los intervalos y subintervalos intermedios. x' puede ser un número entero de 5-15. x' puede ser un número entero de 9-13. x' puede ser un número entero de 1-5. x' puede ser 1.

20 y puede ser un número entero de 1-5, por ejemplo, 1, 2, 3, 4 o 5, lo que incluye todos los intervalos y subintervalos intermedios. y puede ser un número entero de 2-5. y puede ser un número entero de 3-5. y puede ser 3 o 4. y puede ser 4 o 5. y puede ser 3. y puede ser 4. y puede ser 5.

z puede ser un número entero de 1-10, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, lo que incluye todos los intervalos y subintervalos intermedios.

25 z' puede ser un número entero de 1-23, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 o 23, lo que incluye todos los intervalos y subintervalos intermedios. z' puede ser un número entero de 5-15. z' puede ser un número entero de 9-13. z' puede ser 11.

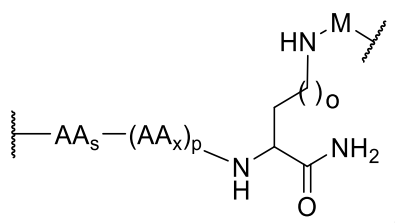
30 Como se analizó anteriormente, el enlazador o M (en donde M es parte del enlazador) puede unirse covalentemente a la carga en cualquier ubicación adecuada en la carga. El enlazador o M (en donde M es parte del enlazador) puede unirse covalentemente al extremo 3' de la carga de oligonucleótidos o al extremo 5' de una carga de oligonucleótidos. El enlazador o M (en donde M es parte del enlazador) puede unirse covalentemente al extremo N-terminal o al extremo C-terminal de una carga peptídica. El enlazador o M (en donde M es parte del enlazador) puede unirse covalentemente a la cadena principal de una carga de oligonucleótidos o péptidos.

El enlazador puede unirse a la cadena lateral del ácido aspártico, ácido glutámico, glutamina, asparagina o lisina, o a una cadena lateral modificada de glutamina o asparagina (por ejemplo, una cadena lateral reducida que tiene un grupo amino), en el cCPP. El enlazador puede unirse a la cadena lateral de la lisina en el cCPP.

- 5 El enlazador puede unirse a la cadena lateral del ácido aspártico, ácido glutámico, glutamina, asparagina o lisina, o a una cadena lateral modificada de glutamina o asparagina (por ejemplo, una cadena lateral reducida que tiene un grupo amino), en una carga peptídica. El enlazador puede unirse a la cadena lateral de la lisina en la carga peptídica.

El enlazador puede tener una estructura:

10



en donde

15

M es un grupo que conjuga L a una carga, por ejemplo, un oligonucleótido;

AA_s es una cadena lateral o extremo terminal de un aminoácido en el cCPP;

cada AA_x es independientemente un residuo de aminoácido;

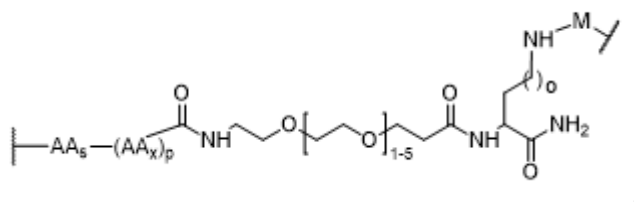
o es un número entero de 0 a 10; y

p es un número entero de 0 a 5.

20

El enlazador puede tener una estructura:

25



en donde

M es un grupo que conjuga L a una carga, por ejemplo, un oligonucleótido;

AA_s es una cadena lateral o extremo terminal de un aminoácido en el cCPP;

cada AA_x es independientemente un residuo de aminoácido;

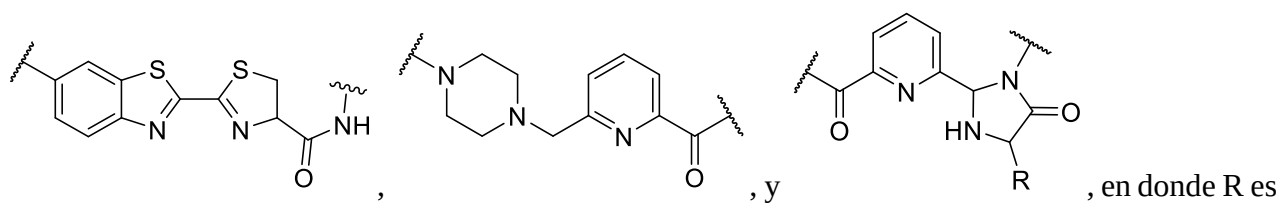
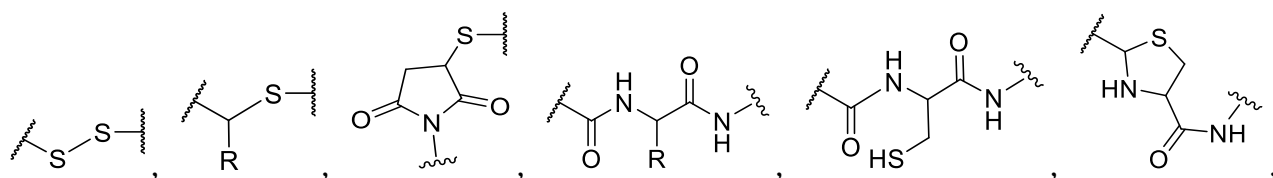
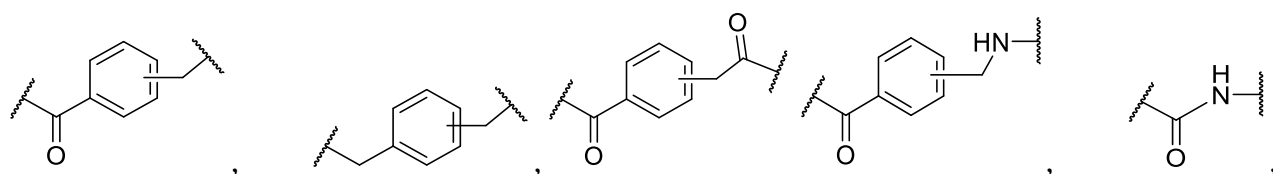
30

o es un número entero de 0 a 10; y

p es un número entero de 0 a 5.

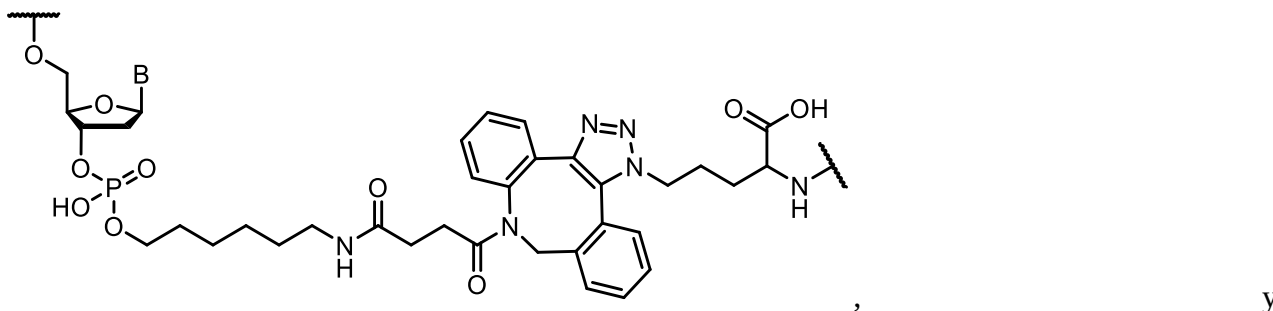
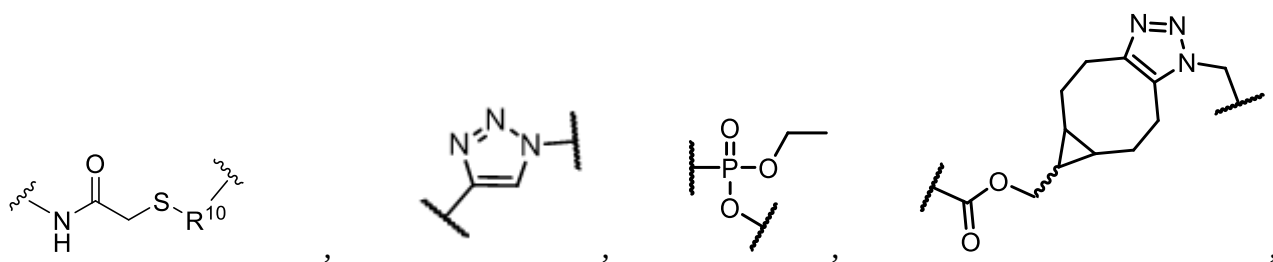
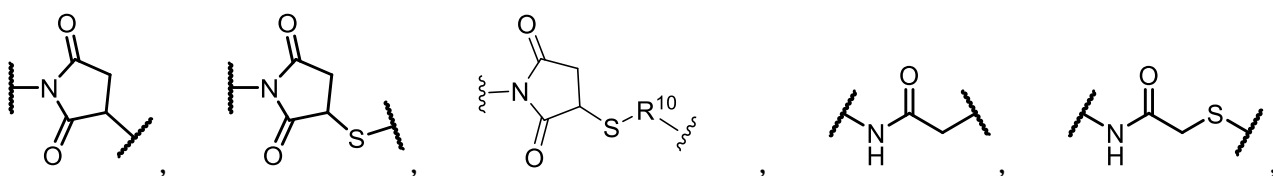
M puede comprender un alquilenilo, alquenilenilo, alquinilenilo, carbociclilo o heterociclilo, cada uno

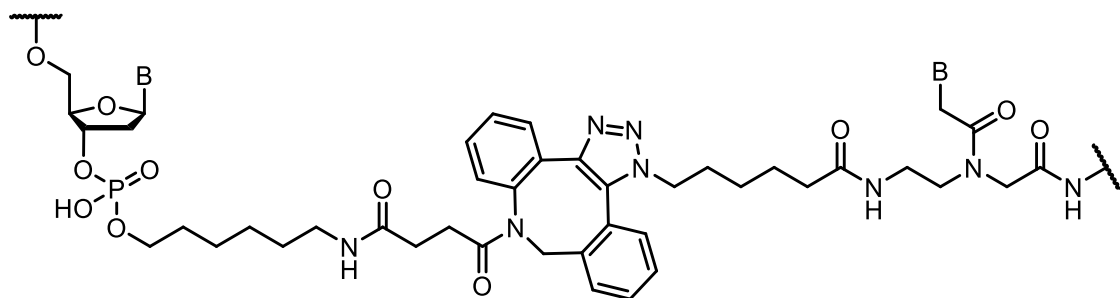
de los cuales se sustituye opcionalmente. M puede seleccionarse de:



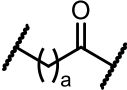
alquilo, alquenilo, alquinilo, carbociclilo o heterociclilo.

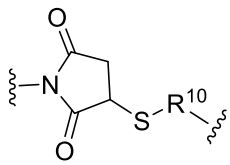
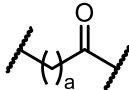
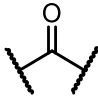
M puede seleccionarse de:

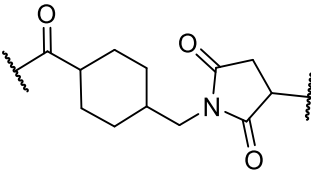




, en donde:

R^{10} es alquileo, cicloalquilo o , en donde a es de 0 a 10.

M puede ser , R^{10} puede ser , y a es de 0 a 10. M puede ser .

M puede ser un reticulador heterobifuncional, por ejemplo, , que se describe en Williams y otros *Curr. Protoc Nucleic Acid Chem.* 2010, 42, 4.41.1-4.41.20, incorporado en la presente descripción como referencia en su totalidad.

M puede ser -C(O)-.

AA_s puede ser una cadena lateral o extremo terminal de un aminoácido en el cCPP. Ejemplos no limitantes de AA_s incluyen ácido aspártico, ácido glutámico, glutamina, asparagina o lisina, o una cadena lateral modificada de la glutamina o asparagina (por ejemplo, una cadena lateral reducida que tiene un grupo amino). AA_x puede ser un AA_{SC} como se define en la presente descripción.

Cada AA_x es independientemente un aminoácido natural o no natural. Uno o más AA_x pueden ser un aminoácido natural. Uno o más AA_x pueden ser un aminoácido no natural. Uno o más AA_x pueden ser un β -aminoácido. El β -aminoácido puede ser β -alanina.

o puede ser un número entero de 0 a 10, por ejemplo, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10. o puede ser 0, 1, 2 o 3. o puede ser 0. o puede ser 1. o puede ser 2. o puede ser 3.

p puede ser de 0 a 5, por ejemplo, 0, 1, 2, 3, 4 o 5. p puede ser 0. p puede ser 1. p puede ser 2. p puede ser 3. p puede ser 4. p puede ser 5.

El enlazador puede tener la estructura:



El enlazador puede tener la estructura:



15

z'' puede ser un número entero de 1 a 50, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 y 50, lo que incluye todos los intervalos y valores intermedios. z'' puede ser un número entero de 5-20. z'' puede ser un número entero de 10-15.

El enlazador puede tener la estructura:

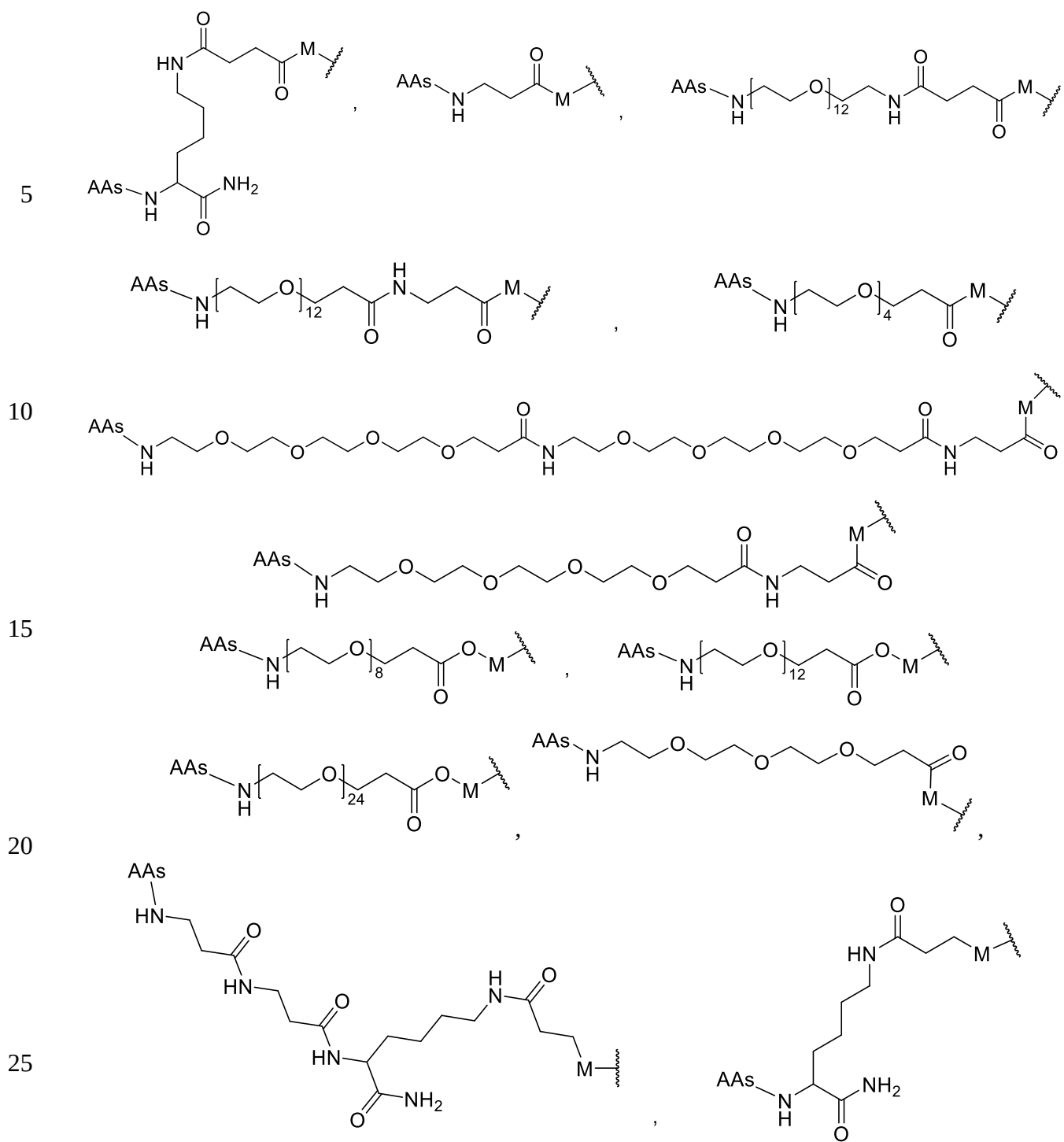


en donde:

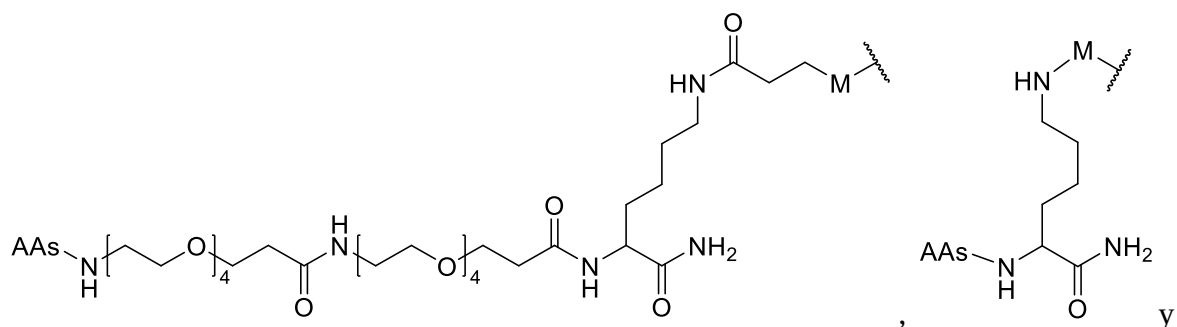
M, AA_s y o son como se define en la presente descripción.

Otros ejemplos no limitantes de enlazadores adecuados incluyen:

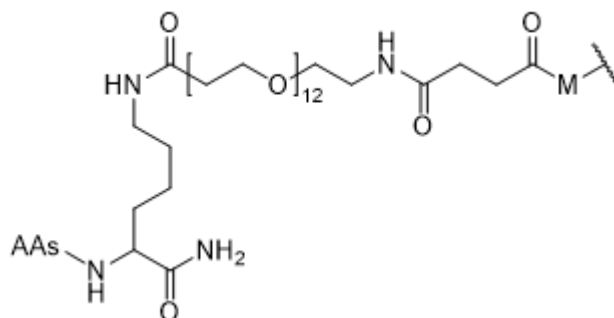




5



10

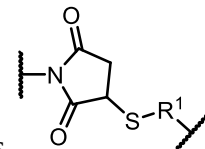


en donde M y AAs son como se define en la presente descripción.

15

En la presente descripción se proporciona un compuesto que comprende un cCPP y un AC que es complementario a un objetivo en una secuencia pre-ARNm que comprende además L, en donde el

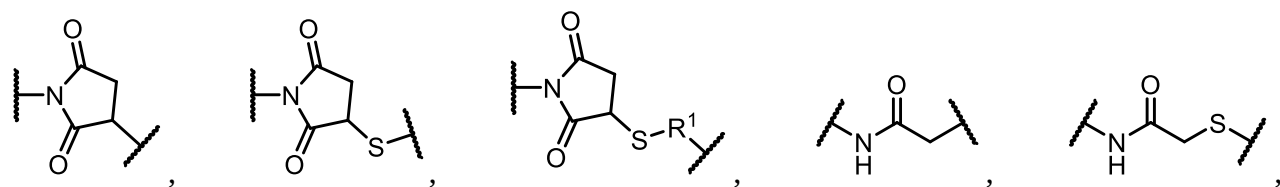
enlazador se conjuga con el AC a través de un grupo de unión (M), en donde M es



20

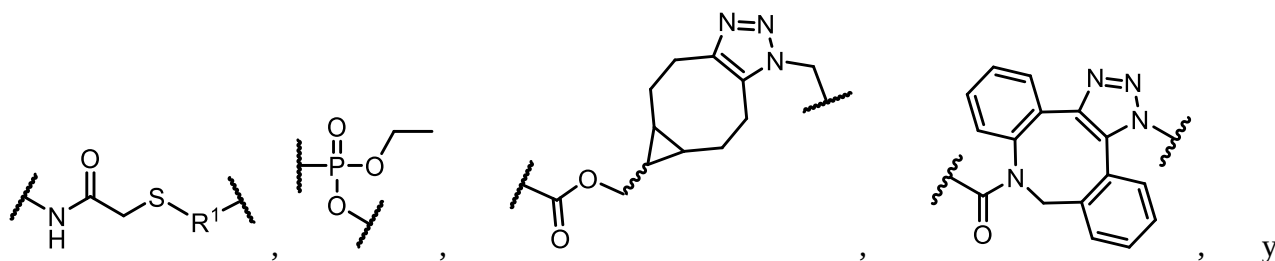
En la presente descripción se proporciona un compuesto que comprende un cCPP y una carga que comprende un compuesto antisentido (AC), por ejemplo, un oligonucleótido antisentido, que es complementario a un objetivo en una secuencia pre-ARNm, en donde el compuesto comprende además L, en donde el enlazador se conjuga al AC a través de un grupo de unión (M), en donde M se selecciona de:

25

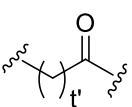


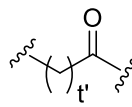
30

5



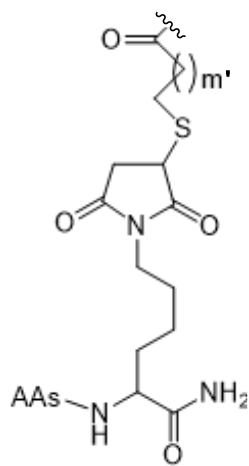
10

; en donde: R¹ es alquileo, cicloalquilo o , en donde t' es de 0 a 10 en donde cada R es independientemente un alquilo, alqueniilo, alquinilo, carbociclilo o heterociclilo, en

donde R¹ es , y t' es 2.

El enlazador puede tener la estructura:

15



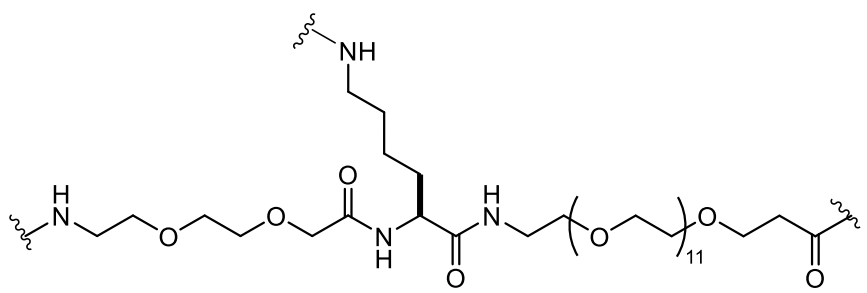
20

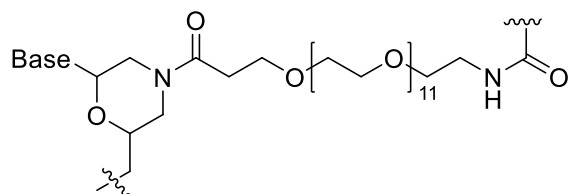
en donde AA_s es como se define en la presente descripción, y m' es 0-10.

25

El enlazador puede ser de la fórmula:

30

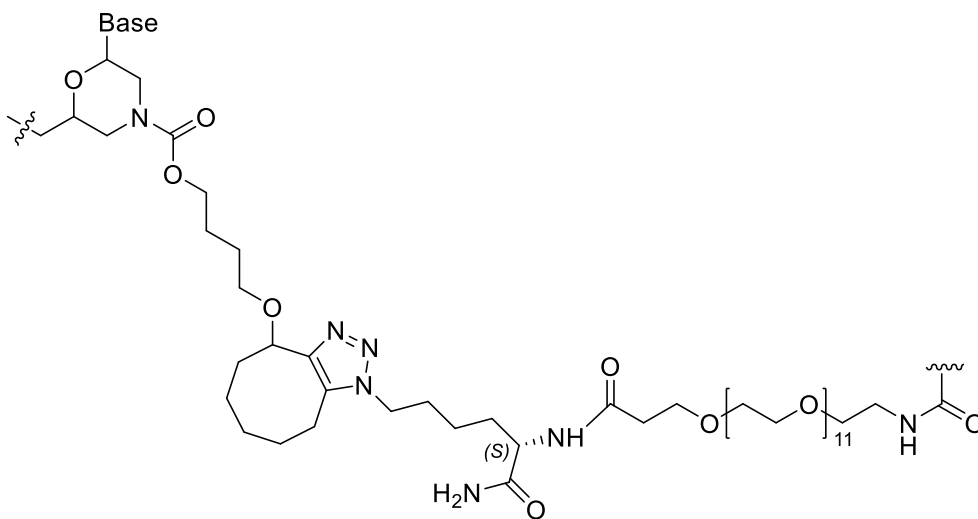




El enlazador puede ser de la fórmula: , en donde “base”

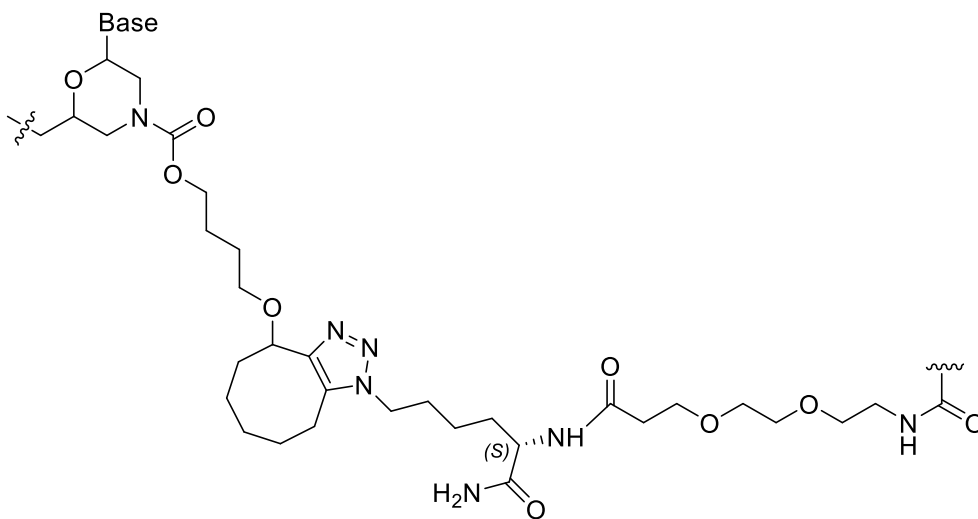
5 es una nucleobase en el extremo 3' de un oligómero de morfolino fosforamidato de carga.

El enlazador puede ser de la fórmula:

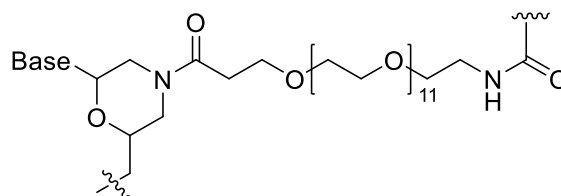


en donde “base” corresponde a una nucleobase en el extremo 3' de un oligómero de morfolino fosforamidato de carga.

El enlazador puede ser de la fórmula:



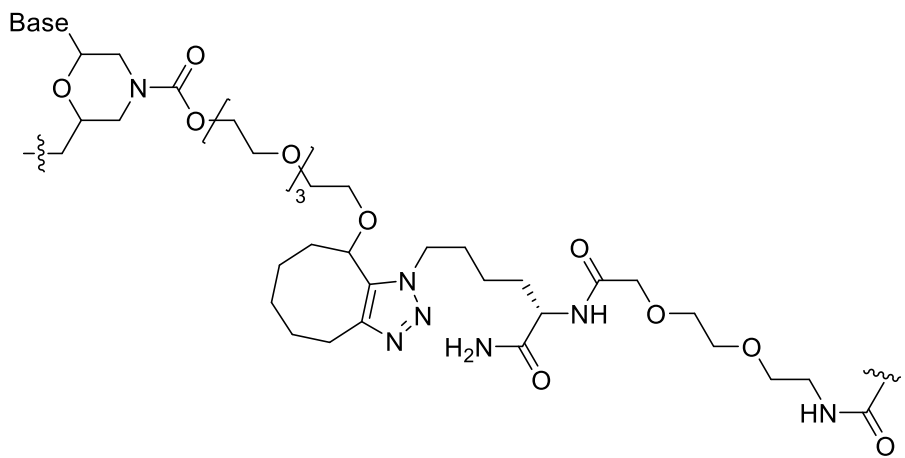
en donde “base” es una nucleobase en el extremo 3' de un oligómero de morfolino fosforamidato de carga.



El enlazador puede ser de la fórmula: , en donde “base”

es una nucleobase en el extremo 3' de un oligómero de morfolino fosforamidato de carga.

El enlazador puede ser de la fórmula:



El enlazador puede unirse covalentemente a una carga en cualquier ubicación adecuada en la carga.

El enlazador se une covalentemente al extremo 3' de la carga o al extremo 5' de una carga oligonucleotídica. El enlazador puede unirse covalentemente a la cadena principal de una carga.

El enlazador puede unirse a la cadena lateral del ácido aspártico, ácido glutámico, glutamina, asparagina o lisina, o a una cadena lateral modificada de glutamina o asparagina (por ejemplo, una cadena lateral reducida que tiene un grupo amino), en el cCPP. El enlazador puede unirse a la cadena lateral de la lisina en el cCPP.

Conjugados cCPP-enlazador

El cCPP puede conjugarse a un enlazador definido en la presente descripción. El enlazador puede conjugarse a un AA_{SC} del cCPP como se define en la presente descripción.

El enlazador puede comprender una subunidad $-(OCH_2CH_2)_{z'}$ (por ejemplo, como espaciador), en donde z' es un número entero de 1 a 23, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 o 23. “ $-(OCH_2CH_2)_{z'}$ ” también se denomina PEG. El conjugado cCPP-enlazador puede tener una estructura seleccionada de la Tabla 7:

Tabla 7: Conjugados cCPP-enlazador y SEQ ID NO

<i>ciclo</i> (FfΦ-4gp-r-4gp-rQ)-PEG ₄ -K-NH ₂	<i>ciclo</i> (SEQ ID NO:118)-PEG ₄ -K-NH ₂
<i>ciclo</i> (FfΦ-Cit-r-Cit-rQ)-PEG ₄ -K-NH ₂	<i>ciclo</i> (SEQ ID NO:119)-PEG ₄ -K-NH ₂
<i>ciclo</i> (FfΦ-Pia-r-Pia-rQ)-PEG ₄ -K-NH ₂	<i>ciclo</i> (SEQ ID NO:120)-PEG ₄ -K-NH ₂
<i>ciclo</i> (FfΦ-Dml-r-Dml-rQ)-PEG ₄ -K-NH ₂	<i>ciclo</i> (SEQ ID NO:121)-PEG ₄ -K-NH ₂
<i>ciclo</i> (FfΦ-Cit-r-Cit-rQ)-PEG ₁₂ -OH	<i>ciclo</i> (SEQ ID NO:122)-PEG ₁₂ -OH
<i>ciclo</i> (fΦR-Cit-R-Cit-Q)-PEG ₁₂ -OH	<i>ciclo</i> (SEQ ID NO:123)-PEG ₁₂ -OH

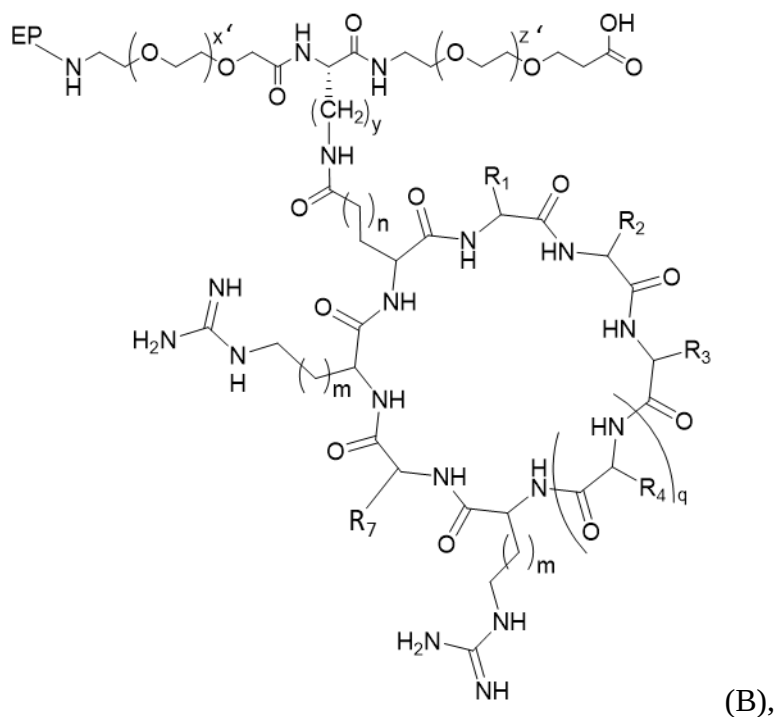
El enlazador puede comprender una subunidad -(OCH₂CH₂)_{z'}, en donde z' es un número entero de 1 a 23, y una subunidad peptídica. La subunidad peptídica puede comprender de 2 a 10 aminoácidos.

El conjugado cCPP-enlazador puede tener una estructura seleccionada de la Tabla 8:

Tabla 8: Conjugado cCPP-enlazador y SEQ ID NO

Ac-PKKKRKV-Lys(<i>ciclo</i> [FfΦ-R-r-Cit-rQ])-PEG ₁₂ -K(N ₃)-NH ₂	Ac-SEQ ID NO:42-Lys(<i>ciclo</i> [SEQ ID NO:127])-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-Lys(<i>ciclo</i> [FfΦ-Cit-r-R-rQ])-PEG ₁₂ -K(N ₃)-NH ₂	Ac-SEQ ID NO:42-Lys(<i>ciclo</i> [SEQ ID NO:128])-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-K(<i>ciclo</i> (FfΦR-cit-R-cit-Q))-PEG ₁₂ -K(N ₃)-NH ₂	Ac-SEQ ID NO:42-K(<i>ciclo</i> (SEQ ID NO:129))-PEG ₁₂ -K(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-B-k(N ₃)-NH ₂	Ac-SEQ ID NO:42-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-B-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG2-k(N ₃)-NH ₂	Ac-SEQ ID NO:42-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG2-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG4-k(N ₃)-NH ₂	Ac-SEQ ID NO:42-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG4-k(N ₃)-NH ₂
Ac-PKKKRKV-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂	Ac-SEQ ID NO:42-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-k(N ₃)-NH ₂
Ac-pkkkrkv-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂	Ac-SEQ ID NO:131-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-k(N ₃)-NH ₂
Ac-rrv-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-OH	Ac-rrv-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-OH
Ac-PKKKRKV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂	Ac-SEQ ID NO:42-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-k(N ₃)-NH ₂
Ac-PKKK-Cit-KV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-k(N ₃)-NH ₂	Ac-SEQ ID NO:126-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-k(N ₃)-NH ₂
Ac-PKKKRKV-PEG2-Lys(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG12-K(N ₃)-NH ₂	Ac-SEQ ID NO:42-PEG2-Lys(<i>ciclo</i> [SEQ ID NO:130])-PEG12-K(N ₃)-NH ₂

Se proporcionan EEV que comprenden un péptido penetrante de células cíclico (cCPP), un enlazador y un péptido exocíclico (EP). Un EEV puede comprender la estructura de fórmula (B):



o una forma protonada de esta,

en donde:

R_1 , R_2 y R_3 son cada uno independientemente H o una cadena lateral aromática o heteroaromática de un aminoácido;

R_4 y R_7 son independientemente H o una cadena lateral de aminoácido;

EP es un péptido exocíclico como se define en la presente descripción;

cada m es independientemente un número entero de 0-3;

n es un número entero de 0-2;

x' es un número entero de 1-20;

y es un número entero de 1-5;

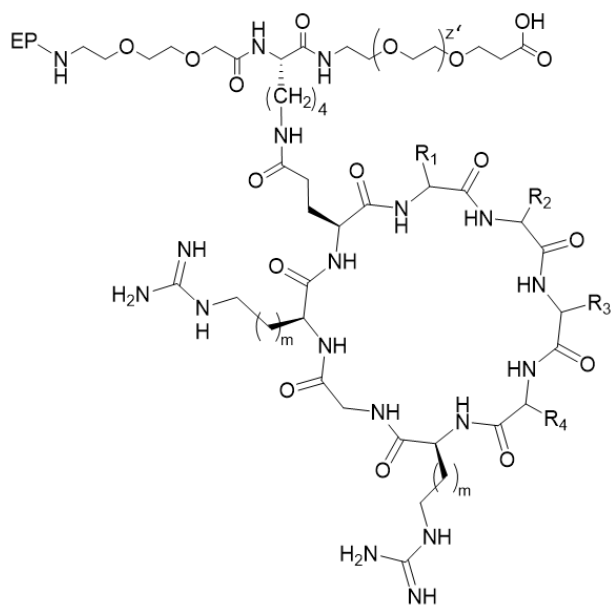
q es 1-4; y

z' es un número entero de 1-23.

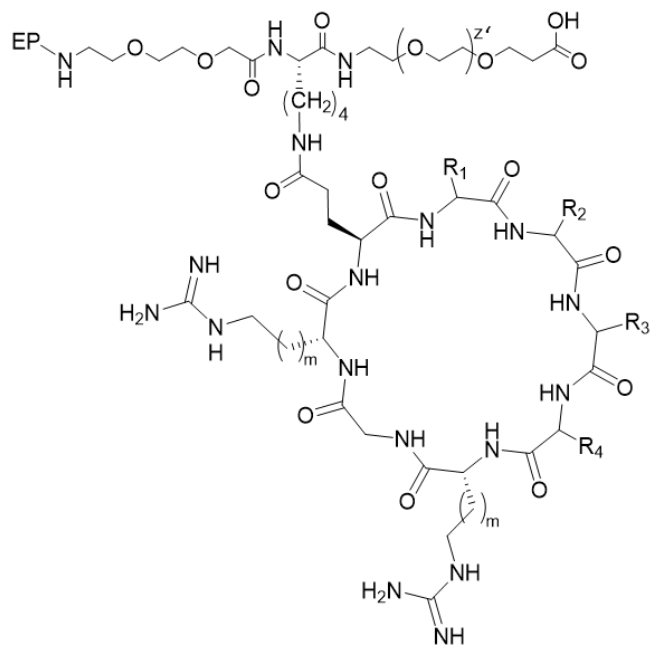
R_1 , R_2 , R_3 , R_4 , R_7 , EP, m, q, y, x' , z' son como se describe en la presente descripción.

n puede ser 0. n puede ser 1. n puede ser 2.

El EEV puede comprender la estructura de fórmula (B-a) o (B-b):



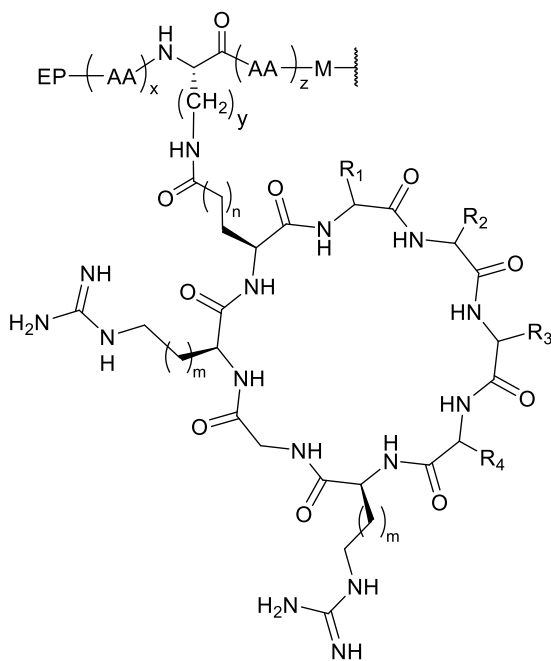
(B-a),



(B-b),

o una forma protonada de estas, en donde EP, R¹, R², R³, R⁴, m y z' son como se definió anteriormente en la fórmula (B).

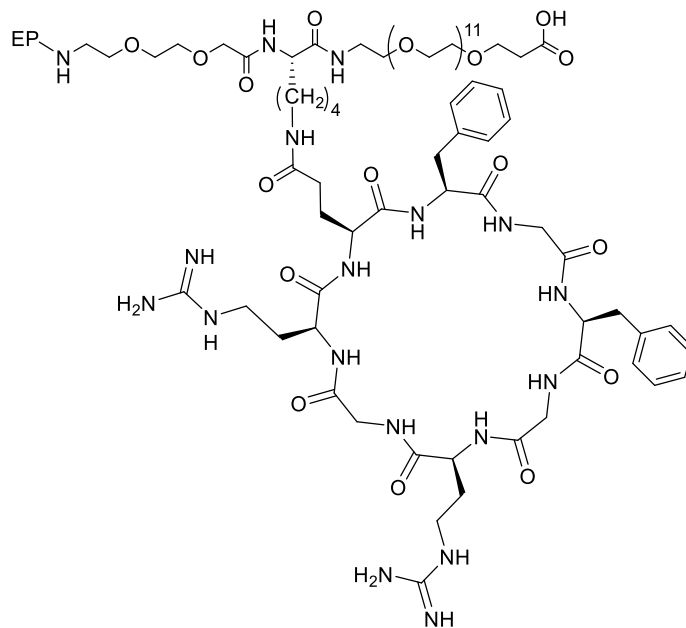
El EEV puede comprender la estructura de fórmula (B-c):



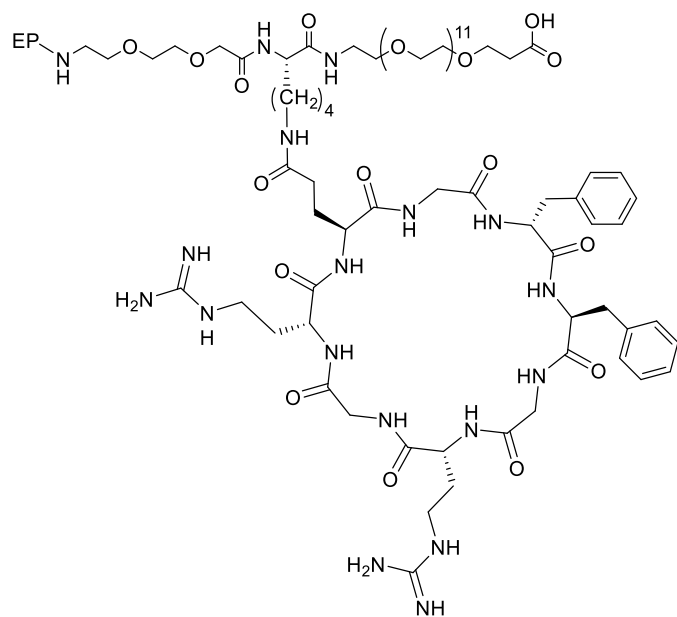
(B-c),

o una forma protonada de esta, en donde EP, R¹, R², R³, R⁴ y m son como se definió anteriormente en la fórmula (B); AA es un aminoácido como se define en la presente descripción; M es como se define en la presente descripción; n es un número entero de 0-2; x es un número entero de 1-10; y es un número entero de 1-5; y z es un número entero de 1-10.

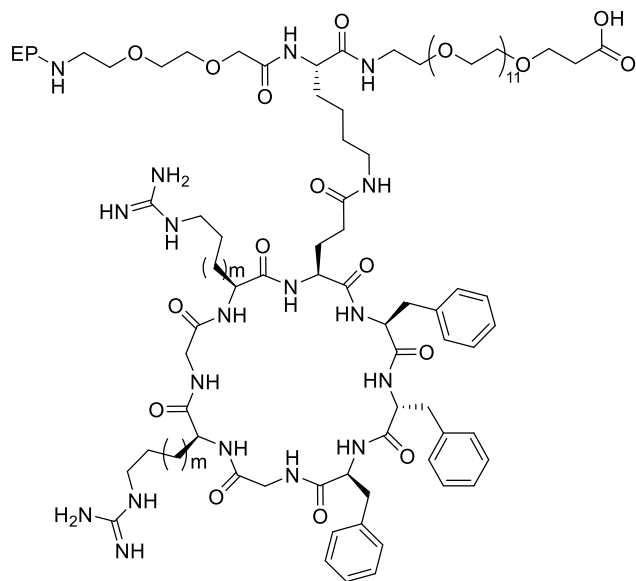
El EEV puede tener la estructura de fórmula (B-1), (B-2), (B-3) o (B-4):



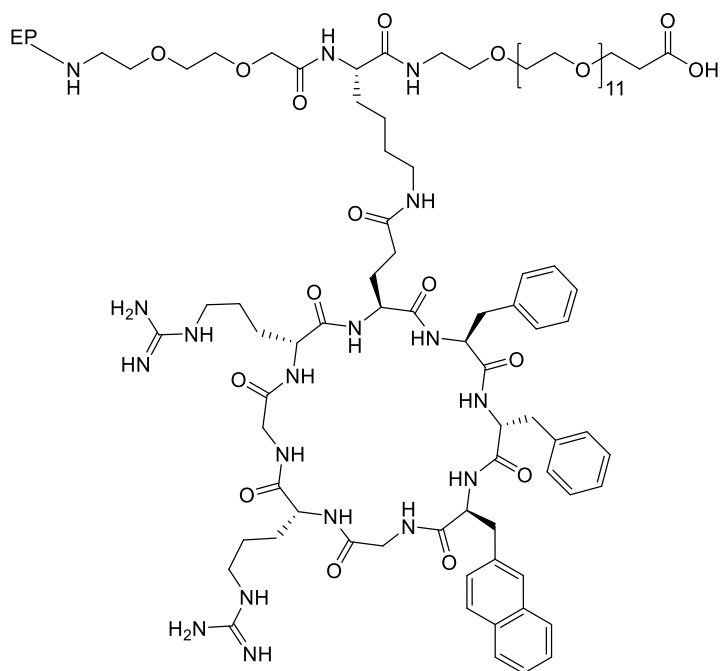
(B-1),



(B-2),



(B-3),

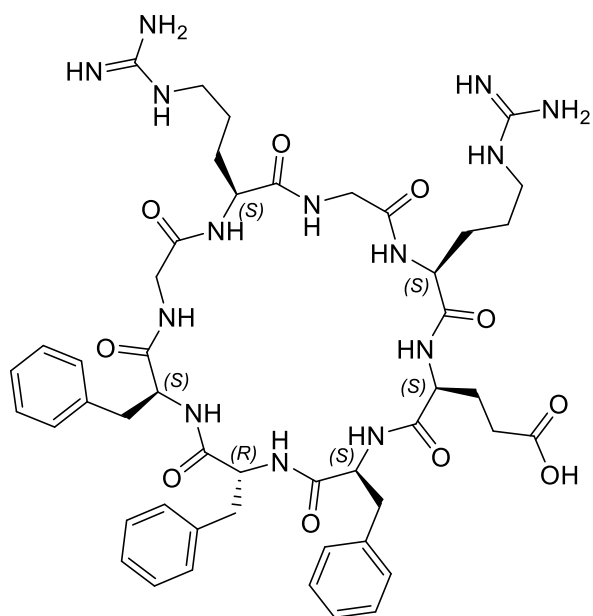


(B-4)

o una forma protonada de estas, en donde EP es como se definió anteriormente en la fórmula (B).

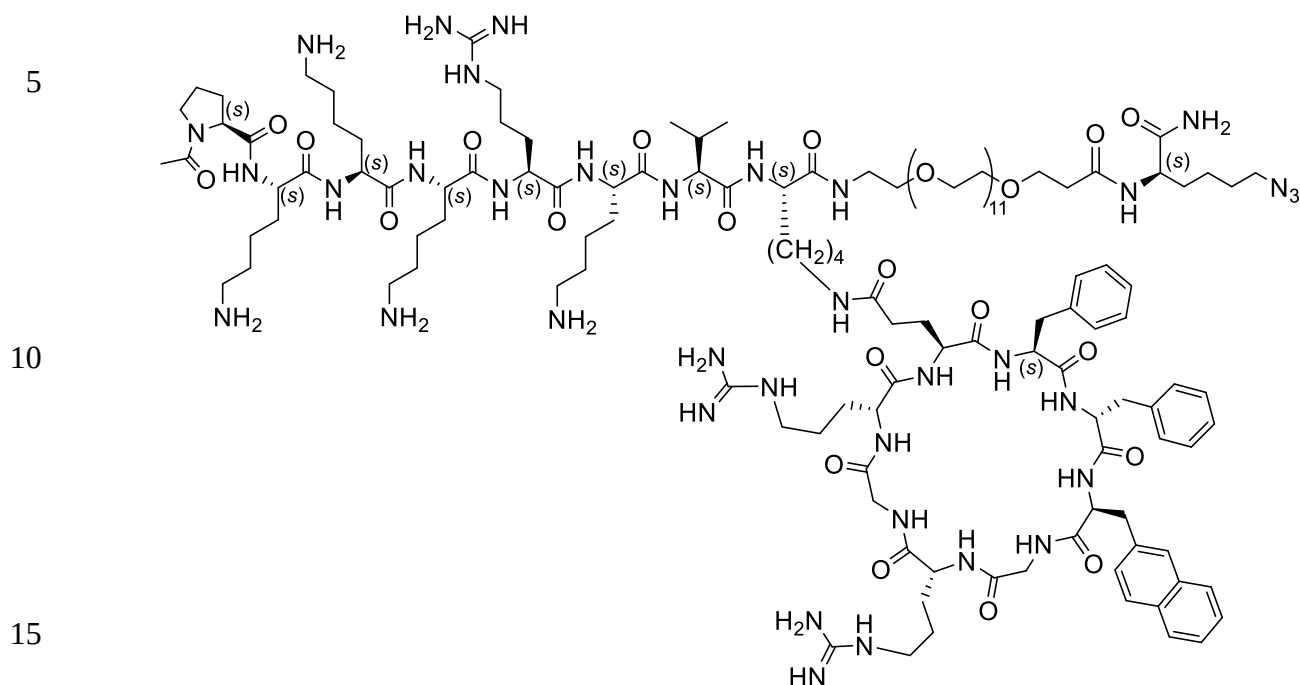
El EEV puede comprender la fórmula (B) y puede tener la estructura: Ac-PKKKRKVAEEA-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH (Ac-SEQ ID NO:132- K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH) o Ac-PK-KKR-KV-AEEA-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH (Ac- SEQ ID NO:133-K(*ciclo*[SEQ ID NO:83])-PEG₁₂-OH).

El EEV puede comprender un cCPP de fórmula:



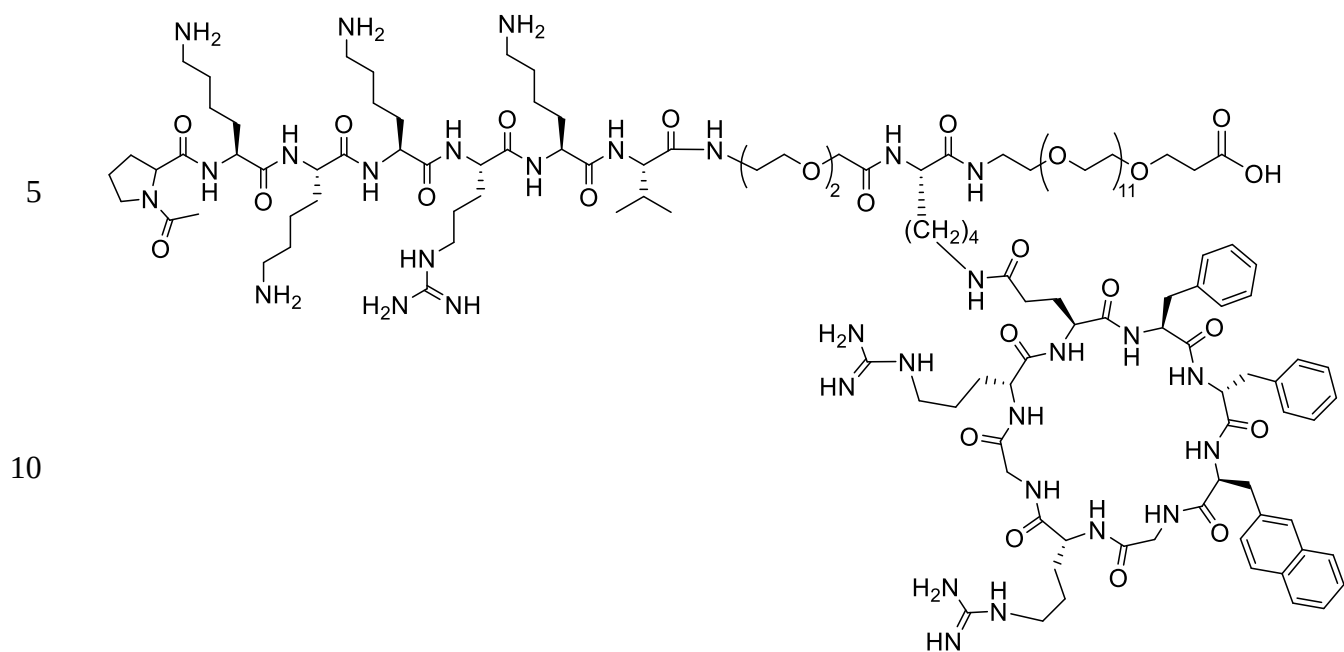
El EEV puede comprender la fórmula: Ac-PKKKRKV-miniPEG₂-Lys(*ciclo*(FfFGRGRQ))-PEG₂-

El EEV puede ser:

[illegible]

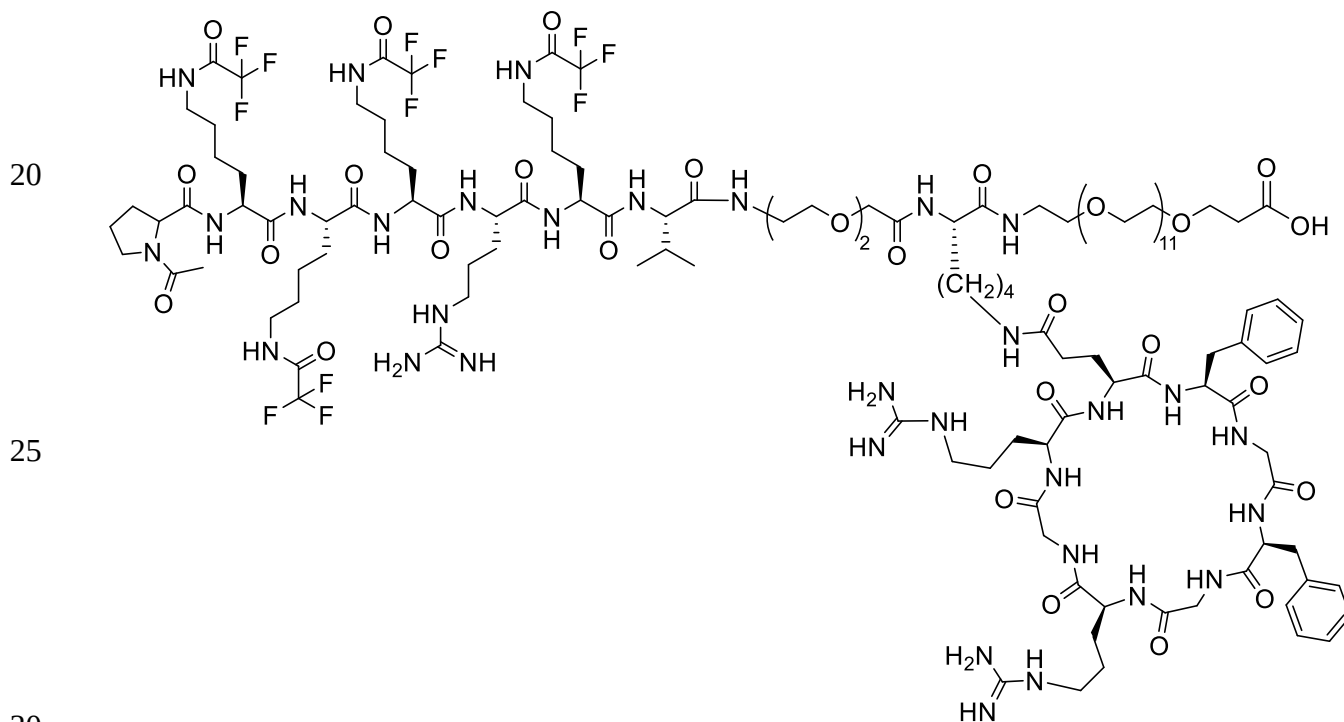
El EEV puede ser Ac-P-K(Tfa)-K(Tfa)-K(Tfa)-R-K(Tfa)-V-miniPEG₂-K(*ciclo*(Ff-Nal-GrGrQ)-PEG₁₂-OH (Ac-SEQ ID NO:134-miniPEG₂-K(*ciclo*(SEQ ID NO:135)-PEG₁₂-OH).

El EEV puede ser

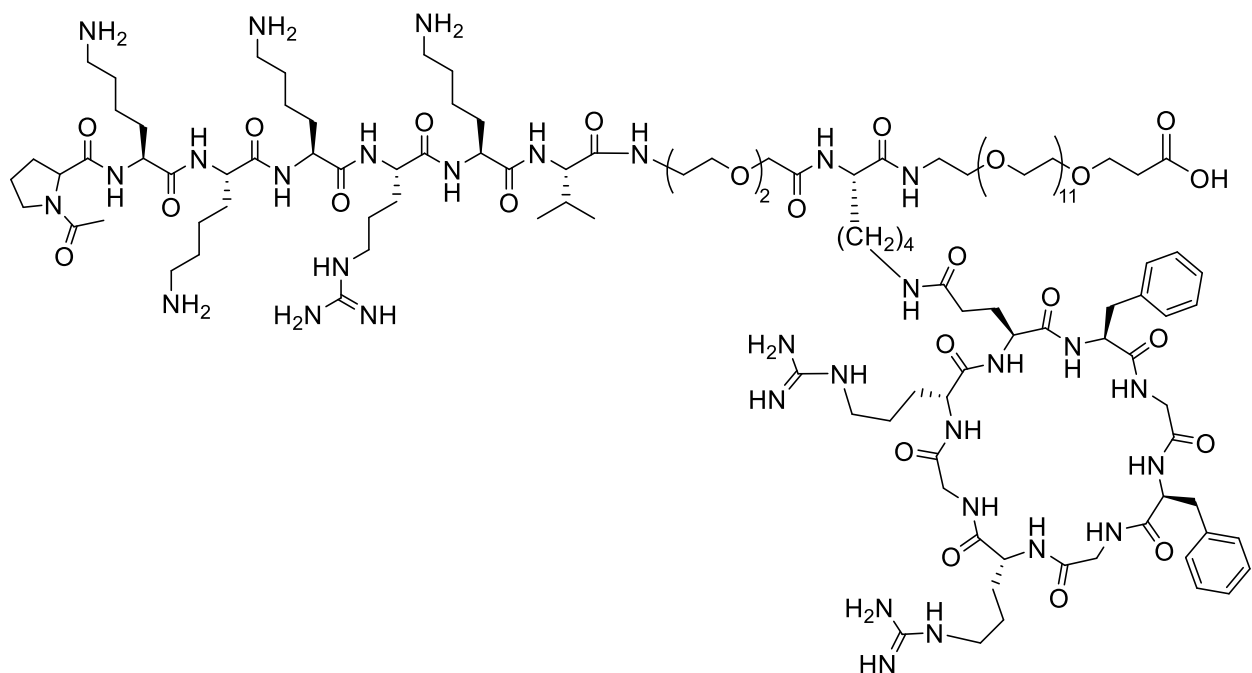


El EEV puede ser Ac-P-K-K-K-R-K-V-miniPEG₂-K(ciclo(Ff-Nal-GrGrQ))-PEG₁₂-OH (Ac- SEQ ID NO:42-miniPEG₂-K(ciclo(SEQ ID NO:135))-PEG₁₂-OH).

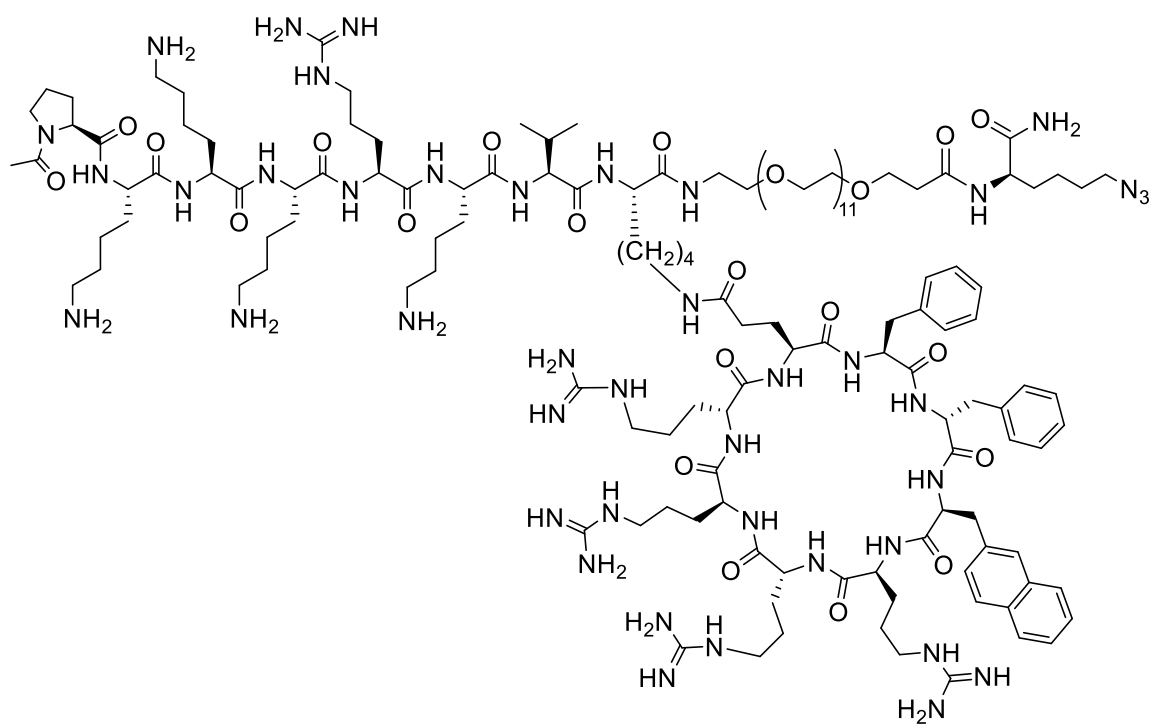
El EEV puede ser



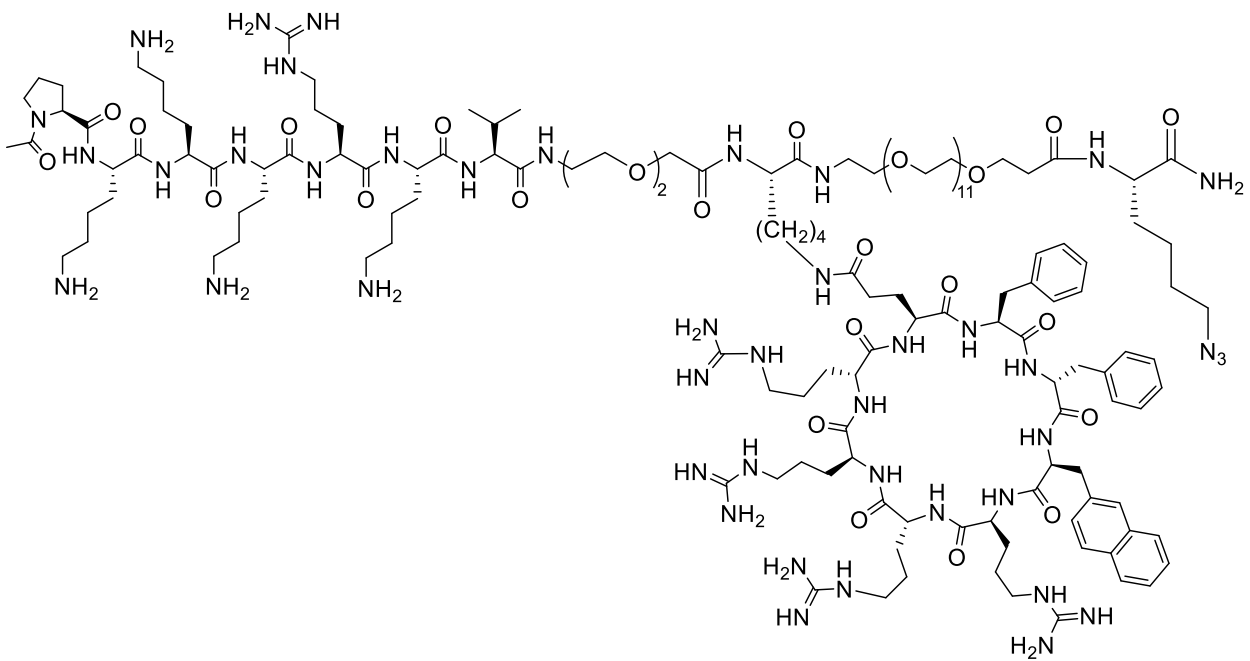
El EEV puede ser



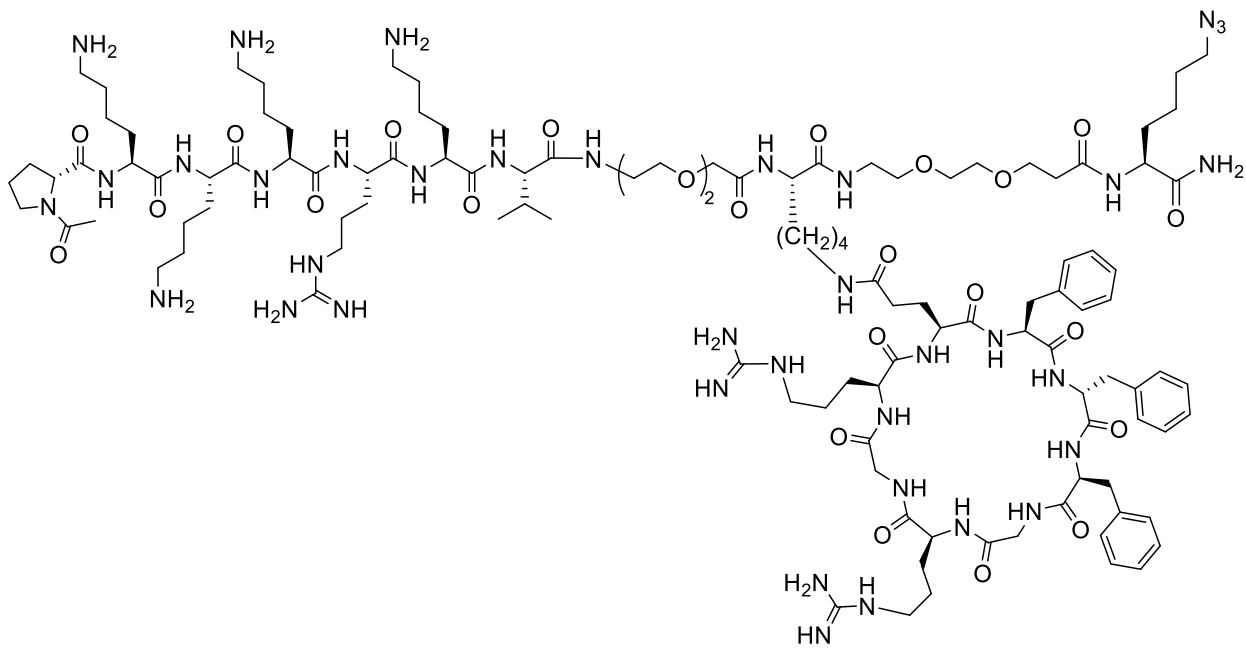
El EEV puede ser



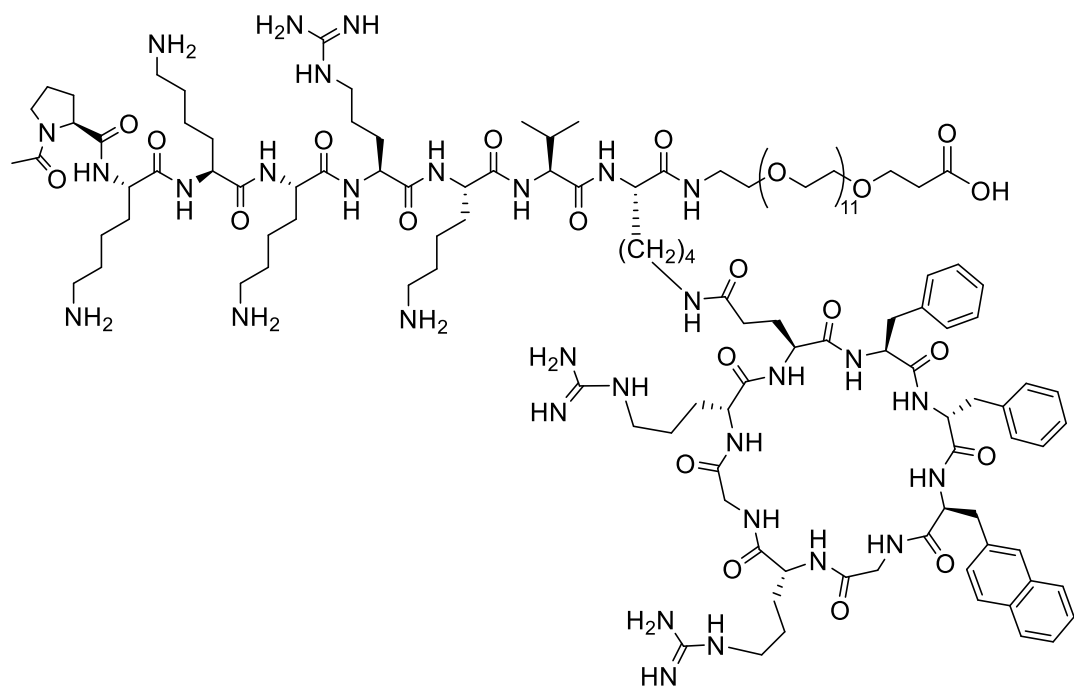
El EEV puede ser



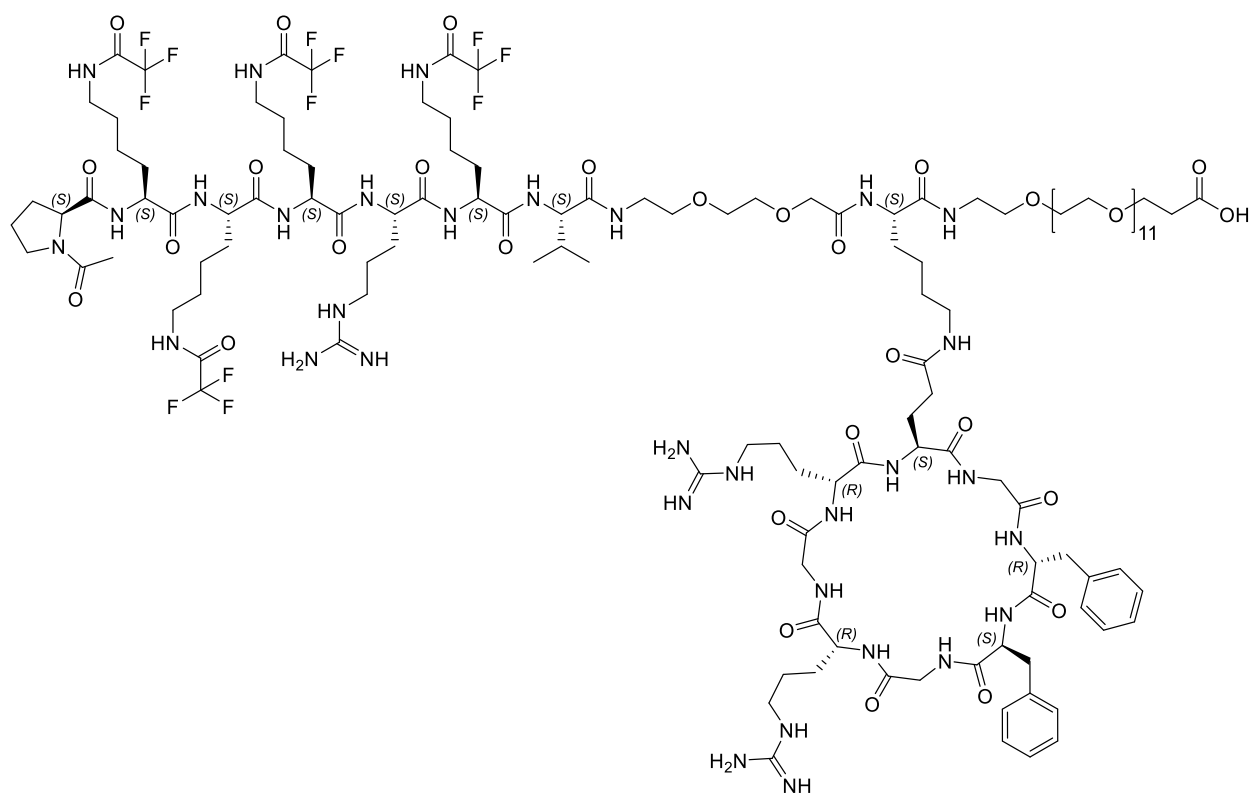
El EEV puede ser



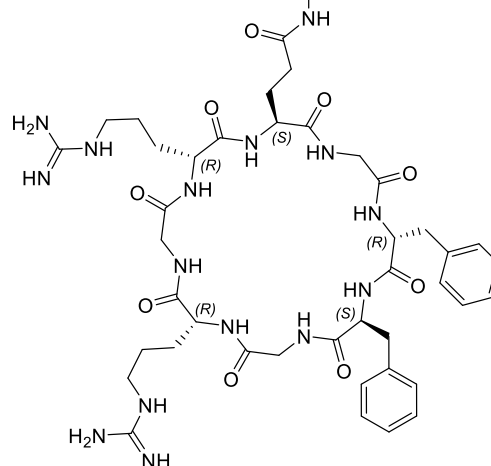
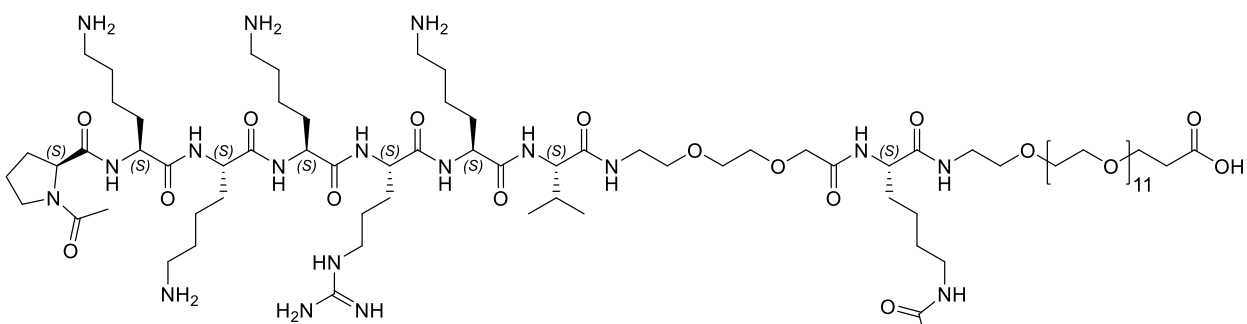
El EEV puede ser



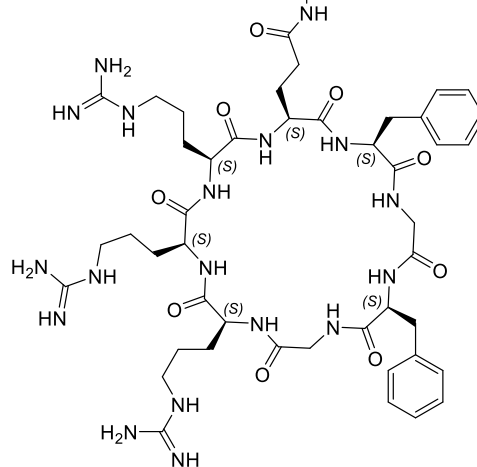
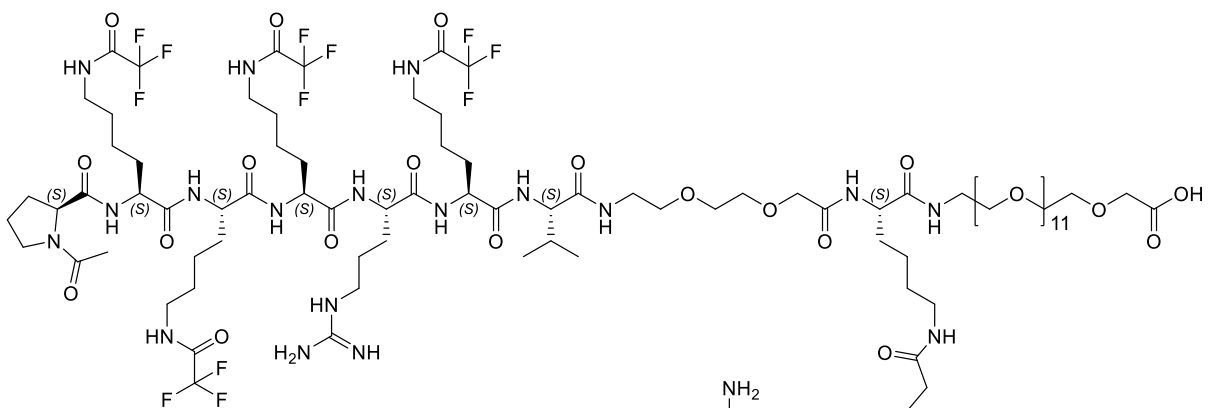
El EEV puede ser:



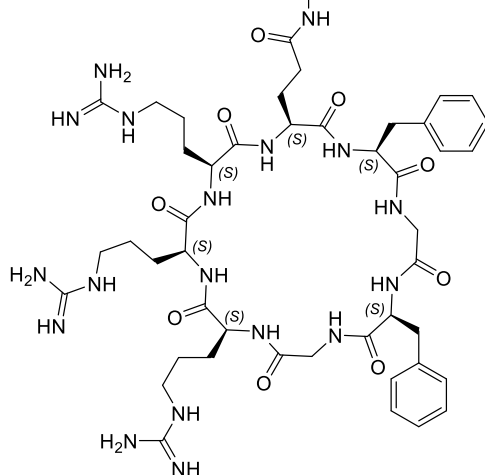
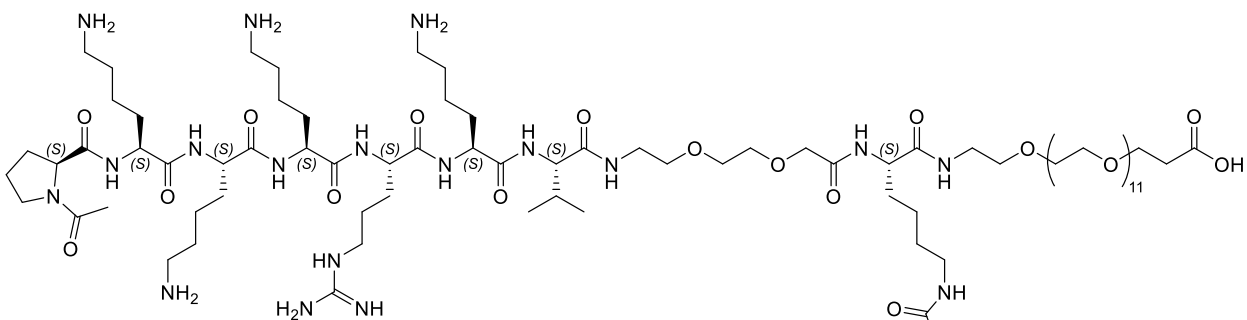
El EEV puede ser



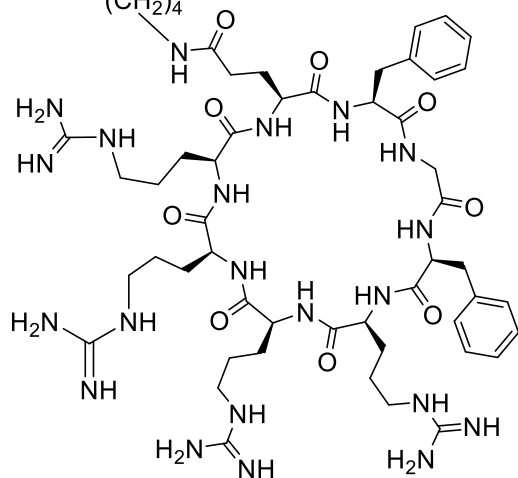
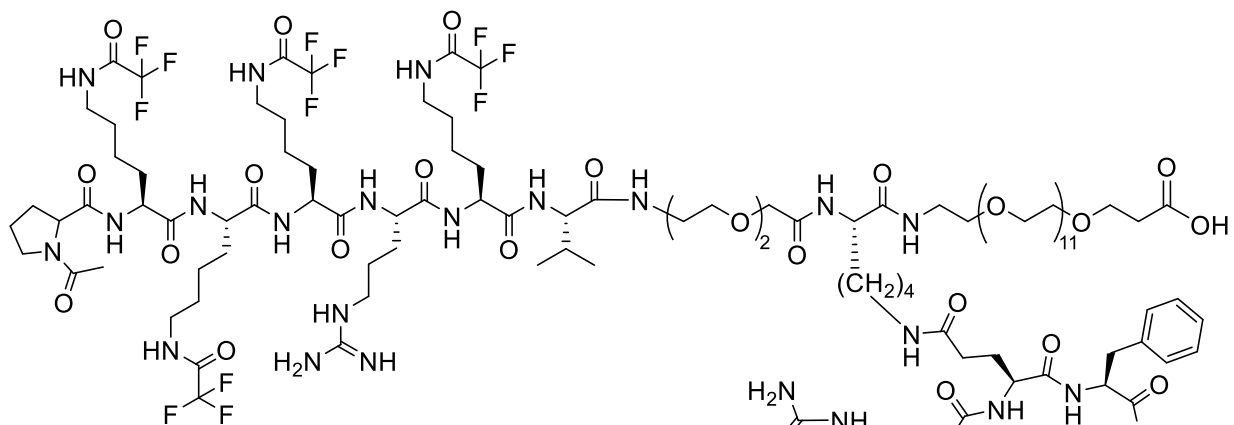
El EEV puede ser



El EEV puede ser



El EEV puede ser



El EEV se puede seleccionar de

5	Ac-rr-miniPEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-rr-miniPEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-frr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-frr-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rfr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-rfr-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rbfbr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-SEQ ID NO:137-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rrr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-rrr-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
10	Ac-rbr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-rbr-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rbrbr-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-SEQ ID NO:138-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-hh-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-hh-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
15	Ac-hbh-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-hbh-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-hbbbh-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-SEQ ID NO: 139-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rbbbh-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac- SEQ ID NO: 140-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
20	Ac-hbrbh-PEG ₂ -Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-PEG ₁₂ -OH	Ac-SEQ ID NO:141-PEG ₂ -Dap(<i>ciclo</i> [SEQ ID NO:136])-PEG ₁₂ -OH
	Ac-rr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rr-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-frr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-frr-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
25	Ac-rfr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rfr-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-rbfbr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:137-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-rrr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rrr-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
30	Ac-rbr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rbr-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-rbrbr-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:138-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-hh-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-hh-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH

5	Ac-hbh-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-hbh-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-hbbbh-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO: 139-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-rbbbh-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-SEQ ID NO: 140-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-hbrbh-Dap(<i>ciclo</i> [FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:141-Dap(<i>ciclo</i> [SEQ ID NO:136])-b-OH
	Ac-KKKK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:7-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:Z80])-miniPEG ₂ -K(N ₃)-NH ₂
10	Ac-KGKK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:13-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:Z80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KKGK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:14-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:Z80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KKK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KKK-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KK-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
15	Ac-KGK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KGK-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KBK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KBK-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KBKBK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:24-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KR-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KR-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
20	Ac-KBR-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac-KBR-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PKKKRKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:42-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PKKKRKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:42-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PGKKRKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:43-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
25	Ac-PKGKRKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:44-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
30		

5	Ac-PKKGRKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:45-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PKKKGKV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:46-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PKKKRGV-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:47-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-PKKKRKG-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:48-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
10	Ac-KKKRK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂	Ac- SEQ ID NO:19-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂
	Ac-KKRK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂ y	Ac- SEQ ID NO:8-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂ y
	Ac-KRK-miniPEG ₂ -Lys(<i>ciclo</i> [FfΦGrGrQ])-miniPEG ₂ -K(N ₃)-NH ₂ .	Ac-KRK-miniPEG ₂ -Lys(<i>ciclo</i> [SEQ ID NO:80])-miniPEG ₂ -K(N ₃)-NH ₂ .

15

El EEV se puede seleccionar de:

Ac-PKKKRKV-Lys(*ciclo*[FfΦGrGrQ])-PEG₁₂-K(N₃)-NH₂

(Ac- SEQ ID NO:42-Lys(*ciclo*[SEQ ID NO:80])-PEG₁₂-K(N₃)-NH₂)

Ac-PKKKRKV-miniPEG₂-Lys(*ciclo*[FfΦGrGrQ])-miniPEG₂-K(N₃)-NH₂

20

(Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:80])-miniPEG₂-K(N₃)-NH₂)

Ac-PKKKRKV-miniPEG₂-Lys(*ciclo*[FGFGRGRQ])-miniPEG₂-K(N₃)-NH₂

(Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:82])-miniPEG₂-K(N₃)-NH₂)

Ac-KR-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂

(Ac-KR-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₂-K(N₃)-NH₂)

25

Ac-PKKKGKV-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂

(Ac- SEQ ID NO:46-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₂-K(N₃)-NH₂)

Ac-PKKKRKG-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂

(Ac- SEQ ID NO:48-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₂-K(N₃)-NH₂)

Ac-KKKRK-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂

30

(Ac- SEQ ID NO:19-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₂-K(N₃)-NH₂)

Ac-PKKKRKV-miniPEG₂-Lys(*ciclo*[FFΦGRGRQ])-miniPEG₂-K(N₃)-NH₂

(Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:80])-miniPEG₂-K(N₃)-NH₂)

Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[βhFfΦGrGrQ])-miniPEG₂-K(N₃)-NH₂
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:142])-miniPEG₂-K(N₃)-NH₂)
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FfΦSrSrQ])-miniPEG₂-K(N₃)-NH₂
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:143])-miniPEG₂-K(N₃)-NH₂).

5 El EEV se puede seleccionar de:

Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*(GfFGrGrQ))-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*(SEQ ID NO:133))-PEG₁₂-OH)
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFKRKRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:144])-PEG₁₂-OH)

10 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFRGRGQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:145])-PEG₁₂-OH)
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFGRGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:146])-PEG₁₂-OH)
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFGRrRQ])-PEG₁₂-OH

15 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:147])-PEG₁₂-OH)
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:84])-PEG₁₂-OH) y
 Ac-PKKKKRKV-miniPEG₂-Lys(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:42-miniPEG₂-Lys(*ciclo*[SEQ ID NO:85])-PEG₁₂-OH).

20 El EEV se puede seleccionar de:

Ac-K-K-K-R-K-G-miniPEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac-SEQ ID NO:148-miniPEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-K-K-K-R-K-miniPEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:19-miniPEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)

25 Ac-K-K-R-K-K-PEG₄-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:22-PEG₄-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-K-R-K-K-K-PEG₄-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:21-PEG₄-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-K-K-K-K-R-PEG₄-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH

30 (Ac- SEQ ID NO:23-PEG₄-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-R-K-K-K-K-PEG₄-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:20-PEG₄-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH) y

Ac-K-K-K-R-K-PEG₄-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
(Ac- SEQ ID NO:19-PEG₄-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH).

El EEV se puede seleccionar de:

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₂-K(N₃)-NH₂

5 (Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₂-K(N₃)-NH₂)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₂-K(N₃)-NH₂

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:133])-PEG₂-K(N₃)-NH₂) y

10 Ac- PKKKRKV-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:133])-PEG₁₂-OH).

La carga puede ser un AC y el EEV se puede seleccionar de:

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:80])-PEG₁₂-OH)

15 Ac-PKKKKRKV-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:79])-PEG₁₂-OH)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:81])-PEG₁₂-OH)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH

20 (Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:133])-PEG₁₂-OH)

Ac-PKKKKRKV-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:84])-PEG₁₂-OH)

25 Ac-PKKKKRKV-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH

(Ac- SEQ ID NO:42-PEG₂-K(*ciclo*[SEQ ID NO:85])-PEG₁₂-OH)

Ac-rr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH

(Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:80])-PEG₁₂-OH)

Ac-rr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH

30 (Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:79])-PEG₁₂-OH)

Ac-rr-PEG₂-K(*ciclo*[FfF-GRGRQ])-PEG₁₂-OH

(Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:81])-PEG₁₂-OH)

Ac-rr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:82]))-PEG₁₂-OH
 Ac-rr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:133]))-PEG₁₂-OH
 5 Ac-rr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:84]))-PEG₁₂-OH
 Ac-rr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac-rr-PEG₂-K(*ciclo*[SEQ ID NO:85]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 10 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:80]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:79]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:81]))-PEG₁₂-OH
 15 Ac-rrr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:82]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:133]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 20 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:84]))-PEG₁₂-OH
 Ac-rrr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac-rrr-PEG₂-K(*ciclo*[SEQ ID NO:85]))-PEG₁₂-OH
 Ac-rhr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:80]))-PEG₁₂-OH
 25 Ac-rhr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:79]))-PEG₁₂-OH
 Ac-rhr-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:81]))-PEG₁₂-OH
 Ac-rhr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 30 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:82]))-PEG₁₂-OH
 Ac-rhr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:133]))-PEG₁₂-OH

Ac-rhr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:84])-PEG₁₂-OH)
 Ac-rhr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac-rhr-PEG₂-K(*ciclo*[SEQ ID NO:85])-PEG₁₂-OH)
 5 Ac-rbr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:80])-PEG₁₂-OH)
 Ac-rbr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:79])-PEG₁₂-OH)
 Ac-rbr-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 10 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:81])-PEG₁₂-OH)
 Ac-rbr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-rbr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:133])-PEG₁₂-OH)
 15 Ac-rbr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:84])-PEG₁₂-OH)
 Ac-rbr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac-rbr-PEG₂-K(*ciclo*[SEQ ID NO:85])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 20 (Ac-SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:80])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:79])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:81])-PEG₁₂-OH)
 25 Ac-rbrbr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:82])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:133])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 30 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:84])-PEG₁₂-OH)
 Ac-rbrbr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:138-PEG₂-K(*ciclo*[SEQ ID NO:85])-PEG₁₂-OH)

- Ac-rbhbr-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:80]))-PEG₁₂-OH)
 Ac-rbhbr-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:79]))-PEG₁₂-OH)
 5 Ac-rbhbr-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:81]))-PEG₁₂-OH)
 Ac-rbhbr-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:82]))-PEG₁₂-OH)
 Ac-rbhbr-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 10 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:133]))-PEG₁₂-OH)
 Ac-rbhbr-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:84]))-PEG₁₂-OH)
 Ac-rbhbr-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:149-PEG₂-K(*ciclo*[SEQ ID NO:85]))-PEG₁₂-OH)
 15 Ac-hbrbh-PEG₂-K(*ciclo*[FfΦGrGrQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:80]))-PEG₁₂-OH)
 Ac-hbrbh-PEG₂-K(*ciclo*[FfΦCit-r-Cit-rQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:79]))-PEG₁₂-OH)
 Ac-hbrbh-PEG₂-K(*ciclo*[FfFGRGRQ])-PEG₁₂-OH
 20 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:81]))-PEG₁₂-OH)
 Ac-hbrbh-PEG₂-K(*ciclo*[FGFGRGRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:82]))-PEG₁₂-OH)
 Ac-hbrbh-PEG₂-K(*ciclo*[GfFGrGrQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:133]))-PEG₁₂-OH)
 25 Ac-hbrbh-PEG₂-K(*ciclo*[FGFGRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:84]))-PEG₁₂-OH)
 Ac-hbrbh-PEG₂-K(*ciclo*[FGFRRRRQ])-PEG₁₂-OH
 (Ac- SEQ ID NO:141-PEG₂-K(*ciclo*[SEQ ID NO:85]))-PEG₁₂-OH),

en donde b es beta-alanina, y la secuencia exocíclica puede tener la estereoquímica D o L.

30 Carga

El péptido penetrante de células (CPP), tal como un péptido penetrante de células cíclico (por ejemplo, cCPP), puede conjugarse a una carga. Como se usa en la presente, “carga” es un compuesto

o resto para el cual se desea el suministro a una célula. La carga puede conjugarse al grupo carbonilo terminal de un enlazador. Al menos un átomo del péptido cíclico puede reemplazarse por una carga o al menos un par solitario puede formar un enlace a una carga. La carga puede conjugarse al cCPP mediante un enlazador. La carga puede conjugarse a un AA_{SC} mediante un enlazador. Al menos un átomo del cCPP puede reemplazarse por la carga o al menos un par solitario del cCPP forma un enlace a la carga. Un grupo hidroxilo en una cadena lateral de aminoácido del cCPP puede reemplazarse por un enlace a la carga. Un grupo hidroxilo en una cadena lateral de glutamina del cCPP puede reemplazarse por una unión a la carga. La carga puede conjugarse al cCPP mediante un enlazador. La carga puede conjugarse a un AA_{SC} mediante un enlazador.

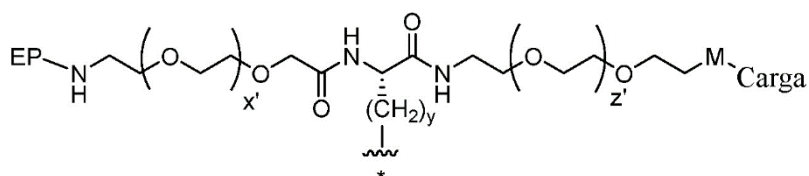
En modalidades, la cadena lateral de aminoácido comprende un grupo químicamente reactivo al cual se conjuga el enlazador o la carga. El grupo químicamente reactivo puede comprender un grupo amina, un ácido carboxílico, una amida, un grupo hidroxilo, un grupo sulfhidrilo, un grupo guanidinilo, un grupo fenólico, un grupo tioéter, un grupo imidazolilo o un grupo indolilo. En modalidades, el aminoácido del cCPP al que se conjuga la carga comprende lisina, arginina, ácido aspártico, ácido glutámico, asparagina, glutamina, homoglutamina, serina, treonina, tirosina, cisteína, arginina, tirosina, metionina, histidina o triptófano.

La carga puede comprender uno o más restos detectables, uno o más restos terapéuticos (TM), uno o más restos de direccionamiento, o cualquiera de sus combinaciones. En modalidades, la carga comprende un TM. En modalidades, la carga comprende un AC.

Péptidos penetrantes de células cíclicos (cCPP) conjugados a un resto de carga

El péptido penetrante de células cíclico (cCPP) puede conjugarse con un resto de carga.

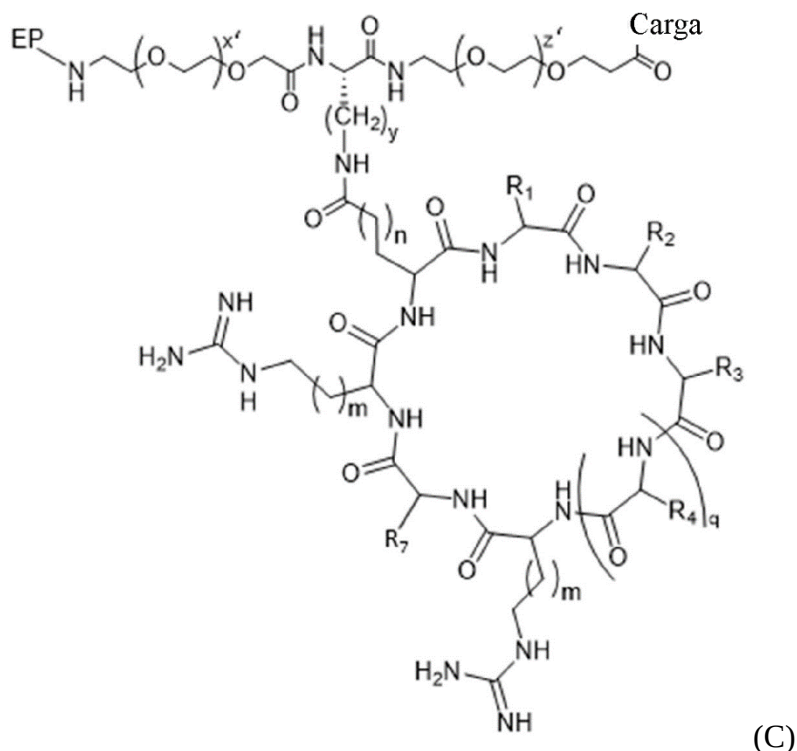
El resto de carga puede conjugarse al enlazador en el grupo carbonilo terminal para proporcionar la siguiente estructura:



, en donde:

EP es un péptido exocíclico y M, AA_{SC}, carga, x', y, y z' son como se definió anteriormente, * es el punto de unión al AA_{SC}. x' puede ser 1. y puede ser 4. z' puede ser 11. -(OCH₂CHππ₂)_{x'}- y/o -(OCH₂CHππ₂)_{z'}- pueden reemplazarse independientemente con uno o más aminoácidos, lo que incluye, por ejemplo, glicina, beta-alanina, ácido 4-aminobutírico, ácido 5-aminopentanoico, ácido 6-aminohexanoico, o sus combinaciones.

Un vehículo de escape endosómico (EEV) puede comprender un péptido penetrante de células cíclico (cCPP), un péptido exocíclico (EP) y un enlazador, y puede conjugarse a una carga para formar un conjugado de EEV que comprende la estructura de fórmula (C):



o una forma protonada de esta,
en donde:

20 R_1 , R_2 y R_3 pueden ser cada uno independientemente H o un residuo de aminoácido que tiene una cadena lateral que comprende un grupo aromático;

R_4 es H o una cadena lateral de aminoácido;

EP es un péptido exocíclico como se define en la presente descripción;

Carga es un resto como se define en la presente descripción;

25 cada m es independientemente un número entero de 0-3;

n es un número entero de 0-2;

x' es un número entero de 2-20;

y es un número entero de 1-5;

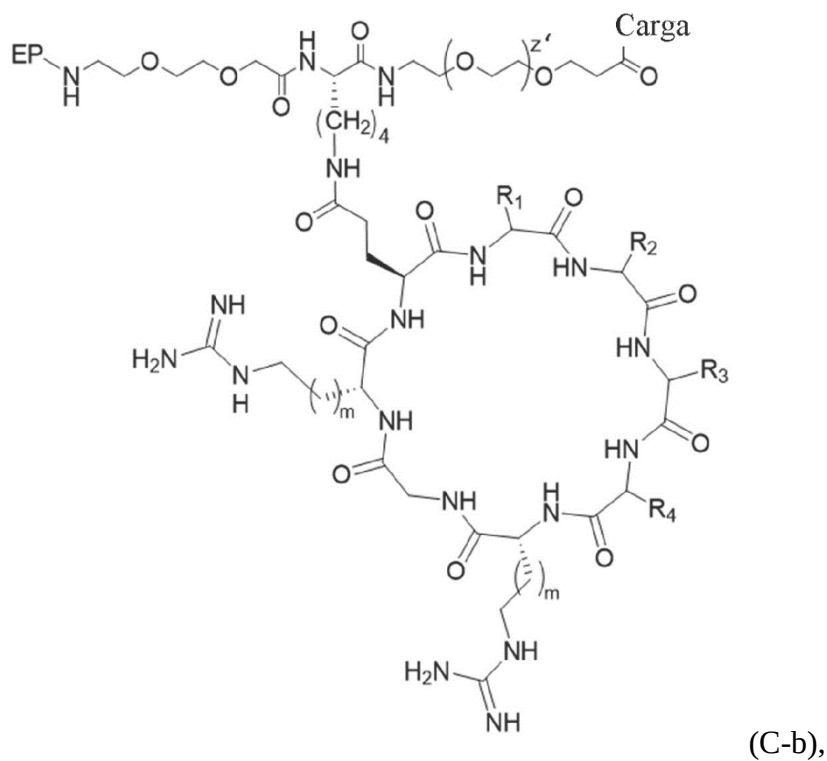
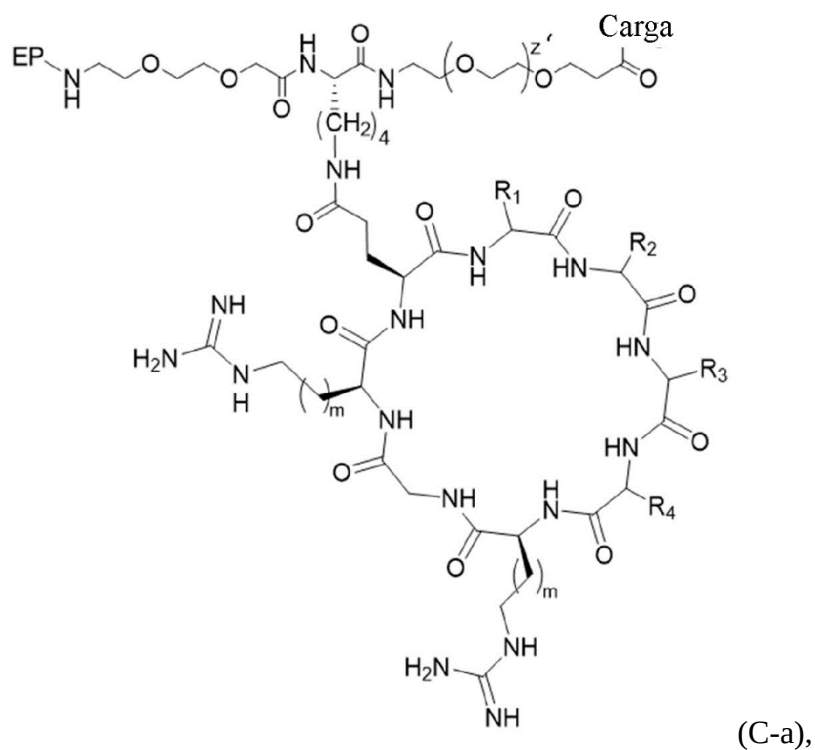
q es un número entero de 1-4; y

30 z' es un número entero de 2-20.

R_1 , R_2 , R_3 , R_4 , EP, carga, m, n, x' , y, q, y z' son como se define en la presente descripción.

El EEV puede conjugarse a una carga y el conjugado de EEV puede comprender la estructura de

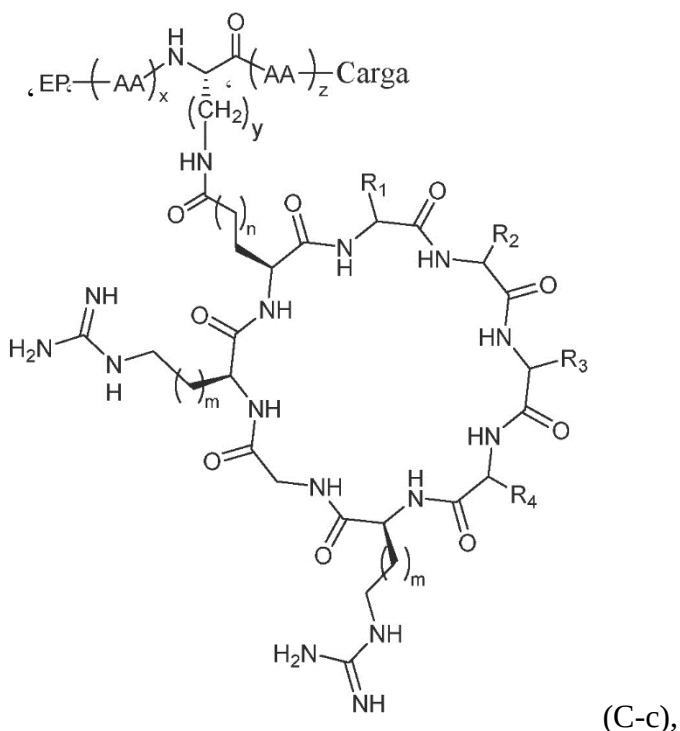
fórmula (C-a) o (C-b):



o una forma protonada de esta, en donde EP, m y z son como se definió anteriormente en la fórmula (C).

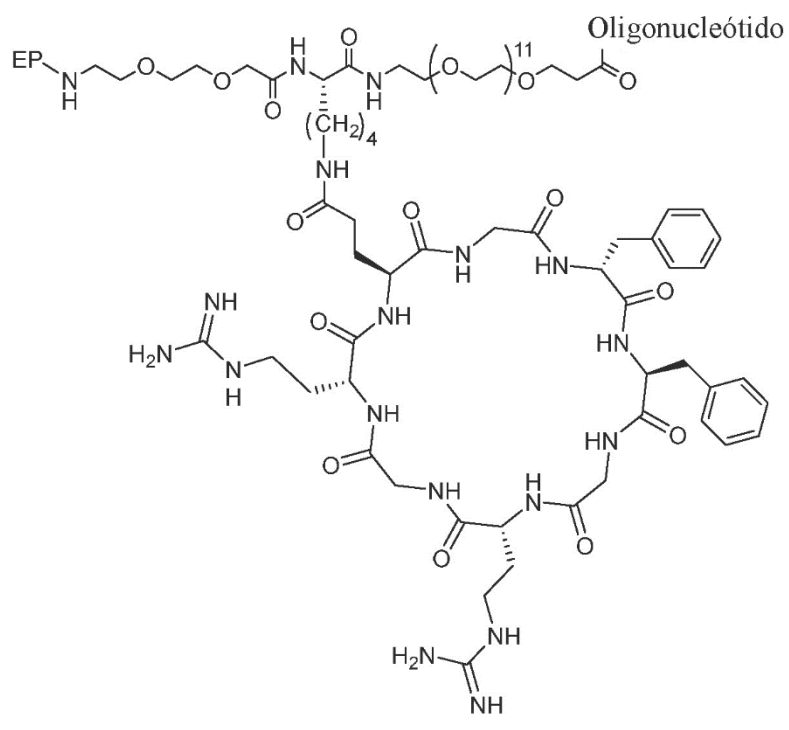
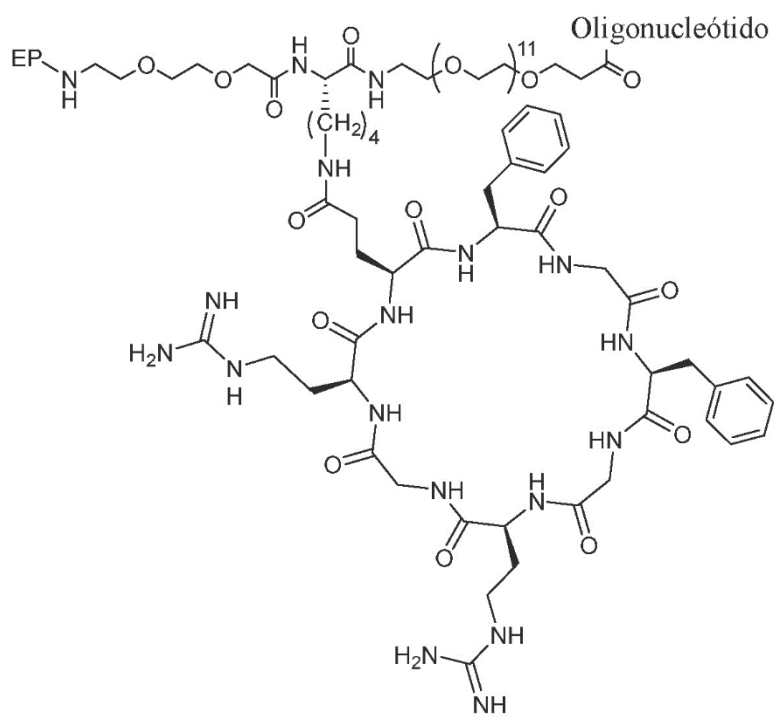
El EEV puede conjugarse a una carga y el conjugado de EEV puede comprender la estructura de

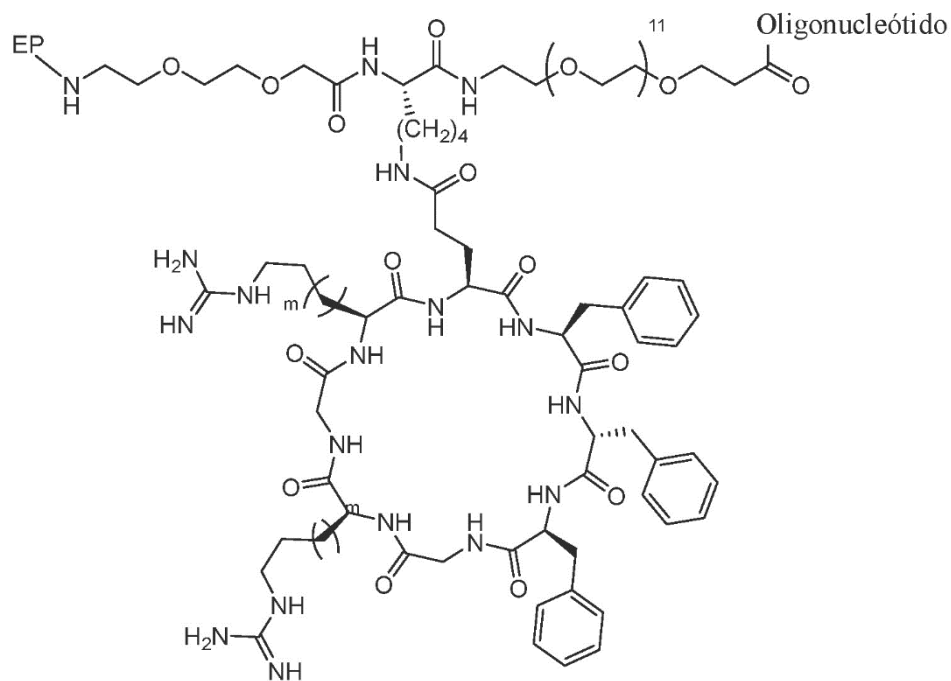
fórmula (C-c):



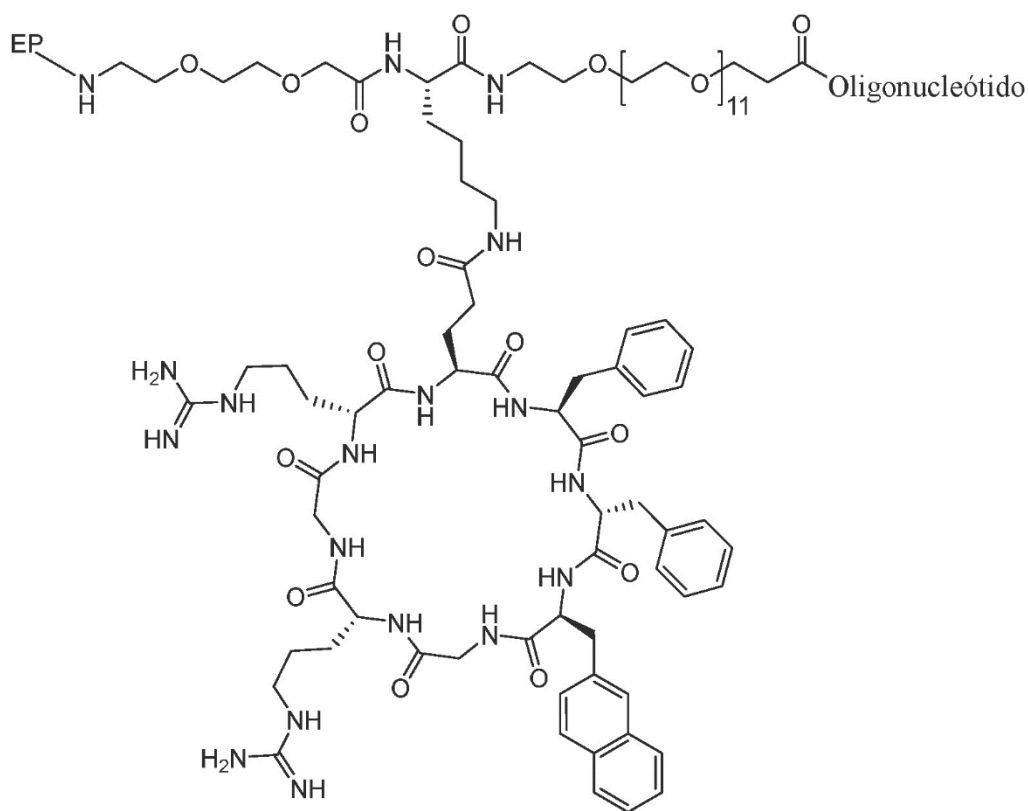
o una forma protonada de esta, en donde EP, R¹, R², R³, R⁴ y m son como se definió anteriormente en la fórmula (III); AA puede ser un aminoácido como se define en la presente descripción; n puede ser un número entero de 0-2; x puede ser un número entero de 1-10; y puede ser un número entero de 1-5; y z puede ser un número entero de 1-10.

[0001] El EEV puede conjugarse a una carga oligonucleotídica y el conjugado EEV-oligonucleótido puede comprender una estructura de fórmula (C-1), (C-2), (C-3) o (C-4):



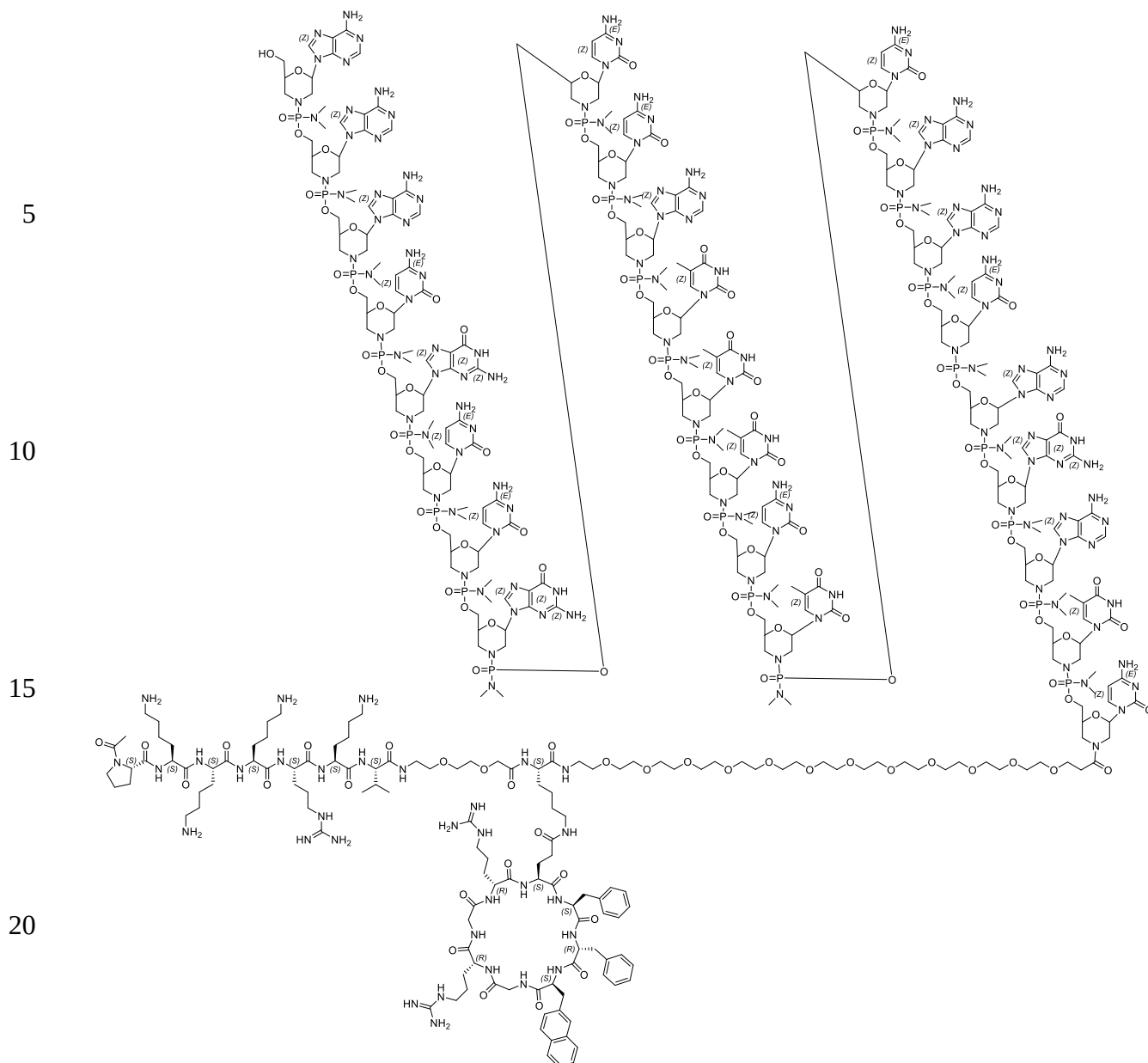


(C-3),



(C-4)

El EEV puede conjugarse a una carga oligonucleotídica y el conjugado de EEV puede comprender la estructura:



Eficiencia del suministro citosólico

25 Las modificaciones a un péptido penetrante de células cíclico (cCPP) pueden mejorar la eficiencia del suministro citosólico. La mejora de la eficiencia de la captación citosólica puede medirse al comparar la eficiencia del suministro citosólico de un cCPP que tiene una secuencia modificada con una secuencia control. La secuencia control no incluye un residuo de aminoácido de reemplazo particular en la secuencia modificada (lo que incluye, pero no se limita a, arginina, fenilalanina y/o

30 glicina), pero es de cualquier otra manera idéntica.

Como se usa en la presente, la eficiencia del suministro citosólico se refiere a la capacidad de un cCPP para atravesar una membrana celular y entrar al citosol de una célula. La eficiencia del

suministro citosólico del cCPP no es dependiente necesariamente de un receptor o un tipo de célula. La eficiencia del suministro citosólico puede referirse a la eficiencia absoluta del suministro citosólico o a la eficiencia relativa del suministro citosólico.

La eficiencia absoluta del suministro citosólico es la relación de la concentración citosólica de un cCPP (o un conjugado cCPP-carga) sobre la concentración del cCPP (o el conjugado cCPP-carga) en el medio de crecimiento. La eficiencia relativa del suministro citosólico se refiere a la concentración de un cCPP en el citosol en comparación con la concentración de un cCPP de control en el citosol. La cuantificación puede lograrse mediante el marcaje fluorescente del cCPP (por ejemplo, con un colorante FITC) y la medición de la intensidad de fluorescencia mediante el uso de técnicas bien conocidas en la técnica.

La eficiencia relativa del suministro citosólico se determina al comparar (i) la cantidad de un cCPP de la invención internalizado por un tipo de célula (por ejemplo, células HeLa) con (ii) la cantidad de un cCPP de control internalizado por el mismo tipo de célula. Para medir la eficiencia relativa del suministro citosólico, el tipo de célula puede incubarse en presencia de un cCPP durante un período de tiempo específico (por ejemplo, 30 minutos, 1 hora, 2 horas, etc.) después de lo cual la cantidad del cCPP internalizado por la célula se cuantifica mediante el uso de métodos conocidos en la técnica, por ejemplo, microscopía de fluorescencia. Por separado, la misma concentración del cCPP de control se incuba en presencia del tipo de célula durante el mismo período de tiempo, y se cuantifica la cantidad del cCPP de control internalizado por la célula.

La eficiencia relativa del suministro citosólico puede determinarse al medir la IC_{50} de un cCPP que tiene una secuencia modificada para un objetivo intracelular y comparar la IC_{50} del cCPP que tiene la secuencia modificada con una secuencia de control (como se describe en la presente descripción). La eficiencia relativa del suministro citosólico de los cCPP puede estar en el intervalo de aproximadamente 50 % a aproximadamente 450 % en comparación con ciclo(FfΦRrRrQ, SEQ ID NO:150), por ejemplo, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente 90 %, aproximadamente 100 %, aproximadamente 110 %, aproximadamente 120 %, aproximadamente 130 %, aproximadamente 140 %, aproximadamente 150 %, aproximadamente 160 %, aproximadamente 170 %, aproximadamente 180 %, aproximadamente 190 %, aproximadamente 200 %, aproximadamente 210 %, aproximadamente 220 %, aproximadamente 230 %, aproximadamente 240 %, aproximadamente 250 %, aproximadamente 260 %, aproximadamente 270 %, aproximadamente 280 %, aproximadamente 290 %, aproximadamente 300 %, aproximadamente 310 %, aproximadamente 320 %, aproximadamente 330 %, aproximadamente 340

%, aproximadamente 350 %, aproximadamente 360 %, aproximadamente 370 %, aproximadamente 380 %, aproximadamente 390 %, aproximadamente 400 %, aproximadamente 410 %, aproximadamente 420 %, aproximadamente 430 %, aproximadamente 440 %, aproximadamente 450 %, aproximadamente 460 %, aproximadamente 470 %, aproximadamente 480 %, aproximadamente 490 %, aproximadamente 500 %, aproximadamente 510 %, aproximadamente 520 %, aproximadamente 530 %, aproximadamente 540 %, aproximadamente 550 %, aproximadamente 560 %, aproximadamente 570 %, aproximadamente 580 %, o aproximadamente 590 %, lo que incluye todos los valores y subintervalos intermedios. La eficiencia relativa del suministro citosólico de los cCPP puede mejorarse en más de aproximadamente 600 % en comparación con un péptido cíclico que comprende ciclo(FfΦRrRrQ, SEQ ID NO:150).

La eficacia absoluta del suministro citosólico de aproximadamente 40 % a aproximadamente 100 %, por ejemplo, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 91 %, aproximadamente 92 %, aproximadamente 93 %, aproximadamente 94 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, lo que incluye todos los valores y subintervalos intermedios.

Los cCPP de la presente descripción pueden mejorar la eficiencia del suministro citosólico en aproximadamente 1.1 veces a aproximadamente 30 veces, en comparación con una secuencia de cualquier otra manera idéntica, por ejemplo, aproximadamente 1.2, aproximadamente 1.3, aproximadamente 1.4, aproximadamente 1.5, aproximadamente 1.6, aproximadamente 1.7, aproximadamente 1.8, aproximadamente 1.9, aproximadamente 2.0, aproximadamente 2.5, aproximadamente 3.0, aproximadamente 3.5, aproximadamente 4.0, aproximadamente 4.5, aproximadamente 5.0, aproximadamente 5.5, aproximadamente 6.0, aproximadamente 6.5, aproximadamente 7.0, aproximadamente 7.5, aproximadamente 8.0, aproximadamente 8.5, aproximadamente 9.0, aproximadamente 10, aproximadamente 10.5, aproximadamente 11.0, aproximadamente 11.5, aproximadamente 12.0, aproximadamente 12.5, aproximadamente 13.0, aproximadamente 13.5, aproximadamente 14.0, aproximadamente 14.5, aproximadamente 15.0, aproximadamente 15.5, aproximadamente 16.0, aproximadamente 16.5, aproximadamente 17.0, aproximadamente 17.5, aproximadamente 18.0, aproximadamente 18.5, aproximadamente 19.0, aproximadamente 19.5, aproximadamente 20, aproximadamente 20.5, aproximadamente 21.0, aproximadamente 21.5, aproximadamente 22.0, aproximadamente 22.5, aproximadamente 23.0,

aproximadamente 23.5, aproximadamente 24.0, aproximadamente 24.5, aproximadamente 25.0, aproximadamente 25.5, aproximadamente 26.0, aproximadamente 26.5, aproximadamente 27.0, aproximadamente 27.5, aproximadamente 28.0, aproximadamente 28.5, aproximadamente 29.0, o aproximadamente 29.5 veces, lo que incluye todos los valores y subintervalos intermedios.

5 Resto detectable

En modalidades, el compuesto descrito en la presente descripción incluye un resto detectable. En modalidades, el resto detectable se une al péptido penetrante de células en el grupo amino, el grupo carboxilato o la cadena lateral de cualquiera de los aminoácidos del resto de péptido penetrante de células (por ejemplo, en el grupo amino, el grupo carboxilato o la cadena lateral de cualquier aminoácido en el CPP). En modalidades, el resto terapéutico incluye un resto detectable. El resto detectable puede incluir cualquier marcador detectable. Los ejemplos de marcadores detectables adecuados incluyen, pero no se limitan a, un marcador UV-Vis, un marcador del infrarrojo cercano, un grupo luminiscente, un grupo fosforescente, un marcador de resonancia de espín magnético, un fotosensibilizador, un resto fotoescindible, un centro quelante, un átomo pesado, un isótopo radiactivo, una etiqueta de resonancia de espín detectable con isótopo, un resto paramagnético, un cromóforo, o cualquiera de sus combinaciones. En modalidades, el marcador es detectable sin la adición de reactivos adicionales.

En modalidades, el resto detectable es un resto detectable biocompatible, de manera que los compuestos pueden ser adecuados para su uso en una variedad de aplicaciones biológicas. “Biocompatible” y “biológicamente compatible”, como se usa en la presente, se refiere generalmente a compuestos que, junto con cualquier metabolito o producto de degradación de estos, son generalmente no tóxicos para células y tejidos, y que no provocan ningún efecto adverso significativo para las células y tejidos cuando las células y tejidos se incuban (por ejemplo, se cultivan) en su presencia.

El resto detectable puede contener un luminóforo tal como un marcador fluorescente o un marcador del infrarrojo cercano. Los ejemplos de luminóforos adecuados incluyen, pero no se limitan a, porfirinas metálicas; benzoporfirinas; azabenzoporfirina; naftoporfirina; ftalocianina; hidrocarburos aromáticos policíclicos tales como perileno diimina, pirenos; colorantes azo; colorantes de xanteno; dipiorometano de boro, dipiorometano de aza-boro, colorantes de cianina, complejos metal-ligando tal como bipyridina, bipyridilos, fenantrolina, cumarina, y acetilacetatos de rutenio e iridio; acridina, derivados de oxazina tales como benzofenoxazina; aza-anuleno, escuaína; 8-hidroxiquinolina, polimetinas, nanopartículas que producen luminiscencia, tales como puntos

cuánticos, nanocristales; carboestirilo; complejo de terbio; fósforo inorgánico; ionóforos tales como éteres de corona afiliados o colorantes derivatizados; o sus combinaciones. Los ejemplos específicos de luminóforos adecuados incluyen, pero no se limitan a, Pd (II) octaetilporfirina; Pt (II)-octaetilporfirina; tetrafenilporfirina Pd (II); Pt (II) tetrafenilporfirina; Pd (II) meso-tetrafenilporfirina tetrabenzoporfina; Pt (II) meso-tetrafenilmetrilbenzoporfirina; Pd (II) octaetilporfirina cetona; Pt (II) octaetilporfirina cetona; Pd (II) mesotetra(pentafluorofenil)porfirina; Pt (II) mesotetra(pentafluorofenil) porfirina; Ru (II) tris(4,7-difenil-1,10-fenantrolina) (Ru (dpp)₃); Ru (II) tris(1,10-fenantrolina) (Ru(fen)₃), cloruro tris(2,2'-bipiridina)rutenio (II) hexahidrato (Ru(bpy)₃); eritrosina B; fluoresceína; isotiocianato de fluoresceína (FITC); eosina; iridio (III) ((N-metil-benzimidazol-2-il)-7-(dietilamino)-cumarina)); 169enzotiazol ((benzotiazol-2-il)-7-(dietilamino)-cumarina))-2-(acetilacetato); colorantes de colágeno; rojo fluorescente Macroflex; amarillo fluorescente Macrolex; rojo Texas; rodamina B; rodamina 6G; rodamina sulfuro; m-cresol; azul timol; azul de xilenol; rojo cresol; azul clorofenol; verde bromocresol; rojo bromocresol; azul de bromotimol; Cy2; un Cy3; un Cy5; un Cy5.5; Cy7; 4-nitrofenol; alizarina; fenolftaleína; o-cresolftaleína; rojo clorofenol; calmagita; bromo-xilenol; rojo fenol; rojo neutro; nitrazina; 3,4,5,6-tetrabromofenolftaleína; rojo congo; fluoresceína; eosina; 2',7'-diclorofluoresceína; 5(6)-carboxifluoresceína; carboxinaftofluoresceína; ácido 8-hidroxipireno-1,3,6-trisulfónico; seminaftorodaflúor; semi-naftofluoresceína; tris (4,7-difenil-1,10-fenantrolina) rutenio (II) dicloruro; (4,7-difenil-1,10-fenantrolina) rutenio (II) tetrafenilboro; platino (II) octaetilporfirina; dialquilcarbocianina; dioctadecilcicloxacarbocianina; cloruro de fluorenilmetiloxicarbonilo; 7-amino-4-metilcumarina (Amc); proteína fluorescente verde (GFP); y derivados o sus combinaciones. En algunos ejemplos, el resto detectable puede incluir rodamina B (Rho), isotiocianato de fluoresceína (FITC), 7-amino-4-metilcumarina (Amc), proteína fluorescente verde (GFP), o derivados o sus combinaciones.

Métodos de fabricación

Los compuestos descritos en la presente descripción pueden prepararse en una variedad de maneras conocidas por un experto en la técnica de la síntesis orgánica o variaciones en la misma como se aprecia por los expertos en la técnica. Los compuestos descritos en la presente descripción pueden prepararse a partir de materiales de partida fácilmente disponibles. Las condiciones de reacción óptimas pueden variar con los reactivos o solventes particulares usados, pero dichas condiciones pueden determinarse por un experto en la técnica.

Las variaciones en los compuestos descritos en la presente descripción incluyen la adición,

sustracción o movimiento de los diversos constituyentes como se describe para cada compuesto. De manera similar, cuando uno o más centros quirales están presentes en una molécula, puede cambiarse la quiralidad de la molécula. Adicionalmente, la síntesis de compuestos puede involucrar la protección y desprotección de diversos grupos químicos. El uso de la protección y desprotección, y la selección de grupos de protección apropiados pueden determinarse por un experto en la técnica. La química de los grupos de protección puede encontrarse, por ejemplo, en Wuts y Greene, *Protective Groups in Organic Synthesis*, 4a Ed., Wiley & Sons, 2006, que se incorpora en la presente descripción como referencia en su totalidad.

Los materiales de partida y reactivos usados en la preparación de los compuestos y composiciones descritos están disponibles de proveedores comerciales tales como Aldrich Chemical Co., (Milwaukee, WI), Acros Organics (Morris Plains, NJ), Fisher Scientific (Pittsburgh, PA), Sigma (St. Louis, MO), Pfizer (Nueva York, NY), GlaxoSmithKline (Raleigh, NC), Merck (Whitehouse Station, NJ), Johnson & Johnson (Nuevo Brunswick, NJ), Aventis (Bridgewater, NJ), AstraZeneca (Wilmington, DE), Novartis (Basilea, Suiza), Wyeth (Madison, NJ), Bristol-Myers-Squibb (Nueva York, NY), Roche (Basilea, Suiza), Lilly (Indianapolis, IN), Abbott (Abbott Park, IL), Schering Plough (Kenilworth, NJ) o Boehringer Ingelheim (Ingelheim, Alemania), o se preparan mediante métodos conocidos para los expertos en la técnica siguiendo los procedimientos expuestos en referencias tales como Fieser and Fieser's *Reagents for Organic Synthesis*, Volúmenes 1-17 (John Wiley and Sons, 1991); Rodd's *Chemistry of Carbon Compounds*, Volúmenes 1-5 y suplementos (Elsevier Science Publishers, 1989); *Organic Reactions*, Volúmenes 1-40 (John Wiley and Sons, 1991); *March's Advanced Organic Chemistry*, (John Wiley and Sons, 4a edición); y *Larock's Comprehensive Organic Transformations* (VCH Publishers Inc., 1989). Otros materiales, tales como los portadores farmacéuticos descritos en la presente descripción, pueden obtenerse de fuentes comerciales.

Las reacciones para producir los compuestos descritos en la presente descripción pueden llevarse a cabo en solventes, que pueden seleccionarse por un experto en la técnica de la síntesis orgánica. Los solventes pueden ser sustancialmente no reactivos con los materiales de partida (reactivos), los intermediarios, o productos bajo las condiciones en las que se llevan a cabo las reacciones, es decir, temperatura y presión. Las reacciones pueden llevarse a cabo en un solvente o una mezcla de más de un solvente. La formación del producto o intermediario puede monitorearse de acuerdo con cualquier método adecuado conocido en la técnica. Por ejemplo, la formación del producto puede monitorearse mediante medios espectroscópicos, tales como espectroscopía de resonancia magnética

nuclear (por ejemplo, ^1H o ^{13}C), espectroscopía infrarroja, espectrofotometría (por ejemplo, UV-visible) o espectrometría de masas, o mediante cromatografía tal como cromatografía líquida de alto rendimiento (HPLC) o cromatografía de capa fina.

Los compuestos descritos pueden prepararse mediante síntesis de péptidos en fase sólida, en donde el extremo α -N-terminal del aminoácido está protegido por un grupo de protección de ácido o base. Dichos grupos de protección deben tener las propiedades de ser estables a las condiciones de formación de los enlaces peptídicos, al tiempo que se puedan eliminar fácilmente sin la destrucción de la cadena peptídica creciente o la racemización de cualquiera de los centros quirales contenidos en la misma. Los grupos de protección adecuados son 9-fluorenilmetiloxycarbonilo (Fmoc), t-butiloxycarbonilo (Boc), benciloxycarbonilo (Cbz), bifenilisopropiloxycarbonilo, t-amiloxycarbonilo, isoborniloxycarbonilo, α,α -dimetil-3,5-dimetoxibenciloxycarbonilo, o-nitrofenilsulfenilo, 2-ciano-t-butiloxycarbonilo, y similares. Se prefiere particularmente el grupo de protección 9-fluorenilmetiloxycarbonilo (Fmoc) para la síntesis de los compuestos descritos. Otros grupos de protección de la cadena lateral preferidos son, para grupos amino de la cadena lateral como lisina y arginina, 2,2,5,7,8-pentametilcroman-6-sulfonilo (pmc), nitro, p-toluenosulfonilo, 4-metoxibenceno-sulfonilo, Cbz, Boc, y adamantiloxycarbonilo; para la tirosina, bencilo, o-bromobenciloxi-carbonilo, 2,6-diclorobencil, isopropilo, t-butilo (t-Bu), ciclohexilo, ciclopentilo y acetilo (Ac); para la serina, t-butilo, bencilo y tetrahidropiranilo; para la histidina, tritilo, bencilo, Cbz, p-toluenosulfonilo y 2,4-dinitrofenilo; para el triptófano, formilo; para el ácido aspártico y el ácido glutámico, bencilo y t-butilo y para la cisteína, trifenilmetilo (tritilo).

En el método de síntesis de péptidos en fase sólida, el aminoácido α -C-terminal se une a un soporte sólido o resina adecuados. Los soportes sólidos adecuados útiles para la síntesis anterior son aquellos materiales que son inertes a los reactivos y las condiciones de reacción de las reacciones de desprotección-condensación escalonadas, así como también que sean insolubles en el medio usado.

Los soportes sólidos para la síntesis de péptidos carboxi α -C-terminales son resina 4-hidroximetilfenoximetil-copoli(estireno-divinilbenceno al 1 %) o 4-(2',4'-dimetoxifenil-Fmoc-aminometil)fenoxiacetamidoetilo disponibles de Applied Biosystems (Foster City, Calif.). El aminoácido α -C-terminal se acopla a la resina por medio de N,N'-díciclohexilcarbodiimida (DCC), N,N'-diisopropilcarbodiimida (DIC) u O-benzotriazol-1-il-N,N,N',N'-tetrametiluroniohexafluorofosfato (HBTU), con o sin 4-dimetilaminopiridina (DMAP), 1-hidroxibenzotriazol (HOBT), cloruro de benzotriazol-1-iloxi-tris(dimetilamino)fosfoniohexafluorofosfato (BOP) o bis(2-oxo-3-oxazolidinil)fosfina (BOPCl),

acoplamiento mediado durante de aproximadamente 1 a aproximadamente 24 horas a una temperatura de entre 10 °C y 50 °C en un solvente tal como diclorometano o DMF. Cuando el soporte sólido es resina 4-(2',4'-dimetoxifenil-Fmoc-aminometil)fenoxi-acetamidoetilo, el grupo Fmoc se escinde con una amina secundaria, preferentemente piperidina, antes del acoplamiento con el aminoácido α -C-terminal como se describió anteriormente. Un método para el acoplamiento a la resina 4-(2',4'-dimetoxifenil-Fmoc-aminometil)fenoxi-acetamidoetilo desprotegida es O-benzotriazol-1-il-N,N,N',N'-tetrametiluroniohexafluorofosfato (HBTU, 1 equiv.) y 1-hidroxibenzotriazol (HOBT, 1 equiv.) en DMF. El acoplamiento de aminoácidos protegidos sucesivos puede llevarse a cabo en un sintetizador de polipéptidos automático. En un ejemplo, el extremo α -N-terminal en los aminoácidos de la cadena peptídica en crecimiento se protege con Fmoc. La eliminación del grupo de protección Fmoc del lado α -N-terminal del péptido en crecimiento se logra mediante el tratamiento con una amina secundaria, preferentemente, piperidina. Cada aminoácido protegido se introduce a continuación en aproximadamente 3 veces el exceso molar, y el acoplamiento se lleva a cabo, preferentemente, en DMF. El agente de acoplamiento puede ser O-benzotriazol-1-il-N,N,N',N'-tetrametiluroniohexafluorofosfato (HBTU, 1 equiv.) y 1-hidroxibenzotriazol (HOBT, 1 equiv.). Al final de la síntesis en fase sólida, el polipéptido se retira de la resina y se desprotege, ya sea sucesivamente o en una única operación. La eliminación del polipéptido y la desprotección pueden llevarse a cabo en una única operación mediante el tratamiento del polipéptido unido a la resina con un reactivo de escisión que comprende tianisol, agua, etanoditiol y ácido trifluoroacético. En los casos en donde el α -C-terminal del polipéptido es una alquilamida, la resina se escinde mediante aminólisis con una alquilamina. Alternativamente, el péptido puede eliminarse por transesterificación, por ejemplo, con metanol, seguido de aminólisis o transamidación directa. El péptido protegido puede purificarse en este punto o llevarse a la siguiente etapa directamente. La eliminación de los grupos de protección de la cadena lateral puede lograrse mediante el uso del cóctel de escisión descrito anteriormente. El péptido completamente desprotegido puede purificarse mediante una secuencia de etapas cromatográficas que emplea cualquiera o todos los tipos siguientes: intercambio iónico en una resina débilmente básica (forma de acetato); cromatografía de adsorción hidrófoba en poliestireno-divinilbenceno no derivatizado (por ejemplo, Amberlite XAD); cromatografía de adsorción en gel de sílice; cromatografía de intercambio iónico en carboximetilcelulosa; cromatografía de reparto, por ejemplo, en Sephadex G-25, LH-20 o distribución contracorriente; cromatografía líquida de alto rendimiento (HPLC), especialmente HPLC de fase inversa en empaques de columna de fase unida de sílice unido a octil u

octadecilsililo.

Los polímeros anteriores, tales como los grupos PEG, pueden unirse a un oligonucleótido, tal como un AC, bajo cualquier condición adecuada. Puede usarse cualquier medio conocido en la técnica, lo que incluye mediante acilación, alquilación reductiva, adición de Michael, alquilación de tiol u otros métodos de conjugación/ligación quimiosselectiva a través de un grupo reactivo en el resto de PEG (por ejemplo, un grupo aldehído, amino, éster, tiol, α -haloacetilo, maleimido o hidrazino) a un grupo reactivo en el AC (por ejemplo, un grupo aldehído, amino, éster, tiol, α -haloacetilo, maleimido o hidrazino). Los grupos activadores que pueden usarse para enlazar el polímero soluble en agua a una o más proteínas incluyen, sin limitación, grupos sulfona, maleimida, sulfhidrilo, tiol, triflato, tresilato, azidirina, oxirano, 5-piridilo y acilo alfa-halogenado (por ejemplo, ácido α -yodo acético, ácido α -bromoacético, ácido α -cloroacético). Si se une al AC mediante alquilación reductiva, el polímero seleccionado debe tener un único aldehído reactivo de manera que se controle el grado de polimerización. Ver, por ejemplo, Kinstler y otros, *Adv. Drug. Delivery Rev.* (2002), 54: 477-485; Roberts y otros, *Adv. Drug Delivery Rev.* (2002), 54: 459-476; y Zalipsky y otros, *Adv. Drug Delivery Rev.* (1995), 16: 157-182.

Para enlazar covalentemente de manera directa el AC o enlazador al CPP, los residuos de aminoácidos apropiados del CPP pueden hacerse reaccionar con un agente derivatizante orgánico que es capaz de reaccionar con una cadena lateral seleccionada o los extremos N o C terminales de un aminoácido. Los grupos reactivos en el resto de péptido o conjugado incluyen, por ejemplo, un grupo aldehído, amino, éster, tiol, α -haloacetilo, maleimido o hidrazino. Los agentes derivatizantes incluyen, por ejemplo, éster de maleimidobenzoyl sulfosuccinimida (conjugación a través de residuos de cisteína), N-hidroxisuccinimida (a través de residuos de lisina), glutaraldehído, anhídrido succínico u otros agentes conocidos en la técnica.

Los métodos para fabricar AC y conjugar el AC al CPP lineal se describen generalmente en la publicación de Estados Unidos núm. 2018/0298383, que se incorpora en la presente descripción como referencia para todos los propósitos. Los métodos pueden aplicarse a los CPP cíclicos descritos en la presente descripción.

Los esquemas sintéticos se proporcionan en la Figura 3A-3D y la Figura 4.

Ejemplos no limitantes de compuestos que incluyen un CPP y un grupo reactivo útil para la conjugación a un AC se muestran en la Tabla 9. También se muestran ejemplos de grupos enlazadores. Los grupos reactivos ilustrativos incluyen éster tetrafluorofenílico (TFP), ácido

carboxílico libre (COOH) y azida (N₃). En la Tabla 9, n es un entero de 0 a 20; Pipa6 es AcRXRRBRRXRYQFLIRXRBRXRB, en donde B es β-alanina y X es ácido aminohexanoico; Dap es ácido 2,3-diaminopropiónico; NLS es una secuencia de localización nuclear; βA es beta alanina; -ss- es un disulfuro; PABC es un dominio C-terminal de proteína de unión a poli(A); C_x donde x es un número es una cadena de alquilo de longitud x; y BCN es biciclo [6.1.0]nonino.

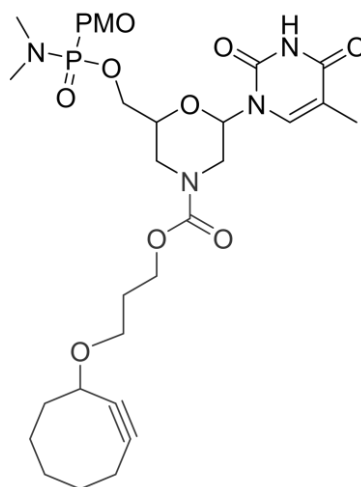
Tabla 9: Compuestos que incluyen un CPP y un grupo reactivo

	TFP-PEG _n -K(CPP)
	TFP-PEG _n -K(CPP)-PEG _n -Dap(palmitoil)
	TFP-PEG _n -K(CPP)-PEG _n -Dap(CPP)
10	TFP-Pip6a
	CPP-PEG _n -TFP
	CPP-PEG _n -K(CPP)-PEG _n -TFP
	CPP-PEG _n -Lys(N ₃)
	CPP-K(CPP)-PEG _n -K(N ₃)
	CPP-PEG _n -K(PEG _n -CPP)-PEG _n -K(N ₃)
15	CPP-PEG _n -K(PEG _n -CPP)-PEG _n -K(N ₃)
	CPP-K(CPP)-K(CPP)-PEG _n -K(N ₃)
	CPP-PEG _n -K(PEG _n -CPP)-K(PEG _n -CPP)-PEG _n -K(N ₃)
	CPP-PEG _n -K(PEG _n -CPP)-K(PEG _n -CPP)-PEG _n -K(N ₃)
	Ac-NLS-Lys(CPP)-PEG _n -K(N ₃)
	K(N ₃)-PEG _n -NLS-ss-PEG _n -CPP
20	BCN-NLS-ss-CPP
	CPP-PEG _n -Val-Cit-PABC-K(N ₃)
	CPP-PEG _n -Cys-ss-Cys-K(N ₃)
	CPP-PEG _n -Cys-ss-Cys-K(N ₃)
	CPP-PEG _n -TFP
	CPP-PEG _n -Lys(N ₃)
	CPP-PEG _n -Cys-prodisulfuro-K(N ₃)
25	CPP-PEG _n -K(N ₃)
	CPP-K(CPP)-PEG _n -K(N ₃)
	CPP-PEG _n -K(CPP)-PEG _n -TFP
	CPP-C ₆ -TFP
	CPP-PEG _n -K(PEG _n -CPP)-PEG _n -K(N ₃)
	Ac-T9-PEG _n -Lys(CPP-PEG _n)-K(N ₃)
	Ac-MSP-PEG _n -K(CPP-PEG _n)-K(N ₃)
30	CPP-PEG _n -TFP (ENTRD 802)
	CPP-C ₆ -TFP (ENTRD 696)
	CPP-PEG _n -K(CPP)-PEG _n -TFP (ENTRD-344)

CPP-PEG _n -COOH
CPP-C ₁₂ -TFP (ENTD-695)
palmitoil-PEG _n -K(CPP)-PEG _n -TFP (ENTD-343)
CPP-PEG _n -K(N ₃) (ENTRD-617)
Ac-T9-PEG _n -K(CPP)-K(N ₃) (ENTRD 673)
Ac-MSP-PEG _n -K(CPP-PEG _n)-K(N ₃) (ENTRD 675)
Ac-NLS-K(CPP)-PEG _n -K(N ₃) (ENTRD 684)
K(N ₃)-PEG _n -NLS-ss-PEG _n -CPP (ETRD-681)
K(N ₃)-PEG _n -NLS-K-βA-βA-CPP (ETRD-682)

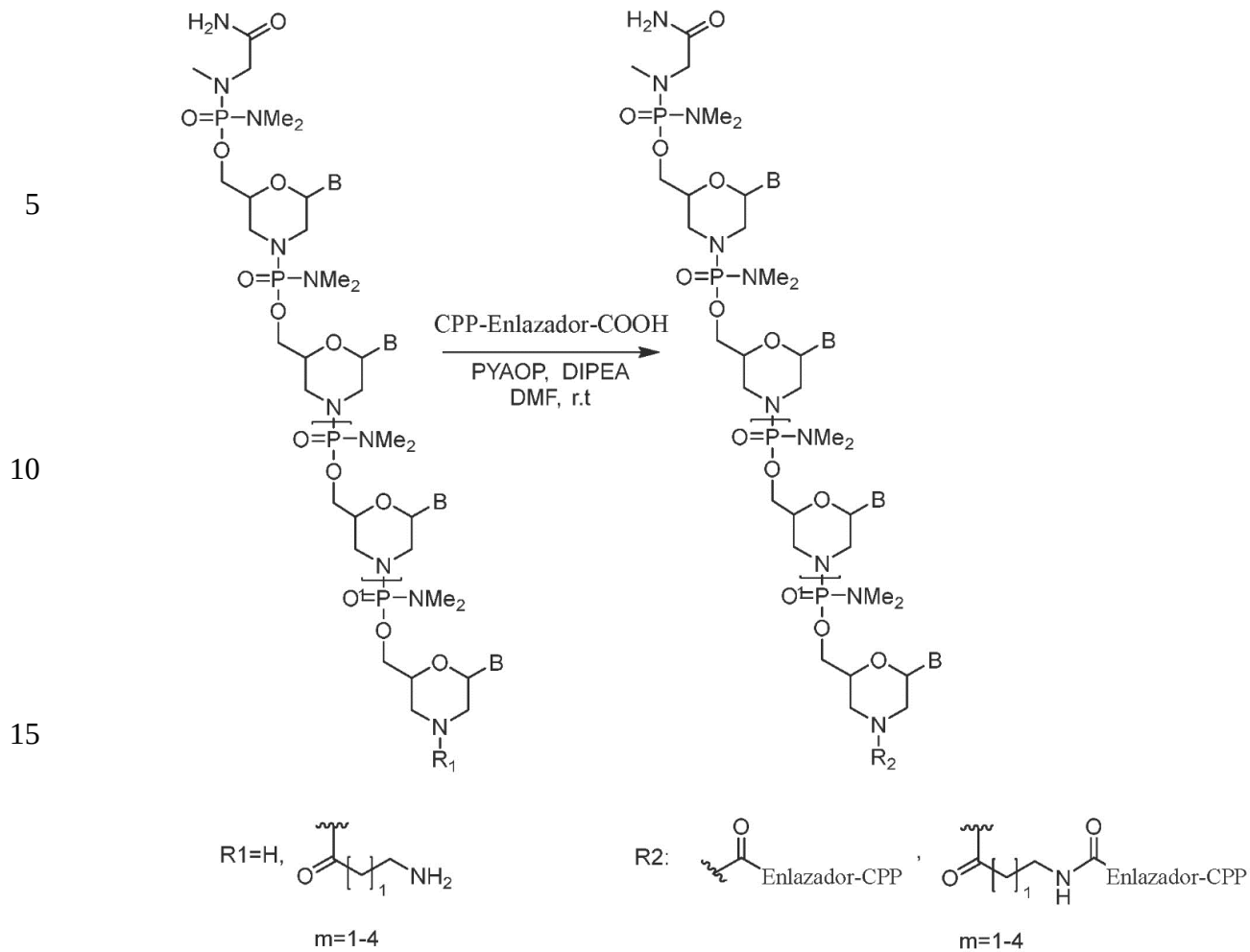
En modalidades, los CPP tienen grupos de ácido carboxílico libres que pueden utilizarse para la conjugación a un AC. En modalidades, los EEV tienen grupos de ácido carboxílico libres que pueden utilizarse para la conjugación a un AC.

La estructura más abajo es un PMO modificado con ciclooctino 3' usado para una reacción clic con



un compuesto que incluye una azida:

Más abajo se muestra un esquema ilustrativo de la conjugación de un CPP y un enlazador al extremo 3' de un AC mediante un enlace amida.

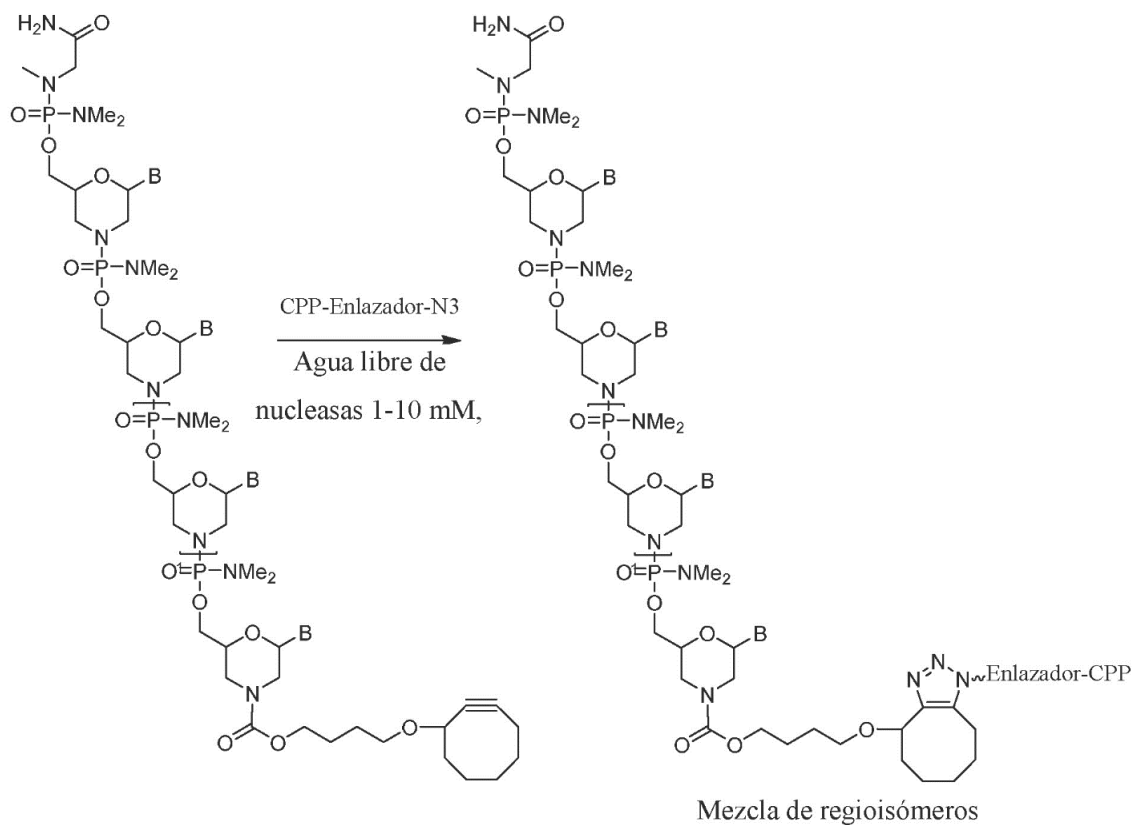


Más abajo se muestra un esquema ilustrativo de la conjugación de un CPP y un enlazador a un PMO modificado con ciclooctino 3' mediante una cicloadición azida-alquino promovida por tensión:

5

10

15



Más abajo se muestra un ejemplo de la química de conjugación usada para conectar un AC y un CPP con un enlazador adicional que contiene un resto de polietilenglicol:

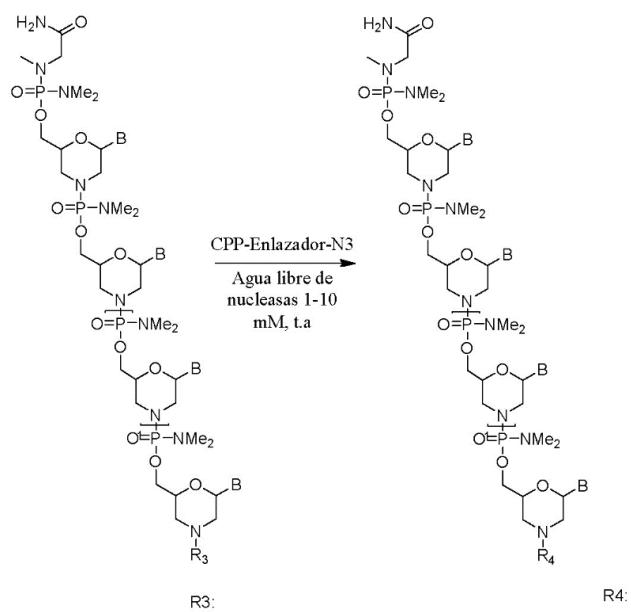
20

25

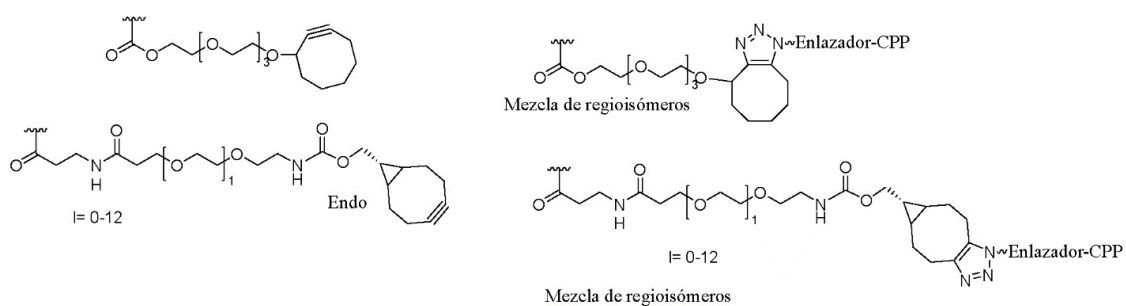
30

5

10



15



20

Más abajo se muestra un ejemplo de la conjugación de un CPP-enlazador a un PMO modificado con ciclooctino 5' mediante una cicloadición azida-alquino promovida por tensión (química clic):

25

30

5

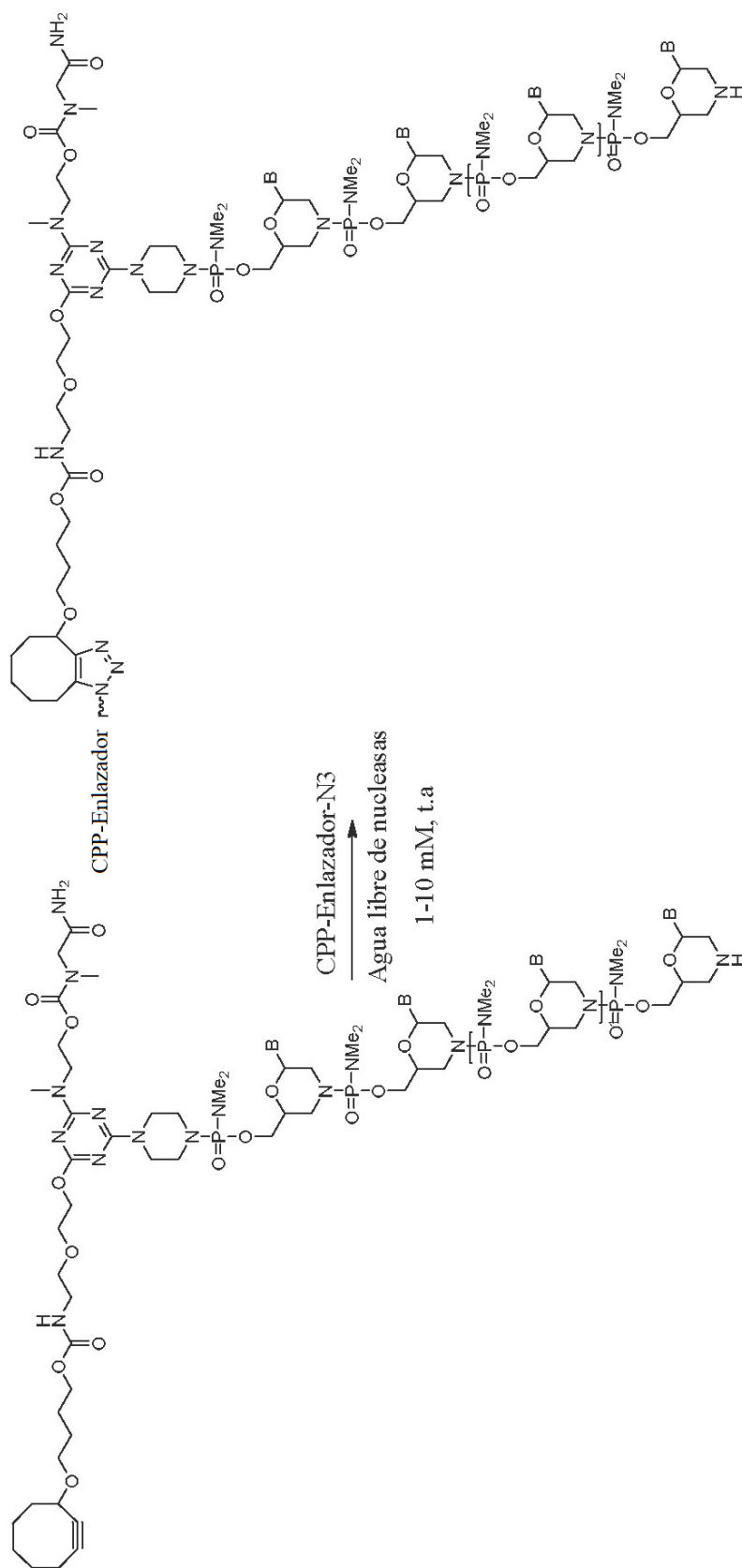
10

15

20

25

30



5

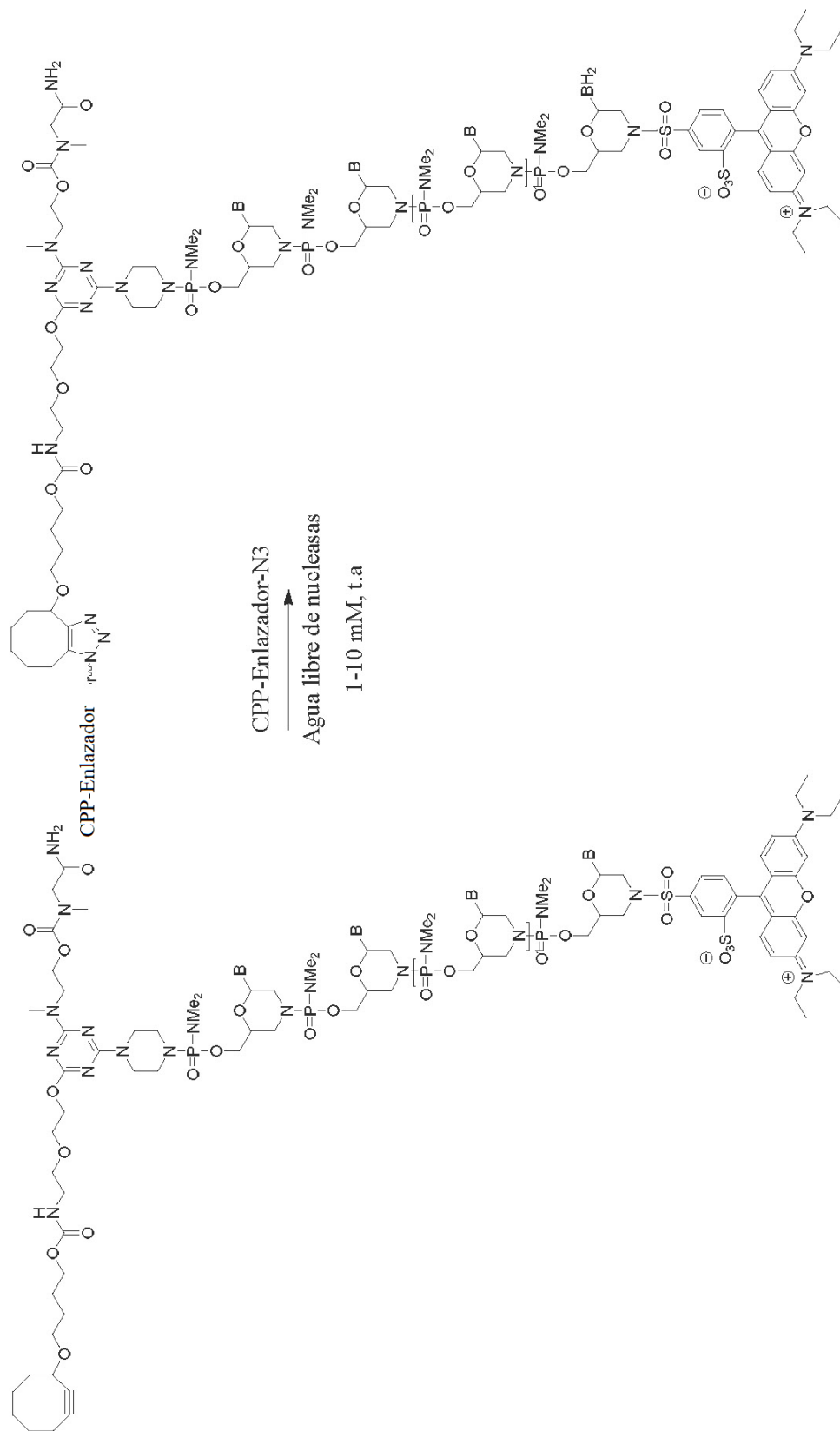
10

15

20

25

30



Los métodos para sintetizar compuestos antisentido oligoméricos se conocen en la técnica. La presente descripción no está limitada por el método de síntesis del AC. En modalidades, en la presente descripción se proporcionan compuestos que tienen grupos de fósforo reactivos útiles para formar enlaces internucleósidos, lo que incluye, por ejemplo, enlaces internucleósidos fosfodiéster y fosforotioato. Los métodos de preparación y/o purificación de precursores o compuestos antisentido no son una limitación de las composiciones o métodos proporcionados en la presente descripción. Los métodos para la síntesis y purificación de ADN, ARN y los compuestos antisentido se conocen bien por los expertos en la técnica.

La oligomerización de nucleósidos modificados y no modificados puede realizarse rutinariamente de acuerdo con los procedimientos de la literatura para el ADN (Protocols for Oligonucleotides and Analogs, Ed. Agrawal (1993), Humana Press) y/o el ARN (Scaringe, Methods (2001), 23, 206-217. Gait y otros, Applications of Chemically synthesized RNA in RNA: Protein Interactions, Ed. Smith (1998), 1-36. Gallo y otros, Tetrahedron (2001), 57, 5,707-5,713).

Los compuestos antisentido proporcionados en la presente descripción pueden fabricarse convenientemente y de manera rutinaria mediante la técnica bien conocida de síntesis en fase sólida. El equipo para dicha síntesis se vende por varios proveedores, lo que incluye, por ejemplo, Applied Biosystems (Foster City, CA). Puede emplearse adicional o alternativamente cualquier otro medio para dicha síntesis conocido en la técnica. Se conoce bien el uso de técnicas similares para preparar oligonucleótidos tales como los fosforotioatos y derivados alquilados. La invención no está limitada por el método de síntesis de los compuestos antisentido.

Los métodos de purificación y análisis de oligonucleótidos son conocidos para los expertos en la técnica. Los métodos de análisis incluyen electroforesis capilar (CE) y espectroscopía de masas-electronebulización. Dichos métodos de síntesis y análisis pueden realizarse en placas de múltiples pocillos. El método de la invención no está limitado por el método de purificación del oligómero.

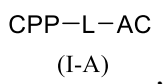
En los compuestos descritos en la presente descripción, el AC se acopla al CPP (por ejemplo, péptido cíclico). Como se usa en la presente, “acoplado” puede referirse a una asociación covalente o no covalente entre el CPP y el AC, lo que incluye la fusión del CPP al AC y la conjugación química del CPP (por ejemplo, péptido cíclico) al AC. Un ejemplo no limitante de un medio para unir no covalentemente el CPP al AC es a través de la interacción estreptavidina/biotina, por ejemplo, mediante la conjugación de la biotina al CPP y la fusión del AC a la estreptavidina. En el compuesto resultante, el CPP se acopla al AC mediante asociación no covalente entre la biotina y la estreptavidina.

En modalidades, el CPP (por ejemplo, péptido cíclico) se conjuga, directa o indirectamente, al AC para formar de esta manera un conjugado CPP-AC. La conjugación del AC al CPP puede producirse en cualquier sitio apropiado en estos restos. Por ejemplo, en modalidades, el extremo 5' o 3' del AC puede conjugarse con el extremo C-terminal, el extremo N-terminal o una cadena lateral de un aminoácido en el CPP.

En modalidades, el AC se enlaza covalentemente al CPP (por ejemplo, péptido cíclico). El enlace covalente, como se usa en la presente, se refiere a constructos donde un resto de CPP se une covalentemente al extremo 5' y/o 3' del resto de AC. Dichos conjugados pueden describirse alternativamente como que tienen un resto de CPP (por ejemplo, un resto de péptido cíclico) y un resto de oligonucleótido. Un conjugado de AC-CPP o CPP-AC enlazado covalentemente, de acuerdo con ciertas modalidades, incluye el componente AC y el componente CPP asociados entre sí mediante un enlazador descrito en la presente descripción.

En modalidades, el AC puede conjugarse al CPP (por ejemplo, péptido cíclico) a través de una cadena lateral de un aminoácido en el CPP. Cualquier cadena lateral de aminoácido en el CPP que sea capaz de formar un enlace covalente, o que pueda modificarse de este modo, puede usarse para enlazar el AC al CPP. El aminoácido en el CPP puede ser un aminoácido natural o no natural. En modalidades, el aminoácido en el CPP usado para conjugar el AC es ácido aspártico, ácido glutámico, glutamina, asparagina, lisina, ornitina, ácido 2,3-diaminopropiónico, o análogos de estos, en donde la cadena lateral se sustituye con un enlace al AC o al enlazador. En modalidades, el aminoácido es lisina, o un análogo de esta. En modalidades, el aminoácido es ácido glutámico, o un análogo de este. En modalidades, el aminoácido es ácido aspártico, o un análogo de este.

En modalidades, el CPP es cíclico. Existen numerosas configuraciones posibles para los compuestos descritos en la presente descripción. En modalidades, los compuestos de la descripción incluyen compuestos en donde el AC se conjuga con la cadena lateral de un aminoácido en el péptido cíclico. En modalidades, los compuestos descritos en la presente descripción tienen una estructura (es decir, exocíclica) de acuerdo con la fórmula I-A:



en donde el enlazador se une covalentemente a la cadena lateral de un aminoácido en el CPP y al extremo 5' del AC, la cadena principal del AC o el extremo 3' del AC.

Enfermedades y genes objetivo

En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno

asociado con uno o más genes que tienen una repetición de nucleótidos expandida (por ejemplo, repeticiones de trinucleótidos expandidas tales como repeticiones de trinucleótidos expandidas). En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno asociado con uno o más genes que tienen una repetición de trinucleótido CTG·CUG expandida. En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno asociado con uno o más genes que tienen una repetición de trinucleótidos CTG·CUG expandida en la UTR 3' del gen. En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno asociado con un gen que tiene un CTG·CUG expandido en la UTR 3' tal como DMPK, ATXN8OS ATXN8 y/o JPH3. En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno asociado con uno o más genes que tienen una repetición de trinucleótidos CTG·CUG expandida en el intrón de un gen. En modalidades, se proporcionan compuestos y métodos para tratar una enfermedad o trastorno asociado con una repetición de trinucleótidos CTG·CUG expandida en un intrón de TCF4. En modalidades, se proporcionan compuestos y métodos para tratar la distrofia miotónica tipo 1 (DM1), la distrofia endotelial corneal de Fuchs (FECF), la ataxia espinocerebelosa-8 (SCA8) y la enfermedad similar a Huntington (HDL2).

Distrofia miotónica tipo 1 (DM1)

En modalidades, se proporcionan compuestos, composiciones y métodos para tratar la distrofia miotónica (DM1 o enfermedad de Steinert). La DM1 es un trastorno multisistémico que se caracteriza con frecuencia por degeneración muscular y miotonía o relajación muscular retardada debido a potenciales de acción repetitivos en las miofibras. La distrofia miotónica tipo 1 (DM1) es la forma más común de distrofia muscular, afectando aproximadamente 1 de cada 8,000 personas. La DM1 es un paradigma de los trastornos genéticos provocados por las expansiones de CTG·CUG. La DM1 es un trastorno neuromuscular provocado por una expansión de la repetición CTG·CUG en la región no traducida (UTR) 3' del gen de la proteína cinasa de la distrofia miotónica (DMPK). A nivel del ARN, el transcrito de DMPK (por ejemplo, la repetición CUG expandida) secuestra proteínas reguladoras del corte y empalme, por ejemplo, la proteína de tipo muscleblind (MBNL), lo que da como resultado un empalme erróneo de un número de pre-ARNm aguas abajo (pre-ARNm que no contienen una repetición CUG expandida) que están regulados por MBNL1. Esta ganancia de función es la causa de la DM1.

El número excesivo de repeticiones CUG imparte actividad tóxica, denominada ganancia de función tóxica. Múltiples proteínas clave se procesan erróneamente y esto contribuye a la naturaleza

multisistémica de la enfermedad, lo que incluye debilidad generalizada de las extremidades, disfunción muscular respiratoria, anomalías cardíacas, fatiga, complicaciones gastrointestinales, cataratas, incontinencia y somnolencia diurna excesiva.

Los pacientes con DM1 con expansiones CTG·CUG dentro de la región no traducida 3' del gen DMPK tienen un mayor riesgo de FECD y forman focos CUGexp-MBNL1 en el endotelio corneal. (Mootha y otros, *Investigative ophthalmology & visual science*, 2017; 58, 4,579-4,585). La asociación de MBNL1 con el ARN mutante afecta el conglomerado celular de MBNL1 libre y desencadena el empalme erróneo de algunos genes objetivo de MBNL1 (por ejemplo, regulados por MBNL1) en el cerebro, los músculos y los tejidos cardíacos afectados (Jiang y otros, *Hum Mol Genet*. 2004; 13: 3,079–3,088). Gattey y otros (*Cornea*. 2014; 33: 96–98) informaron la FECD en cuatro sujetos con DM1, lo que incluye un par madre-hija. Por tanto, es probable que esté presente la asociación entre la DM1 y la FECD (FECD se describe con más detalle en otra parte en la presente descripción).

Sin estar limitado por la teoría, existen al menos dos hipótesis propuestas para explicar la patogénesis de la DM1. Una es que las repeticiones CTG·CUG expandidas inhiben el ARNm de DMPK o la producción de proteínas, lo que da como resultado una haploinsuficiencia de DMPK. Esto estaba respaldado por estudios que demostraron una disminución de la expresión del ARNm y la proteína de la DMPK en el músculo con DM1 (Fu, Y.H. y otros (1993) Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235–238). En modalidades, los compuestos y métodos descritos en la presente descripción mejoran la haploinsuficiencia de DMPK. Otra hipótesis de la ganancia de función del ARN propone que el ARN mutante transcrito a partir del alelo expandido es suficiente para inducir síntomas de la enfermedad. Esto se sugirió mediante observaciones: (i) las repeticiones CTG expandidas se transcriben en repeticiones CUG que se acumulan en focos nucleares discretos, (ii) la expresión de solo la UTR 3' de DMPK con 200 repeticiones CTG es suficiente para inhibir la miogénesis (Davis, B.M., y otros (1997) Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7,388–7,393; Amack, J.D. y otros, (1999) Cis and trans effects of the myotonic dystrophy (DM) mutation in a cell culture model. *Hum. Mol. Genet.* 8, 1975–1984). En modalidades, los compuestos y métodos descritos en la presente descripción reducen la transcripción del ARN mutante que se asocia con el alelo expandido.

Las repeticiones de trinucleótidos CTG·CUG expandidas en la región no traducida 3' del ARNm de

DMPK forma estructuras en horquilla estables imperfectas que se acumulan en el núcleo celular en complejos ribonucleares pequeños o inclusiones microscópicamente visibles, y perjudican la función de las proteínas involucradas en la transcripción, corte y empalme o exportación de ARN. Aunque los genes DMPK con repeticiones CUG se transcriben en ARNm, los transcritos mutantes se secuestran en el núcleo como agregados (focos), lo que da como resultado una disminución en los niveles de ARNm de DMPK citoplasmáticos. Estas agregaciones conducen a la disregulación del corte y empalme alternativo de muchos transcritos diferentes debido al secuestro de dos proteínas de unión al ARN: MBNL1 (de tipo muscleblind 1) y CUGBP1 (proteína de unión a CUG 1), lo que da como resultado la pérdida de función de MBNL1 y la regulación positiva de CUGBP1 (Lee y Cooper. (2009) "Pathogenic mechanisms of myotonic dystrophy", *Biochem Soc Trans.* 37(06): 1281-1286).

En la DM1, la proteína de unión al ARN MBNL1 se secuestra a la estructura de horquilla bicatenaria formada por repeticiones CUG, agotándola del nucleoplasma. A continuación, las repeticiones CUG a las cuales se une la MBNL1 estimulan la activación de la proteína cinasa C (PKC) a través de un mecanismo desconocido, que induce la hiperfosforilación y estabilización de CUGBP1. Los efectos aguas abajo incluyen la alteración del corte y empalme alternativo, la traducción del ARNm y el deterioro del ARNm de los genes aguas abajo. Una característica molecular importante de la DM1 es la regulación errónea del corte y empalme alternativo debido al secuestro de MBNL1 a las repeticiones CUG con estructura de horquilla bicatenaria. Entre más de dos docenas de eventos de corte y empalme regulados erróneamente en la DM1, se conoce que el corte y empalme anómalo del ClC-1 específico del músculo esquelético (canal de cloruro 1) es una de las causas de la miotonía. El aumento de la inclusión de exones que contienen codones de parada prematuros da como resultado una regulación negativa del ARNm y la proteína de ClC-1, lo que es suficiente para provocar miotonía (Charlet-B y otros (2002) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell* 10, 45–53; Mankodi, A. y otros (2002) Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell* 10, 35–44). En modalidades, los compuestos y métodos descritos en la presente descripción mejoran los efectos aguas abajo, lo que incluye la alteración del corte y empalme alternativo, la traducción del ARNm y la degradación del ARNm de los genes aguas abajo. En modalidades, los compuestos y métodos descritos en la presente descripción reducen el número de eventos de corte y empalme regulados erróneamente en la DM1 en comparación con un sujeto con DM1 que no se trata con compuestos o métodos de la

descripción. Por ejemplo, en algunas modalidades, los compuestos y métodos descritos en la presente descripción pueden reducir el número de eventos de corte y empalme regulados erróneamente en uno o más genes aguas abajo tales como 4833439L19Rik, Abcc9, Atp2a1, Arhgef10, Arhgap28, Armcx6, Angel1, Best3, Bin1, Brd2, Cacna1s, Cacna2d1, Cpd, Cpeb3, Ccpg1, Clasp1, Clcn1, Clk4, Cpeb2, Camk2g, Capzb, Copz2, Coch, cTNT, Ctu2, Cyp2s1, Dctn4, Dnm1l, Eya4, Efna3, Efna2, Fbxo31, Fbxo21, Frem2, Fgd4, Fuca1, Fn1, Gogla4, Gpr37l1, Greb1, Heg1, Insr, Impdh2, IR, Itgav, Jag2, Klc1, Kcan6, Kif13a, Ldb3, Lrrfip2, Mapt, Macf1, Map3k4, Mapkap1, Mbnl1, Mllt3, Mbnl2, Mef2c, Mpdz, Mrpl1, Mxra7, Mybpc1, Myo9a, Ncapd3, Ngfr, Ndr3, Ndufv3, Neb, Nfix, Numa1, Opa1, Pacsin2, Pcolce, Pdlim3, Pla2g15, Phactr4, Phka1, Phtf2, Ppp1r12b, Ppp3cc, Ppp1cc, Ramp2, Rapgef1, Rur1, Ryr1, Sorcs2, Spsb4, Scube2, Sema6c, Sfc8a3, Slain2, Sorbs1, Spag9, Tmem28, Tacc1, Tacc2, Ttc7, Tnik, Tnfrsf22, Tnfrsf25, Trappc9, Trim55, Ttn, Txnl4a, Txlnb, Ube2d3 o Vsp39. En modalidades, los compuestos y métodos descritos en la presente descripción reducen el número de exones que contienen codones de parada prematuros, lo que da como resultado la regulación negativa del ARNm de ClC-1 en comparación con un sujeto con DM1 que no se trata con compuestos o métodos de la descripción.

Los niveles de MBNL1 y CUGBP1 en el núcleo controlan un subconjunto de eventos de corte y empalme regulados por el desarrollo que se invierten en la DM1. En la etapa embrionaria, los niveles nucleares de MBNL1 son bajos y los niveles de CUGBP1 son altos. Durante el desarrollo, los niveles nucleares de MBNL1 aumentan mientras que los niveles de CUGBP1 disminuyen, lo que induce una transición embrionario a adulto de los objetivos de corte y empalme aguas abajo (lo que incluye el exón 11 de IR, los exones de ClC-1 que contienen codones de parada y el exón 5 de cTNT). Sin embargo, en la DM1, MBNL1 se secuestra a repeticiones CUG, lo que da como resultado una disminución de MBNL1 funcional, mientras que los niveles de CUGBP1 aumentan debido a la fosforilación y estabilización. Esto simula la condición embrionaria y potencia la expresión de isoformas embrionarias en adultos, lo que da como resultado múltiples síntomas de la enfermedad (Lee y Cooper; 2009). En modalidades, los compuestos y métodos descritos en la presente descripción reducen la cantidad de MBNL1 secuestrada, aumentan la cantidad de MBNL1 funcional, disminuyen los niveles de CUGBP1 en comparación con un sujeto con DM1 que no se trata con compuestos o métodos de la descripción.

MBNL1 y CELF1 (también denominados “CUGBP1”) son reguladores del desarrollo de los eventos de corte y empalme durante la transición fetal a adulto y la modificación de sus actividades en la DM1 conducen a la expresión de un patrón de corte y empalme fetal en tejidos adultos. El impacto

aguas abajo de MBNL1 baja y CELF1 alto incluye la alteración del corte y empalme alternativo, la traducción del ARNm y la degradación del ARNm en proteínas tales como transcritos de troponina cardiaca T (cTNT), receptor de insulina (INSR), canal iónico de cloruro específico del músculo (CLCN1) y ATPasa 1 de calcio del retículo sarcoplasmático/endoplasmático (ATP2A1), además de MBNL1. Konieczny y otros (2017) “Myotonic dystrophy: candidate small molecule therapeutics”, Drug Discovery Today. 22(11):1,740-1,748.

Los compuestos y métodos para tratar la distrofia miotónica mediante el uso de oligómeros antisentido dirigidos a repeticiones poliCUG en la UTR 3' del gen DMPK se describen en los documentos US10106796B2, US10111962B2, US20150080311A1, cada uno de los cuales se incorpora en la presente descripción como referencia en su totalidad para todos los propósitos. Sin embargo, dichos PMO o PPMO dirigidos a las repeticiones CUG para tratar la DM1 podrían tener limitaciones en el suministro de los oligos a los músculos, que son los tejidos afectados por la enfermedad.

En modalidades, la presente descripción enseña el uso de diversos péptidos penetrantes de células (CPP) para suministrar el AC (por ejemplo, PMO o ASO) y una secuencia de degradación descrita en la presente descripción, por ejemplo, en las Tablas 2 y 10, al citosol de la célula. En modalidades, el CPP o EEV conjugado con el AC suministra el AC de interés a la ubicación celular donde se ubica la secuencia objetivo en el pre-ARNm.

En modalidades, la enfermedad es una forma de distrofia miotónica (por ejemplo, distrofia miotónica tipo 1 o distrofia miotónica tipo 2). En modalidades, el gen objetivo es el gen DMPK, que codifica la proteína cinasa miotónica. En modalidades, los compuestos proporcionados en la presente descripción comprenden un AC (por ejemplo, ASO) que se dirige a la DMPK (por ejemplo, la región no traducida 3'/poliadenilación del gen DMPK) para degradar el gen DMPK. Oligonucleótidos ilustrativos que se dirigen a la DMPK para la degradación se proporcionan en la Tabla 10. La secuencia de degradación puede usarse en combinación con una secuencia de AC que comprende de 10-40 repeticiones CAG, lo que incluye, pero no se limita a, el AC proporcionado en la Tabla 2.

Tabla 10. Oligonucleótidos (AC) dirigidos a la DMPK para la degradación.

ID del oligo (AC)	Secuencia (5'-3')	SEQ ID NO:	Objetivo
	5'-CAG CAG CAG CAG CAG CAG CAG CAG-3'-clic-K-PEG12-Lys(CPP12)-NLS-AC (todos los monómeros de PMO)		DMPK
DMPK-A-17	GGGCCTTTTATTCGCGAGGGTCTGGG	151	DMPK

5	DMPK-A-18	GAGGGCCTTTTATTTCGCGAGGGTCG	152	DMPK
	DMPK-A-19	TGGAGGGCCTTTTATTTCGCGAGGGT	153	DMPK
	DMPK-A-20	GATGGAGGGCCTTTTATTTCGCGAGG	154	DMPK
	DMPK-A-21	CAGATGGAGGGCCTTTTATTTCGCGA	155	DMPK
	DMPK-A-22	GGCAGATGGAGGGCCTTTTATTTCGC	156	DMPK
10	DMPK-A-23	TGGGCAGATGGAGGGCCTTTTATTC	157	DMPK
	DMPK-A-24	TTTGGGCAGATGGAGGGCCTTTTAT	158	DMPK
	DMPK-A-25	GCTTTGGGCAGATGGAGGGCCTTTT	159	DMPK
	DMPK-A-26	GAGCTTTGGGCAGATGGAGGGCCTT	160	DMPK
	DMPK-A-27	CAGAGCTTTGGGCAGATGGAGGGCC	161	DMPK
15	DMPK-A-28	TCCAGAGCTTTGGGCAGATGGAGGG	162	DMPK
	DMPK-A-29	AGTCCAGAGCTTTGGGCAGATGGAG	163	DMPK
	DMPK-A-30	GGAGTCCAGAGCTTTGGGCAGATGG	164	DMPK
	DMPK-A-31	GTGGAGTCCAGAGCTTTGGGCAGAT	165	DMPK
	DMPK-A-32	CTGTGGAGTCCAGAGCTTTGGGCAG	166	DMPK
20	DMPK-A-33	CACTGTGGAGTCCAGAGCTTTGGGC	177	DMPK
	DMPK-A-34	GACACTGTGGAGTCCAGAGCTTTGG	178	DMPK
	DMPK-A-35	CGGACACTGTGGAGTCCAGAGCTTT	179	DMPK
	DMPK-A-36	CGCGGACACTGTGGAGTCCAGAGCT	180	DMPK
	DMPK-A-37	ACCGCGGACACTGTGGAGTCCAGAG	181	DMPK
	DMPK-A-38	AAACCGCGGACACTGTGGAGTCCAG	182	DMPK
	DMPK-A-39	GCAAACCGCGGACACTGTGGAGTCC	183	DMPK
	DMPK-A-40	ACGCAAACCGCGGACACTGTGGAGT	184	DMPK
	DMPK-A-41	CAACGCAAACCGCGGACACTGTGGA	185	DMPK

Ataxia espinocerebelosa-8 (SCA8)

En modalidades, se proporcionan compuestos, composiciones y métodos para tratar la ataxia espinocerebelosa-8 (SCA8). La SCA8 es una afección neurodegenerativa hereditaria caracterizada por una ataxia lentamente progresiva. Los síntomas aparecen normalmente entre la tercera y la quinta década de vida. Los síntomas incluyen anomalías del movimiento ocular, neuropatía sensorial, disfagia, ataxia cerebelosa y deterioro cognitivo.

La SCA8 se asocia con la repetición CTG·CUG expandida anómala heterocigota en la UTR 3' de dos genes solapados ATXN8OS y ATXN8. Los individuos sanos generalmente tienen entre 15 y 50 repeticiones CTG·CUG en los genes ATXN8OS y ATXN8. Los pacientes con SCA8 tienen más de 50 repeticiones CTG·CUG, a veces hasta 240 repeticiones CTG·CUG en los genes ATXN8OS y ATXN8.

Enfermedad similar a Huntington-2 (HDL2)

En modalidades, se proporcionan compuestos, composiciones y métodos para tratar la enfermedad similar a Huntington-2 (HDL2). La HDL2 es un trastorno neurodegenerativo autosómico dominante

que está relacionado fenotípicamente con la enfermedad de Huntington. La HDL2 se caracteriza por síntomas que incluyen corea, distonía, rigidez, bradicinesia y síntomas psiquiátricos tales como demencia. Los síntomas de la HDL2 se producen típicamente en la madurez y pueden conducir a una muerte prematura en aproximadamente 10-15 años.

- 5 La HDL2 se asocia con un CTG·CUG expandido en la UTR 3' del gen de la junctofilina 3 (JPH3) (16q24.3). Los individuos sanos tienen generalmente entre 6 y 27 repeticiones CTG·CUG en el gen JPH3. Los pacientes con HDL2 tienen más de 40 repeticiones CTG·CUG, a veces hasta 60 o más repeticiones CTG·CUG en el gen JPH3.

Distrofia endotelial corneal de Fuchs

- 10 En modalidades, se proporcionan compuestos, composiciones y métodos para tratar la distrofia endotelial corneal de Fuchs (FECD). La FECD (MIM 136800) es un trastorno degenerativo relacionado con la edad del endotelio corneal. La FECD se caracteriza por la pérdida progresiva de células endoteliales de la córnea, el engrosamiento de la membrana de Descemet y la deposición de matriz extracelular en forma de gotas. Cuando el número de células endoteliales se vuelve
15 críticamente bajo, la córnea se hincha y provoca pérdida de la visión (Elhalis y otros, Ocul Surf. 2010; 8(4):173-184).

- La FECD puede heredarse como un rasgo autosómico dominante con heterogeneidad genética. Las mutaciones heterocigotas raras en el gen alfa 2 del colágeno tipo VIII (COL8A2, MIM 120252) pueden dar lugar a una distrofia endotelial corneal de inicio temprano. Otros genes tales como el
20 transportador de borato de sodio de la familia 4 de portadores de soluto, miembro 11 (SLC4A11, MIM 610206), el factor de transcripción 8 (TCF8, MIM 189909), los dominios de homología de lipoxigenasa 1 (LOXHD1, MIM 613267) y similar a la proteína de unión a ATP/GTP 1 (AGBL1, MIM 615523) se asocian colectivamente con una pequeña fracción de casos de FECD de inicio en adultos. Los estudios de asociación en todo el genoma de la FECD de inicio en adultos han sugerido
25 que el factor de transcripción 4 (TCF4, MIM 602272) y más recientemente la proteína que contiene el motivo KN y el dominio de repetición de anquirina 4 (KANK4, MIM 614612), gamma laminina-1 (LAMC1, MIM150290), ATPasa transportadora de Na⁺/K⁺ y el polipéptido beta-1 (ATP1B1, MIM 182330), con el locus TCF4 observado, tienen un efecto predominante sobre la FECD (Mootha y otros, Investigative ophthalmology & visual science, 2017; 58, 4,579-4,585).

- 30 Las repeticiones de trinucleótidos expandidas en el locus CTG18.1 en el intrón 2 de TCF4 se asocian con la FECD (Wieben y otros, PLoS One. 2012; 7(11):e49083). Cada copia del alelo CTG18.1 expandido de más de 40 repeticiones de trinucleótidos CTG·CUG conduce a un riesgo significativo

de desarrollo de la FECD (Mootha y otros, Invest Ophthalmol Vis Sci. 2014; 55: 33–42). Se han informado focos nucleares de ARN, un rasgo distintivo de ARN con ganancia de función tóxica, en trastornos neurodegenerativos provocados por expansiones de repeticiones simples. El ARN con repeticiones CUG expandidas se acumula como focos nucleares en el endotelio corneal de sujetos con FECD con la expansión de la repetición del triplete CTG18.1, mientras que está ausente en muestras de control que carecen de la expansión del triplete (Mootha y otros, Invest Ophthalmol Vis Sci. 2015;56(3):2003-2011). El ARN con repeticiones CUG expandidas se colocaliza con el factor de corte y empalme del ARNm, de tipo muscleblind 1 (MBNL1), en focos nucleares en el endotelio como una marca distintiva molecular. La expansión de la repetición del triplete en el locus CTG18.1 puede mediar la disfunción endotelial a través del corte y empalme de genes aberrantes, como resultado de que los transcritos de ARN con CUG mutantes secuestran a MBNL1 (Du y otros, J Biol Chem. 2015; 290: 5,979–5,990). Por tanto, dos repeticiones de tripletes distintas convergen en los focos de ARN y la FECD, y es probable que los focos puedan desempeñar un papel causal para la FECD.

En modalidades, compuestos y métodos útiles en el tratamiento de la distrofia endotelial corneal de Fuchs (FECD) que reducen el ARN con repeticiones CUG expandidas con oligonucleótidos antisentido se describen en los documentos WO2018165541A1, US10760076B2, cada uno de los cuales se incorpora en la presente descripción como referencia en su totalidad para todos los propósitos. Sin embargo, dichos oligómeros de morfolino fosforodiamidato (PMO) o PMO conjugados con péptidos (PPMO) dirigidos a repeticiones CUG tales como las descritas en la técnica anterior, podrían tener limitaciones para tratar la DM1 en el suministro del oligo al tejido objetivo (por ejemplo, capas endoteliales en la córnea), que son los tejidos afectados por la enfermedad.

En modalidades, la presente descripción enseña el uso de diversos péptidos penetrantes de células (CPP) o vehículos de escape endosómico (EEV) para suministrar el AC (por ejemplo, PMO o ASO) descritos en la presente descripción, por ejemplo, en la Tabla 6, al citosol de la célula. En modalidades, el CPP o el EEV conjugado al AC suministra el AC de interés a la ubicación celular donde se ubica la secuencia objetivo en el pre-ARNm.

En modalidades, la enfermedad es la distrofia endotelial corneal de Fuchs (FECD). En modalidades, el gen objetivo es el *TCF4*, que codifica el factor de transcripción 4 (TCF-4), que también se conoce como factor de transcripción de inmunoglobulina 2 (ITF-2). En modalidades, los compuestos proporcionados en la presente descripción comprenden un oligonucleótido antisentido que se dirige a *TCF4*. Los oligonucleótidos ilustrativos que pueden usarse para dirigirse a *TCF4* se proporcionan

en la Tabla 2 y la Tabla 11.

Tabla 11. Oligonucleótidos ilustrativos dirigidos a la repetición de triplete expandida de TCF4

	Química del oligo	Diseño	Objetivo
5	PMO 21 mer	5'-CAG CAG CAG CAG CAG CAG CAG CAG -3' (todos los monómeros de PMO; SEQ ID NO:146)	TCF4 (CUG) _n
	PMO 25 mer	5'-CAG CAG CAG CAG CAG CAG CAG CAG C-3' (todos los monómeros de PMO; SEQ ID NO:147)	TCF4 (CUG) _n
10	EEV-NLS- PMO 21 mer	EEV-NLS_PMO_CAG 21 mer (conjugación de SEQ ID NO:146)	TCF4 (CUG) _n
	PMO 30 mer	5'-CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG-3' (todos los monómeros de PMO; SEQ ID NO:148)	TCF4 (CUG) _n
	PMO 30 mer	5'-AGC AGC AGC AGC AGC AGC AGC AGC AGC AGC AGC-3' (todos los monómeros de PMO; SEQ ID NO:149)	TCF4 (CUG) _n
15	PMO 30 mer	5'-GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA-3' (todos los monómeros de PMO; SEQ ID NO:150)	TCF4 (CUG) _n

PMO: Oligómero de morfolino fosforodiamidato

Mootha y otros (2017) informaron que la DM1 y la FECD se originan a partir de expansiones de CTG no codificantes aunque ambas no son enfermedades idénticas. La expansión de DMPK en la DM1 da como resultado una enfermedad multiorgánica que involucra diversos tejidos en el ojo, lo que incluye el cristalino, la retina y el endotelio corneal. Por el contrario, la expansión de la repetición en TCF4 parece afectar el endotelio corneal sin ninguna secuela aparente clínicamente a otros tejidos oculares u órganos corporales. Las expansiones mutantes en DMPK y TCF4 comparten similitudes importantes, tales como (i) focos nucleares que contienen repeticiones CUG expandidas, (ii) asociación de los focos con la proteína MBNL1, y (iii) una capacidad para provocar la FECD. Se sugiere que las expansiones de triplete tanto en la DMPK como en TCF4 pueden provocar el mismo fenotipo de tejido endotelial corneal de la FECD a través de mecanismos moleculares compartidos. Ver la patente de Estados Unidos núm. 10760076B2, publicación de solicitud internacional núm. WO2018165541A1, la publicación de solicitud de patente de Estados Unidos núm. 2016/0355796 y la publicación de solicitud de patente de Estados Unidos núm. 2018/0344817, cada una de los cuales se incorpora como referencia en la presente descripción, y las cuales describen enfermedades y genes

correspondientes propensos a formar y/o expandir repeticiones de nucleótidos en tándem.

Composiciones y métodos de administración

Los compuestos de la presente descripción pueden formularse en composiciones adecuadas para aplicaciones in vivo. Los compuestos y/o composiciones pueden administrarse a un paciente que
5 tiene, o se sospecha que tiene, una enfermedad asociada con una repetición de trinucleótidos expandida.

La aplicación *in vivo* de los compuestos descritos y las composiciones que los contienen, puede llevarse a cabo mediante cualquier método y técnica adecuados conocidos actualmente o prospectivamente para los expertos en la técnica. Por ejemplo, los compuestos descritos pueden
10 formularse en una composición fisiológica o farmacéuticamente aceptable y administrarse por cualquier vía adecuada conocida en la técnica lo que incluye, por ejemplo, vías de administración orales y parenterales. Como se usa en la presente, el término parenteral incluye la administración subcutánea, intradérmica, intravenosa, intramuscular, intraperitoneal, intraesternal e intratecal, tal como mediante inyección. La administración de los compuestos o composiciones descritos puede
15 ser en una administración única, o a intervalos continuos o distintos como puede determinarse fácilmente por un experto en la técnica.

Los compuestos descritos en la presente descripción y las composiciones que los comprenden, también pueden administrarse mediante el uso de tecnología de liposomas, cápsulas de liberación lenta, bombas implantables y contenedores biodegradables. Estos métodos de suministro pueden
20 proporcionar, ventajosamente, una dosificación uniforme durante un período extendido de tiempo. Los compuestos también pueden administrarse en sus formas derivadas de sal o sus formas cristalinas.

Los compuestos descritos en la presente descripción pueden formularse en composiciones farmacéuticas de acuerdo con métodos conocidos para preparar composiciones farmacéuticamente
25 aceptables. Las formulaciones se describen en detalle en una serie de fuentes que se conocen bien y están fácilmente disponibles para los expertos en la técnica. Por ejemplo, *Remington's Pharmaceutical Science* de E.W. Martin (1995) describe formulaciones que pueden usarse en relación con los métodos descritos. En general, los compuestos descritos en la presente descripción pueden formularse de manera que una cantidad efectiva del compuesto se combina con un portador
30 adecuado para facilitar la administración efectiva del compuesto. Las composiciones usadas también pueden estar en una variedad de formas. Estas incluyen, por ejemplo, formas de dosificación sólidas, semisólidas y líquidas, tales como comprimidos, píldoras, polvos, soluciones o suspensiones

líquidas, supositorios, soluciones inyectables e infusibles, y aerosoles. La forma depende del modo de administración y la aplicación terapéutica prevista. Las composiciones incluyen también portadores y diluyentes farmacéuticamente aceptables convencionales que son conocidos para los expertos en la técnica. Ejemplos de portadores o diluyentes para su uso con los compuestos incluyen

5 etanol, dimetilsulfóxido, glicerol, alúmina, almidón, solución salina y portadores y diluyentes equivalentes. Para proporcionar la administración de dichas dosificaciones para el tratamiento terapéutico deseado, las composiciones descritas en la presente descripción pueden comprender ventajosamente entre aproximadamente 0.1 % y 100 % en peso del total de uno o más de los compuestos objeto basado en el peso de la composición total, lo que incluye el portador o diluyente.

10 Las formulaciones adecuadas para la administración incluyen, por ejemplo, soluciones acuosas para inyección estériles, que pueden contener antioxidantes, tampones, bacteriostáticos y solutos que hacen que la formulación sea isotónica con la sangre del receptor previsto; y suspensiones estériles acuosas y no acuosas, que pueden incluir agentes de suspensión y agentes espesantes. Las formulaciones pueden presentarse en contenedores de dosis unitaria o multidosis, por ejemplo,

15 ampollas y viales sellados, y pueden almacenarse en una condición de secado por congelación (liofilización) que requiere solo la condición del portador líquido estéril, por ejemplo, agua para inyecciones, antes del uso. Pueden prepararse suspensiones y soluciones para inyección extemporáneas a partir de polvo estéril, gránulos, comprimidos, etc. Debe entenderse que además de los ingredientes particularmente mencionados anteriormente, las composiciones descritas en la

20 presente descripción pueden incluir otros agentes convencionales en la técnica que tienen en cuenta el tipo de formulación en cuestión.

Los compuestos descritos en la presente descripción y las composiciones que los comprenden, pueden suministrarse a una célula ya sea a través del contacto directo con la célula o mediante un medio portador. Los medios portadores para suministrar compuestos y composiciones a las células

25 se conocen en la técnica e incluyen, por ejemplo, encapsular la composición en un resto de liposoma. Otro medio para el suministro de compuestos y composiciones descritos en la presente descripción a una célula comprende unir los compuestos a una proteína o ácido nucleico que se dirige para el suministro a la célula objetivo. La patente de Estados Unidos núm. 6,960,648 y las publicaciones de solicitud de Estados Unidos núms. 20030032594 y 20020120100 describen secuencias de

30 aminoácidos que pueden acoplarse a otra composición y que permiten la translocación de la composición a través de las membranas biológicas. La publicación de solicitud de los Estados Unidos núm. 20020035243 describe además composiciones para transportar restos biológicos a través de las

membranas celulares para el suministro intracelular. Los compuestos también pueden incorporarse en polímeros, ejemplos de los cuales incluyen polímero de poli (D-L lactida-co-glicolida) para tumores intracraneales; poli[bis(p-carboxifenoxi) propano:ácido sebácico] en una relación molar 20:80 (como se usa en GLIADEL); condroitina; quitina; y quitosano.

- 5 Los compuestos y composiciones descritos en la presente descripción, lo que incluye sales o profármacos farmacéuticamente aceptables de estos, pueden administrarse por vía intravenosa, intramuscular o intraperitoneal mediante infusión o inyección. Las soluciones del agente activo o sus sales pueden prepararse en agua, mezcladas opcionalmente con un surfactante no tóxico. Las dispersiones también pueden prepararse en glicerol, polietilenglicoles líquidos, triacetina y mezclas
10 de estos y en aceites. En condiciones ordinarias de almacenamiento y uso, estas preparaciones pueden contener un conservante para impedir el crecimiento de microorganismos.

- Las formas de dosificación farmacéuticas adecuadas para inyección o infusión pueden incluir soluciones o dispersiones acuosas estériles o polvos estériles que comprenden el ingrediente activo, que se adaptan para la preparación extemporánea de soluciones o dispersiones inyectables o
15 infusibles estériles, opcionalmente encapsuladas en liposomas. La forma de dosificación final debe ser estéril, fluida y estable en las condiciones de fabricación y almacenamiento. El portador o vehículo líquido puede ser un solvente o medio de dispersión líquido que comprende, por ejemplo, agua, etanol, un poliol (por ejemplo, glicerol, propilenglicol, polietilenglicoles líquidos, y similares), aceites vegetales, ésteres de glicerilo no tóxicos y mezclas adecuadas de estos. La fluidez apropiada
20 puede mantenerse, por ejemplo, mediante la formación de liposomas, mediante el mantenimiento del tamaño de partícula requerido en el caso de las dispersiones o mediante el uso de surfactantes. Opcionalmente, la prevención de la acción de microorganismos puede producirse mediante diversos otros agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, pueden incluirse agentes isotónicos, por ejemplo,
25 azúcares, tampones o cloruro de sodio. La absorción prolongada de las composiciones inyectables puede producirse mediante la inclusión de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

- Las soluciones inyectables estériles se preparan mediante la incorporación de un compuesto y/o agente descrito en la presente descripción en la cantidad requerida en el solvente apropiado, con
30 diversos otros ingredientes enumerados anteriormente, según sea necesario, seguido de esterilización por filtro. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos de preparación incluyen el secado al vacío y las técnicas de secado por congelación, que

producen un polvo del ingrediente activo más cualquier ingrediente deseado adicional presente en las soluciones previamente esterilizadas por filtración.

Las dosificaciones útiles de los compuestos y agentes y composiciones farmacéuticas descritas en la presente descripción pueden determinarse mediante la comparación de su actividad *in vitro*, y su actividad *in vivo*, en modelos animales. Los métodos para la extrapolación de las dosificaciones efectivas en ratones y otros animales, a los humanos, se conocen en la técnica.

Los intervalos de dosificación para la administración de las composiciones son los suficientemente grandes para producir el efecto deseado en el que se afectan los síntomas o trastornos. La dosificación no debe ser tan grande como para provocar efectos secundarios adversos, tales como reacciones cruzadas no deseadas, reacciones anafilácticas, y similares. Generalmente, la dosificación variará con la edad, condición, sexo y extensión de la enfermedad en el paciente y puede determinarse por un experto en la técnica. La dosificación puede ajustarse por el médico individual en caso de cualquier contraindicación. La dosificación puede variar, y puede administrarse en una o más administraciones de dosis diarias, durante uno o varios días.

Se describen además composiciones farmacéuticas que comprenden un compuesto descrito en la presente descripción, en combinación con un portador farmacéuticamente aceptable. Las composiciones farmacéuticas adaptadas para la administración oral, tópica o parenteral, que comprenden una cantidad de un compuesto, se describen en la presente descripción. La dosis administrada a un paciente, particularmente un ser humano, debe ser suficiente para lograr una respuesta terapéutica en el paciente durante un marco de tiempo razonable, sin toxicidad letal, y que no provoque más que un nivel aceptable de efectos secundarios o morbilidad. Un experto en la técnica reconocerá que la dosificación dependerá de una variedad de factores, lo que incluye la condición (salud) del sujeto, el peso corporal del sujeto, el tipo de tratamiento concurrente, si lo hubiera, la frecuencia del tratamiento, la relación terapéutica, así como también la gravedad y la etapa de la condición patológica.

Se describen además kits que comprenden un compuesto descrito en la presente descripción y/o composiciones farmacéuticas que contienen el mismo, en uno o más contenedores. Los kits descritos pueden incluir opcionalmente portadores y/o diluyentes farmacéuticamente aceptables. En una modalidad, un kit incluye uno o más otros componentes, adjuntos o adyuvantes como se describe en la presente descripción. En una modalidad, un kit incluye instrucciones o materiales de empaque que describen cómo administrar un compuesto o composición del kit. Los contenedores del kit pueden ser de cualquier material adecuado, por ejemplo, vidrio, plástico, metal, etc., y de cualquier tamaño,

forma o configuración adecuada. En una modalidad, un compuesto y/o agente descrito en la presente descripción se proporciona en el kit como un sólido, tal como una forma de comprimido, píldora o polvo. En otra modalidad, un compuesto y/o agente descrito en la presente descripción se proporciona en el kit como líquido o solución. En una modalidad, el kit comprende una ampolla o jeringa que contiene un compuesto y/o agente descrito en la presente descripción en forma de líquido o solución.

En modalidades, el compuesto y/o composición de la descripción se administra a un paciente diagnosticado con una enfermedad asociada con una expansión de repetición de nucleótidos a una dosis de entre aproximadamente 0.1 mg/kg y aproximadamente 1,000 mg/kg, por ejemplo,

aproximadamente 0.1 mg/kg, aproximadamente 0.2 mg/kg, aproximadamente 0.3 mg/kg, aproximadamente 0.4 mg/kg, aproximadamente 0.5 mg/kg, aproximadamente 0.6 mg/kg, aproximadamente 0.7 mg/kg, aproximadamente 0.8 mg/kg, aproximadamente 0.9 mg/kg, aproximadamente 1 mg/kg, aproximadamente 2 mg/kg, aproximadamente 3 mg/kg, aproximadamente 4 mg/kg, aproximadamente 5 mg/kg, aproximadamente 6 mg/kg, aproximadamente 7 mg/kg, aproximadamente 8 mg/kg, aproximadamente 9 mg/kg, aproximadamente 10 mg/kg, aproximadamente 11 mg/kg, aproximadamente 12 mg/kg, aproximadamente 13 mg/kg, aproximadamente 14 mg/kg, aproximadamente 15 mg/kg, aproximadamente 16 mg/kg, aproximadamente 17 mg/kg, aproximadamente 18 mg/kg, aproximadamente 19 mg/kg, aproximadamente 20 mg/kg, aproximadamente 21 mg/kg, aproximadamente 22 mg/kg, aproximadamente 23 mg/kg, aproximadamente 24 mg/kg, aproximadamente 25 mg/kg, aproximadamente 26 mg/kg, aproximadamente 27 mg/kg, aproximadamente 28 mg/kg, aproximadamente 29 mg/kg, aproximadamente 30 mg/kg, aproximadamente 31 mg/kg, aproximadamente 32 mg/kg, aproximadamente 33 mg/kg, aproximadamente 34 mg/kg, aproximadamente 35 mg/kg, aproximadamente 36 mg/kg, aproximadamente 37 mg/kg, aproximadamente 38 mg/kg, aproximadamente 39 mg/kg, aproximadamente 40 mg/kg, aproximadamente 41 mg/kg, aproximadamente 42 mg/kg, aproximadamente 43 mg/kg, aproximadamente 44 mg/kg, aproximadamente 45 mg/kg, aproximadamente 46 mg/kg, aproximadamente 47 mg/kg, aproximadamente 48 mg/kg, aproximadamente 49 mg/kg, aproximadamente 50 mg/kg, aproximadamente 51 mg/kg, aproximadamente 52 mg/kg, aproximadamente 53 mg/kg, aproximadamente 54 mg/kg, aproximadamente 55 mg/kg, aproximadamente 56 mg/kg, aproximadamente 57 mg/kg, aproximadamente 58 mg/kg, aproximadamente 59 mg/kg, aproximadamente 60 mg/kg,

	aproximadamente 61 mg/kg,	aproximadamente 62 mg/kg,	aproximadamente 63 mg/kg,
	aproximadamente 64 mg/kg,	aproximadamente 65 mg/kg,	aproximadamente 66 mg/kg,
	aproximadamente 67 mg/kg,	aproximadamente 68 mg/kg,	aproximadamente 69 mg/kg,
	aproximadamente 70 mg/kg,	aproximadamente 71 mg/kg,	aproximadamente 72 mg/kg,
5	aproximadamente 73 mg/kg,	aproximadamente 74 mg/kg,	aproximadamente 75 mg/kg,
	aproximadamente 76 mg/kg,	aproximadamente 77 mg/kg,	aproximadamente 78 mg/kg,
	aproximadamente 79 mg/kg,	aproximadamente 80 mg/kg,	aproximadamente 81 mg/kg,
	aproximadamente 82 mg/kg,	aproximadamente 83 mg/kg,	aproximadamente 84 mg/kg,
	aproximadamente 85 mg/kg,	aproximadamente 86 mg/kg,	aproximadamente 87 mg/kg,
10	aproximadamente 88 mg/kg,	aproximadamente 89 mg/kg,	aproximadamente 90 mg/kg,
	aproximadamente 91 mg/kg,	aproximadamente 92 mg/kg,	aproximadamente 93 mg/kg,
	aproximadamente 94 mg/kg,	aproximadamente 95 mg/kg,	aproximadamente 96 mg/kg,
	aproximadamente 97 mg/kg,	aproximadamente 98 mg/kg,	aproximadamente 99 mg/kg,
	aproximadamente 100 mg/kg,	aproximadamente 110 mg/kg,	aproximadamente 120 mg/kg,
15	aproximadamente 130 mg/kg,	aproximadamente 140 mg/kg,	aproximadamente 150 mg/kg,
	aproximadamente 160 mg/kg,	aproximadamente 170 mg/kg,	aproximadamente 180 mg/kg,
	aproximadamente 190 mg/kg,	aproximadamente 200 mg/kg,	aproximadamente 210 mg/kg,
	aproximadamente 220 mg/kg,	aproximadamente 230 mg/kg,	aproximadamente 240 mg/kg,
	aproximadamente 250 mg/kg,	aproximadamente 260 mg/kg,	aproximadamente 270 mg/kg,
20	aproximadamente 280 mg/kg,	aproximadamente 290 mg/kg,	aproximadamente 300 mg/kg,
	aproximadamente 310 mg/kg,	aproximadamente 320 mg/kg,	aproximadamente 330 mg/kg,
	aproximadamente 340 mg/kg,	aproximadamente 350 mg/kg,	aproximadamente 360 mg/kg,
	aproximadamente 370 mg/kg,	aproximadamente 380 mg/kg,	aproximadamente 390 mg/kg,
	aproximadamente 400 mg/kg,	aproximadamente 410 mg/kg,	aproximadamente 420 mg/kg,
25	aproximadamente 430 mg/kg,	aproximadamente 440 mg/kg,	aproximadamente 450 mg/kg,
	aproximadamente 460 mg/kg,	aproximadamente 470 mg/kg,	aproximadamente 480 mg/kg,
	aproximadamente 490 mg/kg,	aproximadamente 500 mg/kg,	aproximadamente 510 mg/kg,
	aproximadamente 520 mg/kg,	aproximadamente 530 mg/kg,	aproximadamente 540 mg/kg,
	aproximadamente 550 mg/kg,	aproximadamente 560 mg/kg,	aproximadamente 570 mg/kg,
30	aproximadamente 580 mg/kg,	aproximadamente 590 mg/kg,	aproximadamente 600 mg/kg,
	aproximadamente 610 mg/kg,	aproximadamente 620 mg/kg,	aproximadamente 630 mg/kg,
	aproximadamente 640 mg/kg,	aproximadamente 650 mg/kg,	aproximadamente 660 mg/kg,

aproximadamente 670 mg/kg, aproximadamente 680 mg/kg, aproximadamente 690 mg/kg, aproximadamente 700 mg/kg, aproximadamente 710 mg/kg, aproximadamente 720 mg/kg, aproximadamente 730 mg/kg, aproximadamente 740 mg/kg, aproximadamente 750 mg/kg, aproximadamente 760 mg/kg, aproximadamente 770 mg/kg, aproximadamente 780 mg/kg, 5 aproximadamente 790 mg/kg, aproximadamente 800 mg/kg, aproximadamente 810 mg/kg, aproximadamente 820 mg/kg, aproximadamente 830 mg/kg, aproximadamente 840 mg/kg, aproximadamente 850 mg/kg, aproximadamente 860 mg/kg, aproximadamente 870 mg/kg, aproximadamente 880 mg/kg, aproximadamente 890 mg/kg, aproximadamente 900 mg/kg, aproximadamente 910 mg/kg, aproximadamente 920 mg/kg, aproximadamente 930 mg/kg, 10 aproximadamente 940 mg/kg, aproximadamente 950 mg/kg, aproximadamente 960 mg/kg, aproximadamente 970 mg/kg, aproximadamente 980 mg/kg, aproximadamente 990 mg/kg, o aproximadamente 1,000 mg/kg, lo que incluye todos los valores e intervalos en ellos e intermedios.

Métodos de tratamiento

La presente descripción proporciona un método para tratar una enfermedad en un sujeto que lo 15 necesita, que comprende administrar un compuesto y/o una composición que contiene el compuesto descrito en la presente descripción. En modalidades, la enfermedad es cualquiera de las enfermedades proporcionadas en la presente descripción. En modalidades, el gen o transcrito génico objetivo es cualquiera de los genes o transcritos génicos objetivo proporcionados en la presente descripción.

20 En modalidades, el paciente se identifica como que tiene, o está en riesgo de tener, cualquier enfermedad como se describe en la presente descripción. En modalidades, se proporciona un método para tratar una enfermedad asociada con una repetición CTG·CUG en una región no traducida 3' de un gen/transcrito. En modalidades, se proporciona un método para tratar la distrofia miotónica. En modalidades, se proporciona un método para tratar la distrofia miotónica tipo 1 (DM1). En 25 modalidades, se proporciona un método para tratar la SCA8. En modalidades, se proporciona un método para tratar la HDL2. En modalidades, se proporciona un método para tratar la FECD.

En modalidades, el tratamiento se refiere a mitigación parcial o completa, mejora, alivio, inhibición, retraso en el inicio, reducción de la gravedad y/o incidencia de uno o más síntomas en un sujeto.

El tratamiento de la enfermedad y/o los síntomas de la enfermedad puede producirse a través de una 30 variedad de mecanismos moleculares, tales como los descritos en la presente descripción.

En modalidades, se proporciona un método para alterar la expresión y/o actividad de un gen objetivo en un sujeto que lo necesita, que comprende administrar un compuesto descrito en la presente

descripción. En modalidades, el tratamiento da como resultado la disminución en la expresión de una proteína objetivo a partir de un transcrito objetivo. En modalidades, el tratamiento da como resultado la reducción de los niveles de un transcrito objetivo. En modalidades, el tratamiento da como resultado la modulación del corte y empalme de transcritos génicos aguas abajo que están regulados por el transcrito objetivo y/o proteínas que se unen al transcrito objetivo. En modalidades, la modulación del corte y empalme de transcritos génicos aguas abajo da como resultado un aumento en los transcritos aguas abajo y/o las isoformas de proteínas aguas abajo que se asocian con fenotipos sanos. En modalidades, el corte y empalme alternativo da como resultado una disminución en los transcritos aguas abajo y/o las isoformas de proteínas aguas abajo que se asocian con fenotipos de la enfermedad.

En modalidades, se proporciona un método para tratar la DM1 mediante la reducción del secuestro de al menos una proteína de unión al ARN a un pre-ARNm que comprende al menos una repetición CUG expandida. En modalidades, se proporciona un método para tratar la DM1 mediante la reducción de la acumulación de un pre-ARNm que comprende al menos una repetición CUG expandida. En modalidades, se proporciona un método para tratar la DM1 mediante la corrección de los defectos de corte y empalme de transcritos génicos aguas abajo.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución del nivel del transcrito objetivo y/o la expresión de los genes del transcrito objetivo (por ejemplo, DMPK, TCF4, JPH3, ATXN8OS y/o ATXN8) en más de aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción. En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución del nivel del transcrito objetivo (por ejemplo, DMPK, TCF4, JPH3, ATXN8OS y/o ATXN8) y/o la expresión del transcrito objetivo en aproximadamente 5 % a aproximadamente 100 %, aproximadamente 10 % a aproximadamente 100 %, aproximadamente 20 % a aproximadamente 100 %, aproximadamente 50 % a aproximadamente 100 %, aproximadamente 70 % a aproximadamente 100 %, aproximadamente

80 % a aproximadamente 100 %, aproximadamente 90 % a aproximadamente 100 %, aproximadamente 95 % a aproximadamente 100 %, aproximadamente 40 % a aproximadamente 95 %, aproximadamente 50 % a aproximadamente 95 %, aproximadamente 70 % a aproximadamente 95 % o aproximadamente 90 % a aproximadamente 95 % en comparación con el nivel promedio del transcrito y/o proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución del número de focos nucleares de ARN con repeticiones CUG de un gen objetivo (por ejemplo, DMPK, TCF4, JPH3, ATXN8OS y/o ATXN8) en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de focos en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción. En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución del número de focos nucleares de ARN con repeticiones CUG de un gen objetivo (por ejemplo, DMPK, TCF4, JPH3, ATXN8OS y/o ATXN8) en aproximadamente 5 % a aproximadamente 100 %, aproximadamente 10 % a aproximadamente 100 %, aproximadamente 20 % a aproximadamente 100 %, aproximadamente 50 % a aproximadamente 100 %, aproximadamente 70 % a aproximadamente 100 %, aproximadamente 80 % a aproximadamente 100 %, aproximadamente 90 % a aproximadamente 100 %, aproximadamente 95 % a aproximadamente 100 %, aproximadamente 40 % a aproximadamente 95 %, aproximadamente 50 % a aproximadamente 95 %, aproximadamente 70 % a aproximadamente 95 % o aproximadamente 90 % a aproximadamente 95 %, en comparación con el nivel promedio de focos en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una

disminución del nivel del transcrito aguas abajo y/o la expresión de un producto génico aguas abajo que se asocia con un fenotipo de enfermedad en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción. En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución del nivel del transcrito aguas abajo y/o la expresión de un producto génico aguas abajo que se asocia con un fenotipo de la enfermedad en aproximadamente 5 % a aproximadamente 100 %, aproximadamente 10 % a aproximadamente 100 %, aproximadamente 20 % a aproximadamente 100 %, aproximadamente 50 % a aproximadamente 100 %, aproximadamente 70 % a aproximadamente 100 %, aproximadamente 80 % a aproximadamente 100 %, aproximadamente 90 % a aproximadamente 100 %, aproximadamente 95 % a aproximadamente 100 %, aproximadamente 40 % a aproximadamente 95 %, aproximadamente 50 % a aproximadamente 95 %, aproximadamente 70 % a aproximadamente 95 % o aproximadamente 90 % a aproximadamente 95 % en comparación con el nivel promedio de la proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento del nivel del transcrito aguas abajo y/o la expresión de un producto génico aguas abajo que se asocia con un fenotipo sano en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no

conjugado con un CPP cíclico descrito en la presente descripción. En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento del nivel del transcrito aguas abajo y/o la expresión de un producto génico aguas abajo que se asocia con un fenotipo sano en aproximadamente 5 % a aproximadamente 100 %, aproximadamente 10 % a aproximadamente 100 %, aproximadamente 20 % a aproximadamente 100 %, aproximadamente 50 % a aproximadamente 100 %, aproximadamente 70 % a aproximadamente 100 %, aproximadamente 80 % a aproximadamente 100 %, aproximadamente 90 % a aproximadamente 100 %, aproximadamente 95 % a aproximadamente 100 %, aproximadamente 40 % a aproximadamente 95 %, aproximadamente 50 % a aproximadamente 95 %, aproximadamente 70 % a aproximadamente 95 % o aproximadamente 90 % a aproximadamente 95 % en comparación con el nivel promedio de la proteína en el sujeto antes del tratamiento o de uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución en la expresión de una isoforma de la proteína asociada con un fenotipo de enfermedad en el tejido corneal de un sujeto, tejido muscular, tejido del diafragma, cuádriceps, tríceps, tibial anterior, gastrocnemio o corazón en más de aproximadamente un 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína en el tejido corneal del sujeto, tejido muscular, tejido del diafragma, cuádriceps, tríceps, tibial anterior, gastrocnemio o corazón antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción. En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución en la expresión de una isoforma de proteína asociada con un fenotipo de enfermedad en el tejido corneal de un sujeto, tejido muscular, tejido del diafragma, cuádriceps, tríceps, tibial anterior, gastrocnemio o corazón en más de aproximadamente 5 % a aproximadamente 100 %, aproximadamente 10 % a aproximadamente 100 %, aproximadamente 20 % a aproximadamente 100 %, aproximadamente 50 % a aproximadamente 100 %

%, aproximadamente 70 % a aproximadamente 100 %, aproximadamente 80 % a aproximadamente 100 %, aproximadamente 90 % a aproximadamente 100 %, aproximadamente 95 % a aproximadamente 100 %, aproximadamente 40 % a aproximadamente 95 %, aproximadamente 50 % a aproximadamente 95 %, aproximadamente 70 % a aproximadamente 95 % o aproximadamente 90 % a aproximadamente 95 % en comparación con el nivel promedio de la proteína en el tejido corneal del sujeto, tejido muscular, tejido del diafragma, cuádriceps, tríceps, tibial anterior, gastrocnemio o corazón antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento en la expresión de una proteína aguas abajo sometida a corte y empalme alternativo en el tejido corneal de un sujeto, tejido muscular, tejido del diafragma, cuádriceps o corazón en más de aproximadamente un 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 100 %, aproximadamente 150 %, aproximadamente 200 %, aproximadamente 250 %, aproximadamente 300 %, aproximadamente 350 %, aproximadamente 400 %, aproximadamente 450 %, aproximadamente 500 %, aproximadamente 550 %, aproximadamente 600 %, aproximadamente 650 %, aproximadamente 700 %, aproximadamente 750 %, aproximadamente 800 %, aproximadamente 850 %, aproximadamente 900 %, aproximadamente 950 % o aproximadamente 1,000 % o más, en comparación con el nivel promedio de la proteína aguas abajo en el tejido corneal del sujeto, tejido muscular, tejido del diafragma, cuádriceps o corazón antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento o disminución en la expresión de un isómero de la proteína de tipo salvaje en el tejido corneal de un sujeto, tejido muscular, tejido del diafragma, cuádriceps o corazón en más de aproximadamente un 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 100 %, aproximadamente 150 %, aproximadamente 200 %, aproximadamente 250 %, aproximadamente 300 %, aproximadamente 350 %, aproximadamente 400 %, aproximadamente 450 %, aproximadamente 500 %, aproximadamente 550 %, aproximadamente 600 %, aproximadamente 650 %, aproximadamente 700 %, aproximadamente 750 %, aproximadamente 800 %, aproximadamente 850 %, aproximadamente 900 %, aproximadamente 950 % o aproximadamente 1,000 % o más, en comparación con el nivel promedio de la proteína de tipo salvaje en el tejido corneal del sujeto, tejido muscular, tejido del diafragma, cuádriceps o corazón antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio del isómero de la proteína de tipo salvaje en el tejido corneal del sujeto, tejido muscular, tejido del diafragma, cuádriceps o corazón antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado una disminución en la expresión de una proteína en un tejido de interés de un sujeto en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína en el tejido de interés del sujeto antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento en la expresión de una proteína aguas abajo sometida a corte y empalme alternativo en el tejido de interés de un sujeto en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 100 %, aproximadamente 150 %, aproximadamente 200 %, aproximadamente 250 %, aproximadamente 300 %, aproximadamente 350 %, aproximadamente 400 %, aproximadamente 450 %, aproximadamente 500 %, aproximadamente 550 %, aproximadamente 600 %, aproximadamente 650 %, aproximadamente 700 %, aproximadamente 750 %, aproximadamente 800, aproximadamente 850 %, aproximadamente 900 %, aproximadamente 950 % o aproximadamente 1,000 % o más, en comparación con el nivel

promedio de la proteína aguas abajo en el tejido de interés del sujeto antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

- 5 En modalidades, el tratamiento de acuerdo con la presente descripción da como resultado un aumento o disminución en la expresión de un isómero de la proteína de tipo salvaje aguas abajo en el tejido de interés de un sujeto en más de aproximadamente 5 %, por ejemplo, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 % y aproximadamente 100 %, en comparación con el nivel promedio de la proteína aguas abajo en el tejido de interés del sujeto antes del tratamiento, en comparación con uno o más individuos de control con enfermedad similar sin tratamiento, o en comparación con el tratamiento con un AC no conjugado con un CPP cíclico descrito en la presente descripción.

En modalidades, el tejido de interés del sujeto es tejido corneal o tejido muscular.

- Los términos “mejorar”, “aumentar”, “reducir”, “disminuir”, y similares, como se usan en la presente, indican valores que son con relación a un control. En modalidades, un control adecuado es una medición inicial, tal como una medición en el mismo individuo antes del inicio del tratamiento descrito en la presente descripción, o una medición en un individuo de control (o múltiples individuos de control) en ausencia del tratamiento descrito en la presente descripción. Un “individuo de control” es un individuo afectado con la misma enfermedad, que tiene aproximadamente la misma edad y/o género que el individuo que se trata (para garantizar que las etapas de la enfermedad en el individuo tratado y el/los individuo(s) de control sean comparables).

- El individuo (también denominado “paciente” o “sujeto”) que se trata es un individuo (feto, lactante, niño, adolescente o humano adulto) que tiene una enfermedad o que tiene el potencial de desarrollar una enfermedad. El individuo puede tener una enfermedad mediada por la expresión aberrante de un gen o por el corte y empalme aberrante de un gen. En diversas modalidades, el individuo que tiene la enfermedad puede tener niveles de expresión o actividad de proteínas aguas abajo que son menores que aproximadamente 1-99 % de los niveles normales de expresión o actividad de proteínas de tipo salvaje en un individuo no afectado por la enfermedad. En modalidades, el intervalo incluye, pero

no se limita a, menor que aproximadamente 80-99 %, menor que aproximadamente 65-80 %, menor que aproximadamente 50-65 %, menor que aproximadamente 30-50 %, menor que aproximadamente 25-30 %, menor que aproximadamente 20-25 %, menor que aproximadamente 15-20 %, menor que aproximadamente 10-15 %, menor que aproximadamente 5-10 %, menor que aproximadamente 1-5 % los niveles normales de expresión o actividad que las proteínas de tipo salvaje. En modalidades, el individuo puede tener niveles de expresión o actividad de proteínas aguas abajo que son 1-500 % más altos que los niveles normales de expresión o actividad de las proteínas objetivo de tipo salvaje en un individuo no afectado por la enfermedad. En modalidades, el intervalo incluye, pero no se limita a, mayor que aproximadamente 1-10 %, aproximadamente 10-50 %, aproximadamente 50-100 %, aproximadamente 100-200 %, aproximadamente 200-300 %, aproximadamente 300-400 %, aproximadamente 400-500 % o aproximadamente 500-1,000 % los niveles normales de expresión o actividad de las proteínas objetivo de tipo salvaje.

En modalidades, el individuo es un individuo al que se le ha diagnosticado recientemente la enfermedad. Típicamente, el tratamiento temprano (el tratamiento que comienza lo antes posible después del diagnóstico) es importante para minimizar los efectos de la enfermedad y para maximizar los beneficios del tratamiento.

Ciertas definiciones

Como se usa en la descripción y las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el/la” incluyen los referentes plurales a menos que el contexto lo indique claramente de cualquier otra manera. Por tanto, por ejemplo, la referencia a “una composición” incluye mezclas de dos o más de dichas composiciones, la referencia a “un agente” incluye mezclas de dos o más de dichos agentes, la referencia a “el componente” incluye mezclas de dos o más de dichos componentes, y similares.

El término “aproximadamente”, cuando precede inmediatamente a un valor numérico, significa un intervalo (por ejemplo, más o menos 20 %, 10 % o 5 % de ese valor). Por ejemplo, “aproximadamente 50” puede significar de 45 a 55, “aproximadamente 25,000” puede significar de 22,500 a 27,500, etc., a menos que el contexto de la descripción lo indique de cualquier otra manera, o sea inconsistente con dicha interpretación. Por ejemplo, en una lista de valores numéricos tales como “aproximadamente 49, aproximadamente 50, aproximadamente 55,...”, “aproximadamente 50” significa un intervalo que se extiende hasta menos de la mitad del/de los intervalo(s) entre los valores anteriores y posteriores, por ejemplo, más de 49.5 a menos de 52.5. Además, las frases “menor que aproximadamente” un valor o “mayor que aproximadamente” un valor deben entenderse en vista de la definición del término “aproximadamente” proporcionada en la presente descripción.

De manera similar, el término “aproximadamente”, cuando precede a una serie de valores numéricos o un intervalo de valores (por ejemplo, “aproximadamente 10, 20, 30” o “aproximadamente 10-30”) se refiere, respectivamente, a todos los valores en la serie, o los puntos finales del intervalo.

Como se usa en la presente, “péptido penetrante de células” o “CPP” se refiere a un péptido que facilita el suministro de una carga, por ejemplo, un resto terapéutico (TM) a una célula. En modalidades, el CPP es cíclico, y se representa como “cCPP”. En modalidades, el cCPP es capaz de dirigir un resto terapéutico para penetrar la membrana de una célula. En modalidades, el cCPP suministra el resto terapéutico al citosol de la célula. En modalidades, el cCPP suministra un compuesto antisentido (AC) a una ubicación celular donde está ubicado un pre-ARNm.

Como se usa en la presente, el término “vehículo de escape endosómico” (EEV) se refiere a un cCPP que se conjuga mediante un enlace químico (es decir, un enlace covalente o interacción no covalente) a un enlazador y/o un péptido exocíclico (EP). El EEV puede ser un EEV de fórmula (B).

Como se usa en la presente, el término “EEV-conjugado” se refiere a un vehículo de escape endosómico definido en la presente descripción conjugado mediante un enlace químico (es decir, un enlace covalente o interacción no covalente) a una carga. La carga puede ser un resto terapéutico (por ejemplo, un oligonucleótido, péptido o molécula pequeña) que puede suministrarse a una célula mediante el EEV. El EEV-conjugado puede ser un EEV-conjugado de fórmula (C).

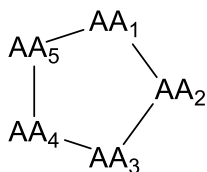
Como se usa en la presente, el término “péptido exocíclico” (EP) y “péptido modulador” (MP) puede usarse indistintamente para referirse a dos o más residuos de aminoácidos enlazados mediante un enlace peptídico que pueden conjugarse a un péptido penetrante de células cíclico (cCPP) descrito en la presente descripción. El EP, cuando se conjuga con un péptido cíclico descrito en la presente descripción, puede alterar la distribución y/o retención en el tejido del compuesto. Típicamente, el EP comprende al menos un residuo de aminoácido cargado positivamente, por ejemplo, al menos un residuo de lisina y/o al menos un residuo de arginina. Ejemplos no limitantes de EP se describen en la presente descripción. El EP puede ser un péptido que se ha identificado en la técnica como una “secuencia de localización nuclear” (NLS). Ejemplos no limitantes de secuencias de localización nuclear incluyen la secuencia de localización nuclear del antígeno T grande del virus SV40, cuya unidad funcional mínima es la secuencia de siete aminoácidos PKKKRKV (SEQ ID NO:42), la NLS bipartita de la nucleoplasmina con la secuencia NLSKRPAAIKKAGQAKKKK (SEQ ID NO:52), la secuencia de localización nuclear de c-myc que tiene la secuencia de aminoácidos PAAKRVKLD (SEQ ID NO:53) o RQRRNELKRSF (SEQ ID NO:54), la secuencia RMRKFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO:50) del dominio

IBB de la importina-alfa, las secuencias VSRKRPRP (SEQ ID NO:57) y PPKKARED (SEQ ID NO:58) de la proteína T del mioma, la secuencia PQQKKKPL (SEQ ID NO:59) del p53 humano, la secuencia SALIKKKKMAP (SEQ ID NO:60) de c-abl IV de ratón, las secuencias DRLRR (SEQ ID NO:61) y PKQKKRK (SEQ ID NO:62) del virus de la influenza NS1, la secuencia RKLKKKIKKL (SEQ ID NO:63) del antígeno del virus de la hepatitis delta, la secuencia REKKKFLKRR (SEQ ID NO:64) de la proteína Mx1 de ratón, la secuencia KRKGDEV DGVDEVAKKKSKK (SEQ ID NO:65) de la poli(ADP-ribosa) polimerasa humana, y la secuencia RKCLQAGMNLEARKTKK (SEQ ID NO:66) de los receptores de hormonas esteroideas (humanos) glucocorticoides. La publicación internacional núm. 2001/038547 describe ejemplos adicionales de NLS, y se incorpora como referencia en la presente descripción en su totalidad.

Como se usa en la presente, “enlazador” o “L” se refiere a un resto que une covalentemente uno o más restos (por ejemplo, un péptido exocíclico (EP) y una carga, por ejemplo, un oligonucleótido, péptido o molécula pequeña) al péptido penetrante de células cíclico (cCPP). El enlazador puede comprender un aminoácido o polipéptido natural o no natural. El enlazador puede ser un compuesto sintético que contiene dos o más grupos funcionales apropiados adecuados para unir el cCPP a un resto de carga, para formar de esta manera los compuestos descritos en la presente descripción. El enlazador puede comprender un resto de polietilenglicol (PEG). El enlazador puede comprender uno o más aminoácidos. El cCPP puede unirse covalentemente a una carga mediante un enlazador.

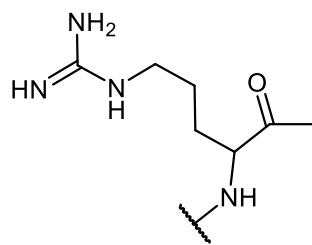
Los términos “péptido”, “proteína” y “polipéptido” se usan indistintamente para referirse a una molécula natural o sintética que comprende dos o más aminoácidos enlazados por el grupo carboxilo de un aminoácido al grupo alfa amino de otro. Dos o más residuos de aminoácidos pueden enlazarse por el grupo carboxilo de un aminoácido al grupo alfa amino. Dos o más aminoácidos del polipéptido pueden unirse mediante un enlace peptídico. El polipéptido puede incluir una modificación de la cadena principal peptídica en la que dos o más aminoácidos se unen covalentemente mediante un enlace que no es un enlace peptídico. El polipéptido puede incluir uno o más aminoácidos no naturales, análogos de aminoácidos, u otras moléculas sintéticas que son capaces de integrarse en un polipéptido. El término polipéptido incluye aminoácidos de origen natural y artificiales. El término polipéptido incluye péptidos, por ejemplo, que incluyen de aproximadamente 2 a aproximadamente 100 residuos de aminoácidos, así como también proteínas, que incluyen más de aproximadamente 100 residuos de aminoácidos, o más de aproximadamente 1,000 residuos de aminoácidos, lo que incluye, pero no se limita a, proteínas terapéuticas tales como anticuerpos, enzimas, receptores, proteínas solubles, y similares.

Como se usa en la presente, el término “contiguo” se refiere a dos aminoácidos, que se conectan mediante un enlace covalente. Por ejemplo, en el contexto de un péptido penetrante de células cíclico



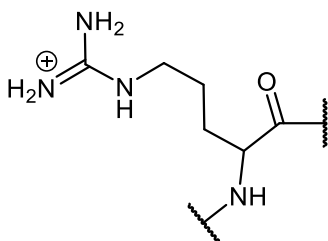
representativo (cCPP) tal como , AA₁/AA₂, AA₂/AA₃, AA₃/AA₄ y AA₅/AA₁ ejemplifican pares de aminoácidos contiguos.

Un residuo de una especie química, como se usa en la presente, se refiere a un derivado de la especie química que está presente en un producto en particular. Para formar el producto, al menos un átomo de la especie se reemplaza por una unión a otro resto, de manera que el producto contiene un derivado, o residuo, de la especie química. Por ejemplo, los péptidos penetrantes de células cíclicas (cCPP) descritos en la presente descripción, tienen aminoácidos (por ejemplo, arginina) incorporados en ellos a través de la formación de uno o más enlaces peptídicos. Los aminoácidos incorporados en el cCPP pueden denominarse residuos, o simplemente como aminoácidos. Por



ejemplo, arginina o un residuo de arginina se refiere a

El término “forma protonada de este/esta” se refiere a una forma protonada de un aminoácido o cadena lateral. Por ejemplo, el grupo guanidina en la cadena lateral de la arginina puede protonarse para formar un grupo guanidinio. La estructura de una forma protonada de la arginina es



Como se usa en la presente, el término “quiralidad” se refiere a una molécula que tiene más de un estereoisómero que difiere en la disposición espacial tridimensional de los átomos, en la que un estereoisómero es una imagen especular no superponible del otro. Los aminoácidos, excepto la glicina, tienen un átomo de carbono quiral adyacente al grupo carboxilo. El término “enantiómero” se refiere a estereoisómeros que son quirales. La molécula quiral puede ser un residuo de aminoácido

que tiene un enantiómero “D” y “L”. Las moléculas sin un centro quiral, tales como la glicina, pueden denominarse “aquirales”.

Como se usa en la presente, el término “hidrófobo” se refiere a un resto que no es soluble en agua o tiene una solubilidad mínima en agua. Generalmente, los restos neutros y/o los restos no polares, o los restos que son predominantemente neutros y/o no polares, son hidrófobos. La hidrofobicidad puede medirse mediante uno de los métodos descritos en la presente descripción.

Como se usa en la presente, “aromático” se refiere a una molécula cíclica insaturada que tiene $4n + 2 \pi$ electrones, en donde n es cualquier número entero. “Heteroaromático”, definido más abajo, es un subconjunto de aromático. Los ejemplos de aminoácidos aromáticos incluyen la fenilalanina y la naftilalanina. El término “no aromática” se refiere a cualquier molécula que no cae dentro de la definición de aromático. Por ejemplo, cualquier molécula lineal, ramificada o cíclica que no caiga dentro de la definición de aromática es no aromática. Los ejemplos de aminoácidos no aromáticos incluyen, pero no se limitan a, glicina y citrulina.

“Alquilo”, “cadena de alquilo” o “grupo alquilo” se refiere a un radical de cadena hidrocarbonada completamente saturada, lineal o ramificada que tiene de uno a cuarenta átomos de carbono, y que se une al resto de la molécula mediante un simple enlace. Se incluyen los alquilos que comprenden cualquier número de átomos de carbono de 1 a 40. Un alquilo que comprende hasta 40 átomos de carbono es un alquilo C_1-C_{40} , un alquilo que comprende hasta 10 átomos de carbono es un alquilo C_1-C_{10} , un alquilo que comprende hasta 6 átomos de carbono es un alquilo C_1-C_6 , y un alquilo que comprende hasta 5 átomos de carbono es un alquilo C_1-C_5 . Un alquilo C_1-C_5 incluye los alquilos C_5 , alquilos C_4 , alquilos C_3 , alquilos C_2 y alquilo C_1 (es decir, metilo). Un alquilo C_1-C_6 incluye todos los restos descritos anteriormente para los alquilos C_1-C_5 , pero también incluye los alquilos C_6 . Un alquilo C_1-C_{10} incluye todos los restos descritos anteriormente para los alquilos C_1-C_5 y alquilos C_1-C_6 , pero también incluye los alquilos C_7 , C_8 , C_9 y C_{10} . De manera similar, un alquilo C_1-C_{12} incluye todos los restos anteriores, pero también incluye los alquilos C_{11} y C_{12} . Los ejemplos no limitantes de alquilo C_1-C_{12} incluyen metilo, etilo, *n*-propilo, *i*-propilo, *sec*-propilo, *n*-butilo, *i*-butilo, *sec*-butilo, *t*-butilo, *n*-pentilo, *t*-amilo, *n*-hexilo, *n*-heptilo, *n*-octilo, *n*-nonilo, *n*-decilo, *n*-undecilo y *n*-dodecilo. A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo alquilo puede sustituirse opcionalmente.

“Alquilenos”, “cadena de alquilenos” o “grupo alquilenos” se refiere a un radical de cadena hidrocarbonada divalente completamente saturada, lineal o ramificada, que tiene de uno a cuarenta átomos de carbono. Ejemplos no limitantes de alquilenos C_2-C_{40} incluyen etileno, propileno, *n*-

butileno, etenileno, propenileno, *n*-butenileno, propinileno, *n*-butinileno, y similares. A menos que se indique de cualquier otra manera específicamente en la descripción, una cadena de alquileo puede sustituirse opcionalmente.

5 “Alquenilo”, “cadena de alquenilo” o “grupo alquenilo” se refiere a un radical de cadena hidrocarbonada lineal o ramificada que tiene de dos a cuarenta átomos de carbono y que tiene uno o más dobles enlaces carbono-carbono. Cada grupo alquenilo se une al resto de la molécula mediante un simple enlace. Se incluyen los grupos alquenilo que comprenden cualquier número de átomos de carbono de 2 a 40. Un grupo alquenilo que comprende hasta 40 átomos de carbono es un alquenilo C₂-C₄₀, un alquenilo que comprende hasta 10 átomos de carbono es un alquenilo C₂-C₁₀, un grupo
10 alquenilo que comprende hasta 6 átomos de carbono es un alquenilo C₂-C₆, y un alquenilo que comprende hasta 5 átomos de carbono es un alquenilo C₂-C₅. Un alquenilo C₂-C₅ incluye los alquenos C₅, alquenos C₄, alquenos C₃ y alquenos C₂. Un alquenilo C₂-C₆ incluye todos los restos descritos anteriormente para los alquenos C₂-C₅, pero también incluye los alquenos C₆. Un alquenilo C₂-C₁₀ incluye todos los restos descritos anteriormente para los alquenos C₂-C₅ y
15 alquenos C₂-C₆, pero también incluye los alquenos C₇, C₈, C₉ y C₁₀. De manera similar, un alquenilo C₂-C₁₂ incluye todos los restos anteriores, pero también incluye los alquenos C₁₁ y C₁₂. Los ejemplos no limitantes de alquenilo C₂-C₁₂ incluyen etenilo (vinilo), 1-propenilo, 2-propenilo (alilo), iso-propenilo, 2-metil-1-propenilo, 1-butenilo, 2-butenilo, 3-butenilo, 1-pentenilo, 2-pentenilo, 3-pentenilo, 4-pentenilo, 1-hexenilo, 2-hexenilo, 3-hexenilo, 4-hexenilo, 5-hexenilo, 1-heptenilo, 2-heptenilo, 3-heptenilo, 4-heptenilo, 5-heptenilo, 6-heptenilo, 1-octenilo, 2-octenilo, 3-octenilo, 4-octenilo, 5-octenilo, 6-octenilo, 7-octenilo, 1-nonenilo, 2-nonenilo, 3-nonenilo, 4-nonenilo, 5-nonenilo, 6-nonenilo, 7-nonenilo, 8-nonenilo, 1-decenilo, 2-decenilo, 3-decenilo, 4-decenilo, 5-decenilo, 6-decenilo, 7-decenilo, 8-decenilo, 9-decenilo, 1-undecenilo, 2-undecenilo, 3-undecenilo, 4-undecenilo, 5-undecenilo, 6-undecenilo, 7-undecenilo, 8-undecenilo, 9-undecenilo,
20 10-undecenilo, 1-dodecenilo, 2-dodecenilo, 3-dodecenilo, 4-dodecenilo, 5-dodecenilo, 6-dodecenilo, 7-dodecenilo, 8-dodecenilo, 9-dodecenilo, 10-dodecenilo y 11-dodecenilo. A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo alquilo puede sustituirse opcionalmente.

30 “Alquenileno”, “cadena de alquenileno” o “grupo alquenileno” se refiere a un radical de cadena hidrocarbonada divalente lineal o ramificada, que tiene de dos a cuarenta átomos de carbono, y que tiene uno o más dobles enlaces carbono-carbono. Los ejemplos no limitantes de alquenileno C₂-C₄₀ incluyen eteno, propileno, buteno, y similares. A menos que se indique de cualquier otra manera

específicamente en la descripción, una cadena de alquenileno puede sustituirse opcionalmente.

“Alcoxi” o “grupo alcoxi” se refiere al grupo -OR, donde R es alquilo, alquenilo, alquinilo, cicloalquilo o heterociclilo, como se define en la presente descripción. A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo alcoxi puede sustituirse

“Acilo” o “grupo acilo” se refiere a los grupos -C(O)R, donde R es hidrógeno, alquilo, alquenilo, alquinilo, carbociclilo o heterociclilo, como se define en la presente descripción. A menos que se indique de cualquier otra manera específicamente en la descripción, el acilo puede sustituirse opcionalmente.

“Alquilcarbamoilo” o “grupo alquilcarbamoilo” se refiere al grupo -O-C(O)-NR_aR_b, donde R_a y R_b son iguales o diferentes y son independientemente un alquilo, alquenilo, alquinilo, arilo, heteroarilo, como se define en la presente descripción, o R_aR_b pueden tomarse conjuntamente para formar un grupo cicloalquilo o grupo heterociclilo, como se define en la presente descripción. A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo alquilcarbamoilo puede sustituirse opcionalmente.

“Alquilcarboxamidilo” o “grupo alquilcarboxamidilo” se refiere al grupo -C(O)-NR_aR_b, donde R_a y R_b son iguales o diferentes y son independientemente un grupo alquilo, alquenilo, alquinilo, arilo, heteroarilo, cicloalquilo, cicloalquenilo, cicloalquinilo o heterociclilo, como se define en la presente descripción, o R_aR_b pueden tomarse conjuntamente para formar un grupo cicloalquilo, como se define en la presente descripción. A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo alquilcarboxamidilo puede sustituirse opcionalmente.

“Arilo” se refiere a un radical de sistema de anillo hidrocarbonado que comprende hidrógeno, de 6 a 18 átomos de carbono, y al menos un anillo aromático. Para los propósitos de esta invención, el radical arilo puede ser un sistema de anillo monocíclico, bicíclico, tricíclico o tetracíclico, que puede incluir sistemas de anillos fusionados o con puente. Los radicales arilo incluyen, pero no se limitan a, radicales arilo derivados de aceantrileno, acenaftileno, acefenantrileno, antraceno, azuleno, benceno, criseno, fluoranteno, fluoreno, *as*-indaceno, *s*-indaceno, indano, indeno, naftaleno, fenaleno, fenantreno, pleiadenio, pireno y trifenileno. A menos que se indique de cualquier otra manera específicamente en la descripción, el término “arilo” pretende incluir radicales arilo que se sustituyen opcionalmente.

“Heteroarilo” se refiere a un radical de sistema de anillo de 5 a 20 miembros que comprende átomos de hidrógeno, de uno a trece átomos de carbono, de uno a seis heteroátomos seleccionados de

nitrógeno, oxígeno y azufre, y al menos un anillo aromático. Para los propósitos de esta invención, el radical heteroarilo puede ser un sistema de anillo monocíclico, bicíclico, tricíclico o tetracíclico, que puede incluir sistemas de anillos fusionados o con puente; y los átomos de nitrógeno, carbono o azufre en el radical heteroarilo pueden oxidarse opcionalmente; el átomo de nitrógeno puede cuaternizarse opcionalmente. Los ejemplos incluyen, pero no se limitan a, azepinilo, acridinilo, benzimidazolilo, benzotiazolilo, benzindolilo, benzodioxolilo, benzofuranilo, benzooxazolilo, benzotiazolilo, benzotiodiazolilo, benzo[b][1,4]dioxepinilo, 1,4-benzodioxanilo, benzonaftofuranilo, benzoxazolilo, benzodioxolilo, benzodioxinilo, benzopiranilo, benzopiranonilo, benzofuranilo, benzofuranonilo, benzotienilo (benzotiofenilo), benzotriazolilo, benzo[4,6]imidazo[1,2-a]piridinilo, carbazolilo, cinolinilo, dibenzofuranilo, dibenzotiofenilo, furanilo, furanonilo, isotiazolilo, imidazolilo, indazolilo, indolilo, indazolilo, isoindolilo, indolinilo, isoindolinilo, isoquinolilo, indolizínilo, isoxazolilo, naftiridinilo, oxadiazolilo, 2-oxoazepinilo, oxazolilo, oxiranilo, 1-oxidopiridinilo, 1-oxidopirimidinilo, 1-oxidopirazinilo, 1-oxidopiridazinilo, 1-fenil-1*H*-pirrolilo, fenazinilo, fenotiazinilo, fenoxazinilo, ftalazinilo, pteridinilo, purinilo, pirrolilo, pirazolilo, piridinilo, pirazinilo, pirimidinilo, piridazinilo, quinazolinilo, quinoxalinilo, quinolinilo, quinuclidinilo, isoquinolinilo, tetrahydroquinolinilo, tiazolilo, tiadiazolilo, triazolilo, tetrazolilo, triazinilo y tiofenilo (es decir, tienilo). A menos que se indique de cualquier otra manera específicamente en la descripción, un grupo heteroarilo puede sustituirse opcionalmente.

El término “sustituido” usado en la presente descripción se refiere a cualquiera de los grupos anteriores (es decir, alquilo, alquenilo, alquinilo, cicloalquilo, cicloalquenilo, cicloalquinilo, heterociclilo, arilo, heteroarilo, alcoxi, ariloxi, acilo, alquilcarbamoilo, alquilcarboxamidilo, alcóxicarbonilo, alquiltio o ariltio) en donde al menos un átomo se reemplaza por átomos que no son de hidrógeno tales como, pero no limitados a: un átomo de halógeno tal como F, IC, Br y I; un átomo de oxígeno en grupos tales como grupos hidroxilo, grupos alcoxi y grupos ésteres; un átomo de azufre en grupos tales como grupos tiol, grupos tioalquilo, grupos sulfona, grupos sulfonilo y grupos sulfóxidos; un átomo de nitrógeno en grupos tales como aminas, amidas, alquilaminas, dialquilaminas, arilaminas, alquilarilaminas, diarilminas, N-óxidos, imidas y enaminas; un átomo de silicio en grupos tales como grupos trialquilsililo, grupos dialquilarilsililo, grupos alquildiarilsililo y grupos triarilsililo; y otros heteroátomos en diversos otros grupos. “Sustituido” también significa cualquiera de los grupos anteriores en los que uno o más átomos se reemplazan por un enlace de orden superior (por ejemplo, un doble o triple enlace) a un heteroátomo tal como oxígeno en los grupos oxo, carbonilo, carboxilo y éster; y nitrógeno en grupos tales como los iminos, oximas,

hidrazonas y nitrilos. Por ejemplo, “sustituido” incluye cualquiera de los grupos anteriores en los que uno o más átomos se reemplazan con $-\text{NR}_g\text{R}_h$, $-\text{NR}_g\text{C}(=\text{O})\text{R}_h$, $-\text{NR}_g\text{C}(=\text{O})\text{NR}_g\text{R}_h$, $-\text{NR}_g\text{C}(=\text{O})\text{OR}_h$, $-\text{NR}_g\text{SO}_2\text{R}_h$, $-\text{OC}(=\text{O})\text{NR}_g\text{R}_h$, $-\text{OR}_g$, $-\text{SR}_g$, $-\text{SOR}_g$, $-\text{SO}_2\text{R}_g$, $-\text{OSO}_2\text{R}_g$, $-\text{SO}_2\text{OR}_g$, $=\text{NSO}_2\text{R}_g$ y $-\text{SO}_2\text{NR}_g\text{R}_h$. “Sustituido” también significa cualquiera de los grupos anteriores en los que uno o más átomos de hidrógeno se reemplazan con $-\text{C}(=\text{O})\text{R}_g$, $-\text{C}(=\text{O})\text{OR}_g$, $-\text{C}(=\text{O})\text{NR}_g\text{R}_h$, $-\text{CH}_2\text{SO}_2\text{R}_g$, $-\text{CH}_2\text{SO}_2\text{NR}_g\text{R}_h$. En lo anterior, R_g y R_h son iguales o diferentes e independientemente hidrógeno, alquilo, alquenilo, alquinilo, alcoxi, alquilamino, tioalquilo, arilo, aralquilo, cicloalquilo, cicloalquenilo, cicloalquinilo, cicloalquilalquilo, haloalquilo, haloalquenilo, haloalquinilo, heterociclilo, *N*-heterociclilo, heterocicliclilalquilo, heteroarilo, *N*-heteroarilo y/o heteroarilalquilo. “Sustituido” significa además cualquiera de los grupos anteriores en los que uno o más átomos se reemplazan por un grupo amino, ciano, hidroxilo, imino, nitro, oxo, tioxo, halo, alquilo, alquenilo, alquinilo, alcoxi, alquilamino, tioalquilo, arilo, aralquilo, cicloalquilo, cicloalquenilo, cicloalquinilo, cicloalquilalquilo, haloalquilo, haloalquenilo, haloalquinilo, heterociclilo, *N*-heterociclilo, heterocicliclilalquilo, heteroarilo, *N*-heteroarilo y/o heteroarilalquilo. “Sustituido” también puede significar un aminoácido en el que uno o más átomos en la cadena lateral se reemplazan por alquilo, alquenilo, alquinilo, acilo, alquilcarboxamidilo, alcoxicarbonilo, carbociclilo, heterociclilo, arilo o heteroarilo. Además, cada uno de los sustituyentes anteriores también puede sustituirse opcionalmente con uno o más de los sustituyentes anteriores.

Como se usa en la presente, el símbolo “ $\overset{\cdot\cdot\cdot}{\underset{\cdot\cdot\cdot}{|}}$ ” (en lo adelante puede denominarse “un punto de enlace de unión”) denota un enlace que es un punto de unión entre dos entidades químicas, una de las cuales se representa como unida al punto de enlace de unión y la otra no se representa como unida al punto de enlace de unión. Por ejemplo, “ $\text{XY}\overset{\cdot\cdot\cdot}{\underset{\cdot\cdot\cdot}{|}}$ ” indica que la entidad química “XY” se une a otra entidad química mediante el punto de enlace de unión. Además, el punto específico de unión a la entidad química no representada puede especificarse por inferencia. Por ejemplo, el compuesto $\text{CH}_3\text{-R}^3$, en donde R^3 es H o “ $\text{XY}\overset{\cdot\cdot\cdot}{\underset{\cdot\cdot\cdot}{|}}$ ” infiere que cuando R^3 es “XY”, el punto de enlace de unión es el mismo enlace que el enlace mediante el cual R^3 se representa como unido a CH_3 .

Como se usa en la presente, por “sujeto” se entiende un individuo. Por tanto, el “sujeto” puede incluir animales domesticados (por ejemplo, gatos, perros, etc.), ganado (por ejemplo, bobinos, caballos,

cerdos, ovejas, cabras, etc.), animales de laboratorio (por ejemplo, ratón, conejo, rata, cobayos, etc.) y aves. “Sujeto” también puede incluir un mamífero, tal como un primate o un ser humano. Por tanto, el sujeto puede ser un paciente humano o veterinario. El término “paciente” se refiere a un sujeto bajo el tratamiento de un clínico, por ejemplo, un médico.

5 Los términos “inhibir” o “inhibición” se refieren a una disminución en una actividad, expresión, función u otro parámetro biológico y pueden incluir, pero no requieren la ablación completa de la actividad, expresión, función u otro parámetro biológico. La inhibición puede incluir, por ejemplo, al menos aproximadamente una reducción del 10 % en la actividad, respuesta, afección o enfermedad en comparación con un control. En modalidades, la expresión, actividad o función de un gen o
10 proteína se disminuye en una cantidad estadísticamente significativa. En modalidades, la actividad o función se disminuye en al menos aproximadamente 10 %, aproximadamente 20 %, aproximadamente 30 %, aproximadamente 40 %, aproximadamente 50 % y hasta aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente 90 % o aproximadamente 100 %.

15 Por “reducir” u otras formas de la palabra, tales como “reducción”, se entiende disminuir un evento o característica (por ejemplo, crecimiento tumoral). Se entiende que esto está típicamente en relación con algún estándar o valor esperado, que en otras palabras es relativo, pero que no siempre es necesario que se haga referencia al estándar o valor relativo. Por ejemplo, “reduce el crecimiento tumoral” significa reducir la velocidad de crecimiento de un tumor con relación a un estándar o un
20 control (por ejemplo, un tumor no tratado).

Como se usa en la presente, “tratar”, “tratamiento” y variantes de estos, se refiere a cualquier administración de los compuestos descritos que parcial o completamente mitiga, mejora, alivia, inhibe, retrasa el inicio, reduce la gravedad y/o reduce la incidencia de uno o más síntomas o características de una enfermedad como se describe en la presente descripción. En referencia a un
25 paciente, el término “tratamiento” se refiere al tratamiento médico de un paciente con la intención de curar, mejorar, estabilizar o prevenir una enfermedad, condición patológica o trastorno. Este término incluye el tratamiento activo, es decir, el tratamiento dirigido específicamente a la mejora de una enfermedad, condición patológica o trastorno, y también incluye el tratamiento causal, es decir, el tratamiento dirigido a la eliminación de la causa de la enfermedad, condición patológica o
30 trastorno asociados. Además, este término incluye el tratamiento paliativo, es decir, el tratamiento diseñado para el alivio de los síntomas en lugar de curar la enfermedad, condición patológica o trastorno; el tratamiento preventivo, es decir, el tratamiento dirigido a minimizar o inhibir parcial o

completamente el desarrollo de la enfermedad, condición patológica o trastorno asociados; y el tratamiento de apoyo, es decir, el tratamiento empleado para complementar otra terapia específica dirigida a la mejora de la enfermedad, condición patológica o trastorno asociados.

El término “terapéuticamente efectiva” se refiere a la cantidad del compuesto y/o composición descrita usada que es una cantidad suficiente para mejorar una o más causas o síntomas de una enfermedad o trastorno. Dicha mejora solo requiere una reducción o alteración, no necesariamente la eliminación.

El término “farmacéuticamente aceptable” se refiere a aquellos compuestos, materiales, composiciones y/o formas de dosificación que, dentro del alcance del buen juicio médico, son adecuados para su uso en contacto con los tejidos de seres humanos y/o animales sin toxicidad indebida, irritación, respuesta alérgica u otros problemas o complicaciones en consonancia con una relación beneficio/riesgo razonable.

El término “portador” significa un compuesto, composición, sustancia o estructura que, en combinación con un compuesto o composición de la presente descripción, ayuda o facilita la preparación, almacenamiento, administración, suministro, efectividad, selectividad o cualquier otra característica del compuesto o composición para su uso o propósito previsto, o sus combinaciones. Por ejemplo, un portador puede seleccionarse para minimizar cualquier degradación del ingrediente activo y para minimizar cualquier efecto secundario adverso en el sujeto.

Como se usa en la presente, el término "portador farmacéuticamente aceptable" se refiere a un portador adecuado para la administración a un paciente. Un portador farmacéutico puede ser una sustancia que ayuda o facilita la preparación, almacenamiento, administración, suministro, efectividad, selectividad, o cualquier otra característica del compuesto o composición de la presente descripción para su uso o propósito previsto, o sus combinaciones. Por ejemplo, se puede seleccionar un portador para reducir la degradación del compuesto o para reducir los efectos secundarios adversos en el paciente. En modalidades, un portador farmacéuticamente aceptable puede ser una solución, dispersión, suspensión o emulsión acuosa o no acuosa estéril, así como también polvos estériles para su reconstitución en soluciones o dispersiones inyectables estériles justo antes del uso. Los ejemplos de portadores, diluyentes, solventes o vehículos acuosos y no acuosos adecuados incluyen agua, etanol, polioles (tales como glicerol, propilenglicol, polietilenglicol y similares), carboximetilcelulosa y mezclas adecuadas de estos, aceites vegetales (tales como aceite de oliva) y ésteres orgánicos inyectables tales como el oleato de etilo. La fluidez adecuada puede mantenerse, por ejemplo, mediante el uso de materiales de recubrimiento tales

como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de las dispersiones y mediante el uso de surfactantes. Estas composiciones también pueden contener adyuvantes tales como conservantes, agentes humectantes, agentes emulsionantes y agentes dispersantes. La prevención de la acción de microorganismos puede garantizarse mediante la inclusión de diversos agentes antibacterianos y antifúngicos tales como parabeno, clorobutanol, fenol, ácido sórbico, y similares. También puede ser conveniente incluir agentes isotónicos tales como azúcares, cloruro de sodio, y similares. Las formulaciones inyectables pueden esterilizarse, por ejemplo, por filtración a través de un filtro de retención de bacterias, o mediante la incorporación de agentes esterilizadores en forma de composiciones sólidas estériles que pueden disolverse o dispersarse en agua estéril u otro medio inyectable estéril antes del uso. Los portadores inertes adecuados pueden incluir azúcares tales como la lactosa.

El término “sales farmacéuticamente aceptables” incluye aquellas obtenidas al hacer reaccionar el compuesto activo que funciona como una base, con un ácido inorgánico u orgánico para formar una sal, por ejemplo, sales de ácido clorhídrico, ácido sulfúrico, ácido fosfórico, ácido metanosulfónico, ácido alcanforsulfónico, ácido oxálico, ácido maleico, ácido succínico, ácido cítrico, ácido fórmico, ácido bromhídrico, ácido benzoico, ácido tartárico, ácido fumárico, ácido salicílico, ácido mandélico, ácido carbónico, etc. Los expertos en la técnica reconocerán además que las sales de adición de ácido pueden prepararse mediante la reacción de los compuestos con el ácido inorgánico u orgánico apropiado, mediante cualquiera de una serie de métodos conocidos. El término “sales farmacéuticamente aceptables” también incluye aquellas obtenidas al hacer reaccionar el compuesto activo que funciona como un ácido, con una base inorgánica u orgánica para formar una sal, por ejemplo, sales de etilendiamina, N-metil-glucamina, lisina, arginina, ornitina, colina, N,N'-dibenciletildiamina, cloroprocaina, dietanolamina, procaina, N-bencilfenetilamina, dietilamina, piperazina, tris-(hidroximetil)-aminometano, hidróxido de tetrametilamonio, trietilamina, dibencilamina, efenamina, deshidroabietilamina, N-etilpiperidina, bencilamina, tetrametilamonio, tetraetilamonio, metilamina, dimetilamina, trimetilamina, etilamina, aminoácidos básicos, y similares. Los ejemplos no limitantes de sales inorgánicas o metálicas incluyen sales de litio, sodio, calcio, potasio, magnesio, y similares.

Como se usa en la presente, el término "administración parenteral" se refiere a la administración mediante inyección o infusión. La administración parenteral incluye, pero no se limita a, la administración subcutánea, la administración intravenosa o la administración intramuscular.

Como se usa en la presente, el término "administración subcutánea" se refiere a la administración

justo por debajo de la piel. "Administración intravenosa" significa la administración en una vena.

Como se usa en la presente, el término "dosis" se refiere a una cantidad específica de un agente farmacéutico proporcionada en una única administración. En modalidades, una dosis puede administrarse en dos o más bolos, comprimidos o inyecciones. En modalidades, cuando se desea la administración subcutánea, la dosis deseada requiere un volumen que no se satisface fácilmente mediante una única inyección. En dichas modalidades, pueden usarse dos o más inyecciones para lograr la dosis deseada. En modalidades, una dosis puede administrarse en dos o más inyecciones para reducir la reacción en el sitio de inyección en un paciente.

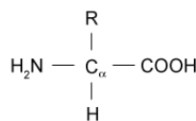
Como se usa en la presente, el término "unidad de dosificación" se refiere a una forma en la que se proporciona un agente farmacéutico. En modalidades, una unidad de dosificación es un vial que incluye los compuestos o composiciones liofilizados descritos en la presente descripción. En modalidades, una unidad de dosificación es un vial que incluye los compuestos o composiciones reconstituidos descritos en la presente descripción.

El término "resto terapéutico" (TM) se refiere a un compuesto que puede usarse para tratar al menos un síntoma de una enfermedad o trastorno, y puede incluir, pero no se limita a, polipéptidos, oligonucleótidos, moléculas pequeñas y otros agentes terapéuticos que pueden usarse para tratar al menos un síntoma de una enfermedad o trastorno. En modalidades, el TM modula la actividad, expresión y/o niveles de un transcrito objetivo. En modalidades, el TM disminuye los niveles del transcrito objetivo a través de mecanismos de degradación. En modalidades, la actividad es la capacidad del transcrito objetivo para unirse a (por ejemplo, secuestrar) una o más proteínas. En modalidades, el TM modula la actividad del transcrito objetivo al reducir la afinidad entre el transcrito objetivo y una o más proteínas que se unen al transcrito objetivo. Como resultado de reducir la afinidad entre el transcrito objetivo y la una o más proteínas, puede modularse la actividad de la una o más proteínas. Por ejemplo, si la una o más proteínas no se unen al transcrito objetivo, están disponibles para llevar a cabo sus funciones, tales como, por ejemplo, facilitar el corte y empalme, el corte y empalme alternativo y/o la omisión de exones de otros transcritos. Como resultado de la función de un TM, puede modularse la actividad, expresión y/o niveles de los genes aguas abajo que están regulados por la una o más proteínas cuya interacción con el transcrito objetivo se interrumpe por el TM.

Los términos "modular" y "modulación" se refieren a una perturbación de la expresión, función o actividad en comparación con el nivel de expresión, función o actividad antes de la modulación. La modulación puede incluir un aumento (estimulación o inducción) o una disminución (inhibición o

reducción) en la expresión, función o actividad. En modalidades, se modula la actividad de un transcrito objetivo. En modalidades, modular la actividad del transcrito objetivo incluye disminuir la capacidad del transcrito objetivo para unirse a una o más proteínas. En modalidades, disminuir la afinidad entre el transcrito objetivo y la una o más proteínas da como resultado la modulación de la actividad de la una o más proteínas que interactúan con el transcrito objetivo. Por ejemplo, si la una o más proteínas no se unen al transcrito objetivo, están disponibles para llevar a cabo sus funciones, tales como, por ejemplo, facilitar el corte y empalme, el corte y empalme alternativo y/o la omisión de exones de otros transcritos (por ejemplo, transcritos aguas abajo). Como tal, modular la actividad del transcrito objetivo puede dar como resultado la modulación de la actividad, expresión y/o niveles de los genes aguas abajo que están regulados por la una o más proteínas cuya interacción con el transcrito objetivo puede interrumpirse.

“Aminoácido” se refiere a un compuesto orgánico que incluye un grupo amino y un grupo ácido



carboxílico y tiene la fórmula general donde R puede ser cualquier grupo orgánico.

Un aminoácido puede ser un aminoácido de origen natural o un aminoácido de origen no natural. Un aminoácido puede ser un aminoácido proteogénico o un aminoácido no proteogénico. Un aminoácido puede ser un L-aminoácido o un D-aminoácido. El término "cadena lateral de aminoácido" o "cadena lateral" se refiere al sustituyente caracterizante (“R”) unido al carbono α de un α -aminoácido natural o no natural. Un aminoácido puede incorporarse en un polipéptido mediante un enlace peptídico.

Como se usa en la presente, un aminoácido “no cargado” es un aminoácido que tiene una cadena lateral que tiene una carga neutra neta a pH 7.35 a 7.45. Los ejemplos de aminoácidos no cargados incluyen, pero no se limitan a, glicina y citrulina.

Como se usa en la presente, un aminoácido “cargado” es un aminoácido que tiene una cadena lateral que tiene una carga neta a un pH de 7.35 a 7.45. Un ejemplo de un aminoácido cargado es la arginina. Como se usa en la presente, el término “identidad de secuencia” se refiere al porcentaje de ácidos nucleicos o aminoácidos entre dos secuencias de oligonucleótidos o polipéptidos, respectivamente, que son iguales y están en la misma posición relativa. Como tal, una secuencia tiene un cierto porcentaje de identidad de secuencia en comparación con otra secuencia. Para la comparación de secuencias, una secuencia actúa típicamente como una secuencia de referencia, a la que se comparan las secuencias de prueba. Los expertos en la técnica apreciarán que dos secuencias se consideran

generalmente “sustancialmente idénticas” si contienen residuos idénticos en las posiciones correspondientes. En modalidades, la identidad de secuencia entre secuencias puede determinarse mediante el uso del algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) según se implementó en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice y otros, Trends Genet. (2000), 16: 276-277), en la versión que existe a la fecha de la presentación. Los parámetros usados son la penalización por apertura de hueco de 10, la penalización por extensión de hueco de 0.5 y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). La salida de Needle etiquetada como “identidad más larga” (obtenida mediante el uso de la opción –nobrief) se usa como el por ciento de identidad y se calcula de la siguiente manera: $(\text{residuos idénticos} \times 100) / (\text{longitud de alineación} - \text{número total de huecos en la alineación})$

En otras modalidades, la identidad de secuencia puede determinarse mediante el uso del algoritmo Smith-Waterman, en la versión que existe a la fecha de la presentación.

Como se usa en la presente, “homología de secuencia” se refiere al porcentaje de aminoácidos entre dos secuencias polipeptídicas que son homólogas y están en la misma posición relativa. Como tal, una secuencia polipeptídica tiene un cierto porcentaje de homología de secuencia en comparación con otra secuencia polipeptídica. Como apreciarán los expertos en la técnica, dos secuencias se consideran generalmente que son “sustancialmente homólogas” si contienen residuos homólogos en las posiciones correspondientes. Los residuos homólogos pueden ser residuos idénticos. Alternativamente, los residuos homólogos pueden ser residuos no idénticos con características estructurales y/o funcionales apropiadamente similares. Por ejemplo, como conocen bien los expertos en la técnica, ciertos aminoácidos se clasifican típicamente como aminoácidos “hidrófobos” o “hidrófilos”, y/o que tienen cadenas laterales “polares” o “no polares”, y la sustitución de un aminoácido por otro del mismo tipo puede considerarse con frecuencia una sustitución “homóloga”.

Como se conoce bien en esta técnica, las secuencias de aminoácidos pueden compararse mediante el uso de cualquiera de una variedad de algoritmos, lo que incluye los disponibles en programas de computadora comerciales tales como BLASTP, gapped BLAST y PSI-BLAST, en existencia a la fecha de la presentación. Dichos programas se describen en Altschul, y otros, *J. Mol. Biol.*, (1990), 215(3): 403-410; Altschul, y otros, *Nucleic Acids Res.* (1997), 25:3,389-3,402; Baxevanis y otros, *Bioinformatics A Practical Guide to the Analysis of Genes and Proteins*, Wiley, 1998; y Misener, y otros, (eds.), *Bioinformatics Methods and Protocols (Methods in Molecular Biology, Vol. 132)*, Humana Press, 1999. Además de identificar secuencias homólogas, los programas

mencionados anteriormente proporcionan típicamente una indicación del grado de homología.

Como se usa en la presente, “resto de direccionamiento celular” se refiere a una molécula o macromolécula que se une específicamente a una molécula, tal como un receptor, en la superficie de una célula objetivo. En modalidades, la molécula de la superficie celular se expresa solo en la superficie de una célula objetivo. En modalidades, la molécula de la superficie celular también está presente en la superficie de una o más células no objetivo, pero la cantidad de la expresión de la molécula de la superficie celular es mayor en la superficie de las células objetivo. Los ejemplos de un resto de direccionamiento celular incluyen, pero no se limitan a, un anticuerpo, un péptido, una proteína, un aptámero o una molécula pequeña.

Como se usa en la presente, los términos "compuesto antisentido" y "AC" se usan indistintamente para referirse a una estructura de ácido nucleico polimérica que es al menos parcialmente complementaria a una molécula de ácido nucleico objetivo a la que se hibrida (el AC). El AC puede ser un polinucleótido u homólogo de polinucleótido corto (en modalidades, menos de 50 bases) que incluye al menos una porción de una secuencia complementaria a una secuencia objetivo. En modalidades, el AC es un polinucleótido u homólogo de polinucleótido que incluye una porción que tiene una secuencia complementaria a una secuencia objetivo en una hebra de pre-ARNm objetivo. El AC puede formarse de ácidos nucleicos naturales, ácidos nucleicos sintéticos, homólogos de ácidos nucleicos o cualquiera de sus combinaciones. En modalidades, el AC incluye oligonucleósidos. En modalidades, el AC incluye oligonucleótidos antisentido. En modalidades, el AC incluye grupos conjugados. Los ejemplos no limitantes de AC incluyen, pero no se limitan a, cebadores, sondas, oligonucleótidos antisentido, oligonucleótidos de secuencia guía externa (EGS), ARNip, oligonucleótidos, oligonucleósidos, análogos de oligonucleótidos, miméticos de oligonucleótidos y combinaciones químicas de estos. Como tal, estos compuestos pueden introducirse en forma de monocatenarios, bicatenarios, circulares, ramificados o en horquillas y pueden contener elementos estructurales tales como abultamientos o lazos internos o terminales. Los compuestos oligoméricos bicatenarios pueden ser dos hebras hibridadas para formar compuestos bicatenarios o una sola hebra con suficiente autocomplementaridad para permitir la hibridación y formación de un compuesto total o parcialmente bicatenario. En modalidades, un AC modula (aumenta, disminuye o cambia) la expresión, los niveles y/o la actividad de un transcrito objetivo (por ejemplo, ácido nucleico objetivo). En modalidades, el AC disminuye el nivel del transcrito objetivo mediante la inducción de mecanismos de degradación. En modalidades, el AC modula la actividad del transcrito objetivo. En modalidades, el AC modula la actividad del transcrito objetivo

al disminuir la capacidad del transcrito objetivo de unirse a una o más proteínas. En modalidades, disminuir la afinidad entre el transcrito objetivo y la una o más proteínas puede dar como resultado la modulación de la actividad de la una o más proteínas. Por ejemplo, si la una o más proteínas no se unen al transcrito objetivo, están disponibles para llevar a cabo sus funciones, tales como, por ejemplo, facilitar el corte y empalme, el corte y empalme alternativo y/o la omisión de exones de otros transcritos (transcritos aguas abajo). Como tal, la modulación mediada por AC de la actividad del transcrito objetivo puede dar como resultado la modulación de la actividad, la expresión y/o los niveles de los genes aguas abajo que están regulados por la una o más proteínas cuya interacción con el transcrito objetivo puede interrumpirse.

Como se usa en la presente, los términos “direccionamiento” o “dirigido a” se refieren a la asociación de un resto terapéutico, por ejemplo, un compuesto antisentido, con una molécula de ácido nucleico objetivo o una región de una molécula de ácido nucleico objetivo. En modalidades, el resto terapéutico incluye un compuesto antisentido que es capaz de hibridarse con un ácido nucleico objetivo en condiciones fisiológicas. En modalidades, el compuesto antisentido se dirige a una porción o sitio específico dentro del ácido nucleico objetivo, por ejemplo, una porción del ácido nucleico objetivo que tiene al menos una estructura, función o característica identificable tal como un exón o intrón en particular, o nucleobases o motivos seleccionados dentro de un exón o intrón.

Como se usa en la presente, los términos “secuencia de ácido nucleico objetivo”, “secuencia de nucleótidos objetivo” y “secuencia objetivo” se refieren a la secuencia de ácido nucleico o la secuencia de nucleótidos a la que se une o se hibrida un resto terapéutico, tal como un compuesto antisentido. Los ácidos nucleicos objetivo incluyen, pero no se limitan a, una porción de un transcrito objetivo, ARN objetivo (lo que incluye, pero no se limita a, pre-ARNm y ARNm o porciones de estos), una porción del ADNc objetivo derivado de dicho ARN, así como también una porción del ARN no traducido objetivo, tal como miARN. Por ejemplo, en modalidades, un ácido nucleico objetivo puede ser una porción de un gen celular objetivo (o el ARNm transcrito a partir de dicho gen) cuya expresión o transcripción se asocia con un trastorno o estado de enfermedad en particular. El término “porción” se refiere a un número definido de nucleótidos contiguos (es decir, enlazados) de un ácido nucleico.

Como se usa en la presente, el término “transcrito” o “transcrito génico” se refiere a una molécula de ARN transcrita a partir de ADN e incluye, pero no se limita a, ARNm, pre-ARNm y ARN parcialmente procesado.

Los términos “transcrito objetivo” y “ARN objetivo” se refieren al pre-ARNm o transcrito de ARNm

al que se une al resto terapéutico. El transcrito objetivo puede incluir una secuencia de nucleótidos objetivo. En modalidades, el transcrito objetivo incluye una secuencia de nucleótidos objetivo que incluye una repetición de trinucleótidos CUG expandida.

El término “gen objetivo” y “gen de interés” se refiere al gen del cual se desea o se pretende modular la expresión y/o actividad. El gen objetivo puede transcribirse en un transcrito objetivo que incluye una secuencia de nucleótidos objetivo. El transcrito objetivo puede traducirse en una proteína de interés.

El término "proteína objetivo" se refiere al polipéptido o proteína codificada por el transcrito objetivo (por ejemplo, ARNm objetivo).

Como se usa en la presente, el término “ARNm” se refiere a una molécula de ARN que codifica una proteína e incluye pre-ARNm y ARNm maduro. "Pre-ARNm" se refiere a una molécula de ARNm eucariota recién sintetizada directamente después de la transcripción del ADN. En modalidades, un pre-ARNm se tapa con una caperuza 5', se modifica con una cola poli-A 3' y/o se empalma para producir una secuencia de ARNm maduro. En modalidades, el pre-ARNm incluye uno o más intrones. En modalidades, el pre-ARNm se somete a un proceso conocido como corte y empalme para eliminar intrones y unir exones. En modalidades, el pre-ARNm incluye uno o más elementos de corte y empalme o elementos reguladores del corte y empalme. En modalidades, el pre-ARNm incluye un sitio de poliadenilación.

Como se usa en la presente, el término “expresión”, "expresión génica", “expresión de un gen”, o similares, se refiere a todas las funciones y etapas mediante las cuales la información codificada en un gen se convierte en un producto génico funcional, tal como un polipéptido o un ARN no codificante, en una célula. Los ejemplos de ARN no codificante incluyen ARN de transferencia (ARNt) y ARN ribosómico. La expresión génica de un polipéptido incluye la transcripción del gen para formar un pre-ARNm, el procesamiento del pre-ARNm para formar un ARNm maduro, la translocación del ARNm maduro del núcleo al citoplasma, la traducción del ARNm maduro al polipéptido y el ensamble del polipéptido codificado. La expresión incluye la expresión parcial. Por ejemplo, la expresión de un gen puede denominarse generación de un transcrito génico. La traducción de un ARNm maduro puede denominarse expresión del ARNm maduro.

Como se usa en la presente, “modulación de la expresión génica” o similares se refiere a la modulación de uno o más de los procesos asociados con la expresión génica. Por ejemplo, la modificación de la expresión génica puede incluir la modificación de una o más de la transcripción del gen, el procesamiento del ARN, la translocación del ARN desde el núcleo al citoplasma y la

traducción del ARNm en una proteína.

Como se usa en la presente, el término "gen" se refiere a una secuencia de ácido nucleico que abarca una región promotora 5' asociada con la expresión del producto génico, y cualquier región de intrones y exones y regiones no traducidas ("UTR") 3' asociadas con la expresión del producto génico.

5 El término "célula inmunitaria" se refiere a una célula de origen hematopoyético y que desempeña un papel en la respuesta inmunitaria. Las células inmunitarias incluyen, pero no se limitan a, linfocitos (por ejemplo, linfocitos B y linfocitos T), células asesinas naturales (NK) y células mieloides. El término "células mieloides" incluye monocitos, macrófagos y granulocitos (por ejemplo, basófilos, neutrófilos, eosinófilos y mastocitos). Los monocitos son linfocitos que circulan
10 a través de la sangre durante 1-3 días, después de lo cual migran hacia los tejidos y se diferencian en macrófagos o células dendríticas inflamatorias o mueren. El término "macrófago", como se usa en la presente, incluye macrófagos derivados del feto (que también pueden denominarse macrófagos residentes de tejido) y macrófagos derivados de monocitos que han migrado desde el torrente sanguíneo hacia un tejido en el cuerpo (que pueden denominarse macrófagos derivados de
15 monocitos). En dependencia de en qué tejido se ubique el macrófago, se denomina célula de Kupffer (hígado), célula mesangial intraglomerular (riñón), un macrófago alveolar (pulmones), un histiocito sinusal (ganglios linfáticos), una célula de hofbauer (placenta), microglia (cerebro y médula espinal) o células de langerhans (piel), entre otros.

Como se usa en la presente, el término "oligonucleótido" se refiere a un compuesto oligomérico que
20 comprende una pluralidad de nucleótidos o nucleósidos enlazados. Pueden modificarse uno o más nucleótidos de un oligonucleótido. Un oligonucleótido puede comprender ácido ribonucleico (ARN) o ácido desoxirribonucleico (ADN). Los oligonucleótidos pueden componerse de nucleobases naturales y/o modificadas, azúcares y enlaces internucleósidos covalentes, y pueden incluir además conjugados sin ácido nucleico.

25 Como se usa en la presente, el término "nucleósido" se refiere a una glicosilamina que incluye una nucleobase y un azúcar. Nucleósidos incluye, pero no se limita a, nucleósidos naturales, nucleósidos sin bases, nucleósidos modificados, y nucleósidos que tienen bases y/o grupos de azúcar miméticos. Un "nucleósido natural" o un "nucleósido no modificado" es un nucleósido que incluye una nucleobase natural y un azúcar natural. Los nucleósidos naturales incluyen nucleósidos de ARN y
30 de ADN.

Como se usa en la presente, el término "azúcar natural" se refiere a un azúcar de un nucleósido que no se modifica a partir de su forma de origen natural en ARN (2'-OH) o ADN (2'-H).

Como se usa en la presente, el término "nucleótido" se refiere a un nucleósido que tiene un grupo fosfato enlazado covalentemente al azúcar. Los nucleótidos pueden modificarse con cualquiera de una variedad de sustituyentes.

Como se usa en la presente, el término "nucleobase" se refiere a la porción de base de un nucleósido o nucleótido. Una nucleobase puede incluir cualquier átomo o grupo de átomos capaces de unirse por hidrógeno a una base de otro ácido nucleico. Una nucleobase natural es una nucleobase que no se modifica a partir de su forma de origen natural en ARN o ADN.

Como se usa en la presente, el término "resto de base heterocíclica" se refiere a una nucleobase que incluye un heterociclo.

Como se usa en la presente, "enlace internucleósido" se refiere a un enlace covalente entre nucleósidos adyacentes.

Como se usa en la presente, "enlace internucleósido natural" se refiere a un enlace fosfodiéster 3' a 5'.

Como se usa en la presente, el término "enlace internucleósido modificado" se refiere a cualquier enlace entre nucleósidos o nucleótidos diferentes de un enlace internucleósido de origen natural.

Como se usa en la presente, "oligonucleósido" se refiere a un oligonucleótido en el que los enlaces internucleósidos no contienen un átomo de fósforo.

Como se usa en la presente, el término "compuesto antisentido quimérico" se refiere a un compuesto antisentido, que tiene al menos un azúcar, nucleobase y/o enlace internucleósido que se modifica diferencialmente en comparación con los otros azúcares, nucleobases y enlaces internucleósidos dentro del mismo compuesto oligomérico. El resto de los azúcares, nucleobases y enlaces internucleósidos pueden modificarse o no modificarse independientemente. En general, un compuesto quimérico oligomérico tendrá nucleósidos modificados que pueden estar en posiciones aisladas o agrupados en regiones que definirán un motivo en particular. Cualquier combinación de modificaciones y/o grupos miméticos puede incluir un compuesto oligomérico quimérico como se describe en la presente descripción.

Como se usa en la presente, el término "oligonucleótido antisentido de cadena principal mezclada" se refiere a un oligonucleótido antisentido en donde al menos un enlace internucleósido del oligonucleótido antisentido es diferente de al menos otro enlace internucleósido del oligonucleótido antisentido.

Como se usa en la presente, el término "complementariedad de nucleobases" se refiere a una nucleobase que es capaz de lograr emparejamiento de bases con otra nucleobase. Por ejemplo, en el

ADN, la adenina (A) es complementaria a la timina (T). Por ejemplo, en el ARN, la adenina (A) es complementaria al uracilo (U). En modalidades, la nucleobase complementaria se refiere a una nucleobase de un compuesto antisentido que es capaz de lograr emparejamiento de bases con una nucleobase de su ácido nucleico objetivo. Por ejemplo, si una nucleobase en una cierta posición de un compuesto antisentido es capaz de unirse por hidrógeno con una nucleobase en una cierta posición de un ácido nucleico objetivo, entonces la posición de unión por hidrógeno entre el oligonucleótido y el ácido nucleico objetivo se considera que es complementaria en ese par de nucleobases.

Como se usa en la presente, el término "nucleobase no complementaria" se refiere a un par de nucleobases que no forman enlaces de hidrógeno entre sí o de cualquier otra manera soportan la hibridación.

Como se usa en la presente, el término "complementario" se refiere a la capacidad de un compuesto oligomérico para hibridar con otro compuesto oligomérico o ácido nucleico a través de la complementariedad de nucleobases. En modalidades, un compuesto antisentido y su objetivo son complementarios entre sí cuando un número suficiente de posiciones correspondientes en cada molécula están ocupadas por nucleobases que pueden unirse entre sí para permitir una asociación estable entre el compuesto antisentido y el objetivo. Un experto en la técnica reconoce que la inclusión de errores de emparejamiento es posible sin eliminar la capacidad de los compuestos oligoméricos de permanecer en asociación. Por lo tanto, en la presente descripción se describen compuestos antisentido que pueden incluir hasta aproximadamente 20 % de nucleótidos con errores de emparejamiento (es decir, no son complementarios en nucleobases a los nucleótidos correspondientes del objetivo). En modalidades, los compuestos antisentido contienen no más de aproximadamente 15 %, por ejemplo, no más de aproximadamente 10 %, por ejemplo, no más de 5 %, o ningún error de emparejamiento. Los nucleótidos restantes son complementarios en nucleobases o de cualquier otra manera no interrumpen la hibridación (por ejemplo, bases universales). Un experto en la técnica reconocería que los compuestos proporcionados en la presente descripción son al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % complementarios en nucleobases a un ácido nucleico objetivo.

Como se usa en la presente, "hibridación" significa el emparejamiento de compuestos oligoméricos complementarios (por ejemplo, un compuesto antisentido y su ácido nucleico objetivo). Aunque no se limita a un mecanismo en particular, el mecanismo más común de emparejamiento involucra la unión por hidrógeno, que puede ser la unión por hidrógeno de Watson-Crick, Hoogsteen o Hoogsteen

invertido, entre bases complementarias de nucleósidos o nucleótidos (nucleobases). Por ejemplo, la base natural adenina es complementaria en nucleobases a las nucleobases naturales timina y uracilo que se emparejan a través de la formación de enlaces de hidrógeno. La base natural guanina es complementaria en nucleobases a las bases naturales citosina y 5-metil citosina. La hibridación puede producirse bajo circunstancias variables.

Como se usa en la presente, el término "hibrida específicamente" se refiere a la capacidad de un compuesto oligomérico para hibridar con un sitio de ácido nucleico con mayor afinidad que con la que hibrida con otro sitio de ácido nucleico. En modalidades, un oligonucleótido antisentido se hibrida específicamente con más de un sitio objetivo. En modalidades, un compuesto oligomérico se hibrida específicamente con su objetivo bajo condiciones de hibridación rigurosas.

Las "condiciones de hibridación rigurosas" y las "condiciones de lavado de hibridación rigurosas" en el contexto de la hibridación de ácidos nucleicos, dependen de la secuencia y son diferentes bajo diferentes parámetros ambientales. Se encuentra una amplia guía para la hibridación de ácidos nucleicos en Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, parte I, capítulo 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, Nueva York (1993). Generalmente, las condiciones de hibridación y lavado altamente rigurosas se seleccionan para que sean aproximadamente 5 °C más bajas que el punto de fusión térmico (T_m) para la secuencia específica a una resistencia iónica y pH definidos. La T_m es la temperatura (bajo la resistencia iónica y el pH definidos) a la que el 50 % de la secuencia objetivo se hibrida con una sonda perfectamente coincidente. Se seleccionan condiciones muy rigurosas para que sean iguales a la T_m para una sonda en particular. Un ejemplo de condiciones de hibridación rigurosas para la hibridación de secuencias de nucleótidos complementarias que tienen más de 100 residuos complementarios en un filtro en una transferencia Southern o Northern es formamida al 50 % con 1 mg de heparina a 42 °C, donde la hibridación se lleva a cabo durante la noche. Un ejemplo de condiciones de lavado altamente rigurosas es NaCl 0.15 M a 72 °C durante aproximadamente 15 minutos. Un ejemplo de condiciones de lavado rigurosas es un lavado de SSC 0.2x a 65 °C durante 15 minutos (ver, Sambrook y Russel, Molecular Cloning: A laboratory Manual, 3ª ed., Cold Spring Harbor Laboratory Press, 2001, para una descripción del tampón de SSC). Frecuentemente, un lavado de rigurosidad alta es precedido por un lavado de rigurosidad baja para eliminar la señal de fondo de la sonda. Un ejemplo de un lavado de rigurosidad media para un dúplex de, por ejemplo, más de 100 nucleótidos, es SSC 1x a 45 °C durante 15 minutos. Un ejemplo de un lavado de rigurosidad baja para un dúplex de, por ejemplo,

más de 100 nucleótidos, es SSC 4-6x a 40 °C durante 15 minutos. Para sondas cortas (por ejemplo, aproximadamente 10 a 50 nucleótidos), las condiciones rigurosas implican típicamente concentraciones de sal de menos de aproximadamente 1.0 M de iones de Na, típicamente aproximadamente 0.01 a 1.0 M de concentración de iones de Na (u otras sales) a pH 7.0 a 8.3, y la temperatura es típicamente al menos aproximadamente 30 °C. Las condiciones rigurosas también pueden lograrse con la adición de agentes desestabilizantes tales como la formamida.

Como se usa en la presente, el término "modificado en 2'" o "sustituido en 2'" significa un azúcar que incluye un sustituyente en la posición 2' distinto de H o OH. Monómeros modificados en 2' incluyen, pero no se limitan a, BNA y monómeros (por ejemplo, nucleósidos y nucleótidos) con sustituyentes en 2', tales como alilo, amino, azido, tio, O-alilo, O-alquilo C1-C10, -OCF3, O-(CH2)2-O-CH3, 2'-O(CH2)2SCH3, O-(CH2)2-O-N(Rm)(Rn) o O-CH2-C(=O)-N(Rm)(Rn), donde están Rm y Rn es, independientemente, H o alquilo C1-C10 sustituido o no sustituido.

Como se usa en la presente, el término "MOE" se refiere a un sustituyente 2'-O-metoxietilo.

Como se usa en la presente, el término "nucleótido modificado de alta afinidad" se refiere a un nucleótido que tiene al menos una nucleobase, enlace internucleósido o resto de azúcar modificado, de manera que la modificación aumenta la afinidad de un compuesto antisentido que incluye el nucleótido modificado por un ácido nucleico objetivo. Las modificaciones de alta afinidad incluyen, pero no se limitan a, BNA, LNA y 2'-MOE.

Como se usa en la presente, el término "mimético" se refiere a grupos que se sustituyen por un azúcar, una nucleobase y/o un enlace de internucleósido en un AC. Generalmente, se usa un mimético en lugar del azúcar o la combinación azúcar-enlace internucleósido, y se mantiene la nucleobase para la hibridación a un objetivo seleccionado. Los ejemplos representativos de un mimético de azúcar incluyen, pero no se limitan a, ciclohexenilo o morfolino. Los ejemplos representativos de un mimético para una combinación de azúcar-enlace internucleósido incluyen, pero no se limitan a, ácidos nucleicos peptídicos (PNA) y grupos morfolino enlazados mediante enlaces aquirales no cargados. En algunos casos, se usa un mimético en lugar de la nucleobase. Los miméticos de nucleobases representativos se conocen bien en la técnica e incluyen, pero no se limitan a, análogos de fenoxazina tricíclicos y bases universales (Berger y otros, Nuc Acid Res. 2000, 28:2,911-14, incorporado en la presente descripción como referencia). Los métodos de síntesis de miméticos de azúcar, nucleósidos y nucleobases se conocen bien por los expertos en la técnica.

Como se usa en la presente, el término "nucleósido bicíclico" o "BNA" se refiere a un nucleósido en donde la porción de furanosa del nucleósido incluye un puente que conecta dos átomos en el anillo

de furanosa, formando de esta manera un sistema de anillo bicíclico. Los BNA incluyen, pero no se limitan a, α -L-LNA, β -D-LNA, ENA, oxiamino BNA (2'-O-N(CH₃)-CH₂-4') y aminooxi BNA (2'-N(CH₃)-O-CH₂-4').

Como se usa en la presente, el término "nucleósido bicíclico 4' a 2'" se refiere a un BNA en donde el puente que conecta dos átomos del anillo de furanosa enlaza el átomo de carbono 4' y el átomo de carbono 2' del anillo de furanosa, formando de esta manera un sistema de anillo bicíclico.

Como se usa en la presente, un "ácido nucleico bloqueado" o "LNA" se refiere a un nucleótido modificado de manera que el grupo hidroxilo 2' del anillo de azúcar ribosilo se enlaza al átomo de carbono 4' del anillo de azúcar mediante un grupo metileno, formando de esta manera un enlace 2'-C,4'-C-oximetileno. Los LNA incluyen, pero no se limitan a, α -L-LNA y β -D-LNA.

Como se usa en la presente, el término "estructura de caperuza" o "resto de caperuza terminal", se refiere a modificaciones químicas que se han incorporado en cualquier extremo de un AC. El término "polipéptido terapéutico" se refiere a una macromolécula de origen natural o producida de manera recombinante que incluye dos o más aminoácidos y tiene actividad terapéutica, profiláctica u otra actividad biológica.

El término "molécula pequeña" se refiere a un compuesto orgánico con actividad farmacológica y un peso molecular de menos de aproximadamente 2,000 daltons, o menor que aproximadamente 1,000 daltons, o menor que aproximadamente 500 daltons. Los agentes terapéuticos de molécula pequeña se fabrican típicamente mediante síntesis química.

"Proteína objetivo de tipo salvaje" se refiere a un isómero de la proteína funcional nativa producido por una versión de tipo salvaje, normal o no mutada del gen objetivo. La proteína objetivo de tipo salvaje también se refiere a una proteína resultante de un pre-ARNm objetivo que se ha reempalmado.

Una "proteína objetivo reempalmada", como se usa en la presente, se refiere a la proteína codificada por el ARNm resultante del corte y empalme del pre-ARNm objetivo al cual se hibrida el AC. La proteína objetivo reempalmada puede ser idéntica a una proteína objetivo de tipo salvaje, puede ser homóloga a una proteína objetivo de tipo salvaje, puede ser una variante funcional de una proteína objetivo de tipo salvaje, puede ser una isoforma de una proteína objetivo de tipo salvaje, o puede ser un fragmento activo de una proteína objetivo de tipo salvaje.

Como se usa en la presente, una "repetición de trinucleótidos expandida", tal como una repetición CUG "expandida" o una repetición CTG "expandida", significa que un gen que contiene o codifica la repetición de trinucleótidos contiene un número de trinucleótidos consecutivos repetidos que es

mayor que el presente en un gen de tipo salvaje. Las repeticiones de nucleótidos expandidas pueden escribirse como XXX·NNN o (XXX·NNN), donde XXX se refiere a la repetición en ADN y NNN se refiere a la repetición en ARN que se transcribe a partir de la repetición en ADN. Por ejemplo, la repetición CTG·CUG se refiere a un gen que tiene una repetición CTG en ADN a partir del cual se transcribe un ARN que tiene una repetición CUG. En modalidades, el número de repeticiones en una repetición de trinucleótidos expandida es 5 o más, 10 o más, 15 o más, o 20 o más que el gen de tipo salvaje. En modalidades, la repetición de trinucleótidos expandida incluye 2x, 3x, 4x, 5x, 10x, 20x, 50x o más repeticiones de trinucleótidos que el gen de tipo salvaje. La repetición de trinucleótidos expandida puede dar como resultado una enfermedad en un sujeto que tiene un gen que contiene la repetición de trinucleótidos expandida. Por ejemplo, un sujeto que tiene una repetición CTG expandida en un gen puede padecer DM1 o FECD. En la DM1, el gen DPMK contiene una repetición CTG expandida. Los sujetos que padecen DM1 pueden tener 50 o más repeticiones CTG en la región no traducida (UTR) 3' del gen DPMK, mientras que los sujetos que no padecen la enfermedad tienen típicamente de 5 a 34 repeticiones CTG en la UTR 3' del gen DPMK. En la FECD, el gen TCF4 contiene una repetición CTG expandida. Los sujetos que padecen FECD pueden tener 40 o más repeticiones CTG en un locus CTG18.1 del gen TCF4, mientras que los sujetos que no padecen la enfermedad tienen típicamente 30 o menos repeticiones CTG en el locus CTG18.1 del gen TCF4. El ARNm transcrito a partir de un gen que tiene una repetición CTG expandida tendrá una repetición CUG expandida.

El término “aguas abajo” en la presente descripción, cuando se refiere a un gen, ARNm o proteína, se refiere a un gen, ARNm o proteína que se ve afectado por la unión del AC al nucleótido objetivo (por ejemplo, transcrito objetivo) pero no es el gen, ARNm o proteína correspondiente al nucleótido objetivo. La unión del AC al nucleótido objetivo puede reducir la agregación o secuestro de una proteína de unión al ARN tal como MBNL1 o CUGBP1 en el ARNm acumulado que tiene repeticiones CUG, lo que puede poner a disposición dichas proteínas de unión al ARN para una transcripción, procesamiento del ARN y/o expresión apropiada de productos génicos aguas abajo. Como se usa en la presente, “fragmento funcional” o “fragmento activo” se refiere a una porción de una proteína objetivo eucariota de tipo salvaje que exhibe una actividad, tal como una o más actividades de una proteína objetivo de tipo salvaje de longitud completa, o que posee otra actividad.

En modalidades, una proteína objetivo reempalmada que comparte al menos una actividad biológica de la proteína objetivo de tipo salvaje se considera un fragmento activo de la proteína objetivo de tipo salvaje. La actividad puede ser cualquier porcentaje de actividad (es decir, más o menos) de la

proteína objetivo de tipo salvaje de longitud completa, lo que incluye, pero no se limita a, aproximadamente 1 % de la actividad, aproximadamente 2 %, aproximadamente 3 %, aproximadamente 4 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 20 %, aproximadamente 30 %, aproximadamente 40 %, aproximadamente 50 %, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, aproximadamente 100 %, aproximadamente 200 %, aproximadamente 300 %, aproximadamente 400 %, aproximadamente 500 %, o más actividad (lo que incluye todos los valores e intervalos entre estos valores) en comparación con la proteína objetivo de tipo salvaje. Por tanto, en modalidades, el fragmento activo puede retener al menos una porción de una o más actividades biológicas de la proteína objetivo de tipo salvaje. En modalidades, el fragmento activo puede potenciar una o más actividades biológicas de la proteína objetivo de tipo salvaje.

"Proteína objetivo de tipo salvaje" se refiere a un isómero de la proteína funcional nativa producido por una versión de tipo salvaje, normal o no mutada del gen objetivo. La proteína objetivo de tipo salvaje también se refiere a la proteína resultante de un pre-ARNm objetivo que se ha empalmado apropiadamente.

Como se usa en la presente, los términos "corte y empalme" y "procesamiento" se refieren a la modificación de un pre-ARNm después de la transcripción, en el que se eliminan los intrones y se unen los exones. El corte y empalme se produce en una serie de reacciones que son catalizadas por un complejo ARN-proteína grande compuesto por cinco ribonucleoproteínas nucleares pequeñas (snRNP) denominado como espliceosoma. Dentro de un intrón, se requieren un sitio de empalme 3', un sitio de empalme 5' y un sitio de ramificación para el corte y empalme. Los componentes de ARN de los snRNP interactúan con el intrón y pueden involucrarse en la catálisis.

Como se usa en la presente, el corte y empalme alternativo se refiere al corte y empalme de diferentes combinaciones de exones presentes en un gen, lo que da como resultado la generación de diferentes transcritos de ARNm a partir de un solo gen.

Una "proteína objetivo reempalmada", como se usa en la presente, se refiere a la proteína codificada por el ARNm resultante del corte y empalme del pre-ARNm objetivo al cual se hibrida el AC. La proteína objetivo reempalmada puede ser idéntica a una proteína objetivo de tipo salvaje, puede ser homóloga a una proteína objetivo de tipo salvaje, puede ser una variante funcional de una proteína objetivo de tipo salvaje, o puede ser un fragmento activo de una proteína objetivo de tipo salvaje.

Todas las publicaciones, patentes y solicitudes de patente mencionadas en la descripción son

indicativas del nivel de conocimientos de los expertos en la técnica a los que pertenece esta invención. Todas las publicaciones, patentes y solicitudes de patente se incorporan en la presente descripción como referencia en la misma medida como si cada publicación o solicitud de patente individual se especificara e indicara individualmente para incorporarse como referencia.

5

EJEMPLOS

Ejemplo 1. Evaluación de los PMO y compuestos PMO-EEV para determinar el impacto sobre la formación de focos de ARN y el rescate del corte y empalme en líneas celulares relacionadas con DM1

10 El efecto de un PMO con repetición (CUG)₇ y un compuesto PMO-EEV con repetición (CUG)₇ (A y D en la Tabla 12) sobre la formación de focos de ARN y el rescate del corte empalme se evaluó en líneas celulares relacionadas con DM1. También se incluyó un compuesto PMO-EEV que tiene una secuencia de PMO con error de emparejamiento y un PMO que tiene una secuencia codificada (B y C en la Tabla 12).

15 Experimentación

Células. Los PMO y los PMO-EEV se evaluaron en un modelo de células HeLa DM1 (HeLa-480), que es una línea celular estable con alta carga de repeticiones CUG y defectos de corte y empalme aguas abajo, y mioblastos DM1 derivados de un paciente con DM1 con aproximadamente 2,600 repeticiones CTG y defectos de corte y empalme aguas abajo. HeLa-480 reproduce las marcas
20 distintivas patogénicas de la DM1, lo que incluye los focos ribonucleares de CUG y el corte y empalme erróneos de los pre-ARNm objetivo de los factores de corte y empalme alternativo de tipo muscleblind (MBNL). Se observó que los mioblastos DM1 crecen bastante lentamente (tiempo de duplicación de aproximadamente 7 días) y no se transfectan bien. Las células HeLa de control y las HeLa-480 se trataron con ENDOPORTER sin compuestos adicionales. Las células HeLa-480
25 representan el estado de enfermedad y las células HeLa representan el estado sin enfermedad.

Análisis de los focos de ARN. Las células HeLa-480 o los mioblastos DM1 se trataron con 1 μ M, 3 μ M o 10 μ M de los compuestos A-D (Tabla 12). Todos los compuestos se transfectaron sin un reactivo de transfección o mediante el uso del agente de transfección ENDOPORTER (disponible de GENETOOLS LLC en Philomath, Oregón) diseñado para suministrar PMO cargados
30 naturalmente a las células. Las células se incubaron durante 24 horas y después se fijaron para el análisis cualitativo de los focos de ARN mediante microscopía. Tras la inspección visual cualitativa de las células HeLa-480 tratadas, el compuesto A (PMO-EEV) mostró la menor cantidad de focos de

ARN. No se extrajo ninguna conclusión de los mioblastos DM1.

Análisis del corte y empalme y RT-PCR. Las células HeLa-480 tratadas como se describió anteriormente se cosecharon después de 24 o 48 horas de la incubación y se extrajo el ARN total. La RT-PCR se realizó para medir y/o cuantificar los patrones de corte y empalme de exones afectados por DM en ARN objetivos tales como MBNL1 y CLASP1. Se evaluó el porcentaje de inclusión de exones de interés.

Tabla 12: PMO y PMO-EEV sometidos a prueba en el Ejemplo 1.

ID	Secuencia (todos los PMO)	Pureza	Solvente
A	CAGCAGCAGCAGCAGCAGCAG-clic-K-PEG12-K(FfΦRrRrQ)-PKKKRKV-Ac (SEQ ID NO: 154)-clic-K-PEG12-(SEQ ID NO:78)-(SEQ ID NO: 42)-Ac	91	Solución salina (NaCl al 0.9 %)
B	GTAAGTGTATTTGGTACTTCC-C3-NH2-PEG4-COT-PEG12-PKKKRKV-(FfΦRrRrQ) (SEQ ID NO: 317)-C3-NH2-PEG4-COT-PEG12-(SEQ ID NO: 78)-(SEQ ID NO:42)-Ac	99	PBSX1
C	AGCCAGAGCACCGCAACCGGACGAG (SEQ ID NO: 318)	81	Solución salina (NaCl al 0.9 %)
D	CAGCAGCAGCAGCAGCAGCAG (SEQ ID NO: 154)	75	Solución salina (NaCl al 0.9 %)

Resultados. La Figura 6A-6D muestra los resultados de RT-PCR de los eventos de corte y empalme alternativo de MBNL1 (6A y 6C, inclusión del exón 5) y CLASP1 (6B y 6D, inclusión del exón 19) 24 horas (6A y 6B) y 48 horas (6C y 6D) después de que las células HeLa-480 se trataran con los compuestos A-D y el agente de transfección ENDOPORTER. Se observó una reducción de la inclusión del exón 5 en MBNL1 para las células tratadas con cada uno de los compuestos A-D en los puntos temporales de 24 y 48 horas (Figura 6A y 6C). Las células tratadas con el compuesto A mostraron el mayor rescate (disminución en la inclusión del exón 5). Se observó un aumento en la inclusión del exón 19 en CLASP1 para las células tratadas con cada uno de los compuestos A-D en los puntos temporales de 24 y 48 horas (Figura 6B y 6D). Las células tratadas con el compuesto A mostraron el mayor rescate (aumento en la inclusión del exón 19). Por el contrario, cuando las células HeLa-480 se trataron con A-D sin el reactivo de transfección ENDOPORTER, no se observó ningún cambio en los eventos de corte y empalme de MBNL1 (Figura 6E) o CLASP1 (Figura 6F), excepto para el compuesto A.

La Figura 7A-7B muestra los resultados de RT-PCR de eventos de corte y empalme alternativo de MBNL1 y CLASP1 después de que los mioblastos DM1 se trataran con los compuestos A-D, el control negativo DM-04 o el control positivo DM-05. El tratamiento con todos los compuestos dio como resultado el rescate de los eventos de corte y empalme de MBNL1 (Figura 7A) y CLASP1 (Figura 7B).

Ejemplo 2. Evaluación de EEV-PMO 221-1106 para determinar el impacto sobre rescate del corte y empalme en el modelo de ratón con DM1

El efecto de un PMO solamente y un PMO-EEV (Tabla 13) sobre el rescate del corte y empalme se evaluó in vivo mediante el uso de un modelo de ratón con DM1 HSA-LR.

Experimentación.

Modelo de ratón. HSA-LR es un modelo de ratón transgénico que tiene repeticiones CUG largas expandidas (LR) en la UTR 3' de un transgén de actina esquelética humana (HSA) y expresa ARN CUGexp (por ejemplo, ARN con CUG expandido) a niveles altos en músculo esquelético (Mankodi y otros, Science 2000, 289(5485):1,769-1,773). El ratón HSA-LR muestra el fenotipo miotónico junto con los defectos de corte y empalme. Se usó un modelo de ratón Friend Virus B NIH Jackson (FVB/NJ) como control y para producir los ratones transgénicos HSA-LR.

Diseño experimental. Los compuestos A-D se administraron a los ratones mediante inyección retroorbital o inyección intravenosa (i.v.) a una dosis única de 100 µl de solución del compuesto por 20 g de peso corporal. La escala fue proporcional al peso corporal de cada ratón (por ejemplo, 150 µl por 30 g de peso corporal).

Tabla 13: PMO y PMO-EEV sometidos a prueba en el Ejemplo 2.

ID	Compuesto	Secuencia (todos los PMO)	Con. (mg/ml)	Solvente
(A)	Salina	Vehículo control	NA	Solución salina (NaCl al 0.9 %)
(B)	221	5'-CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154)	6	Solución salina (NaCl al 0.9 %)

(C)	221-1106 20 mpk	a	5'-CAG CAG CAG CAG CAG CAG CAG-3'- PEG4COT-clic-K-miniPEG2- Lys(FfFGRGRE)-miniPEG2-VKRKKKP-Ac (SEQ ID NO: 154)-PEG4COT-clic-K- miniPEG2-K(SEQ ID NO: 76)-miniPEG2-(SEQ ID NO: 42)-Ac	4	Solución salina (NaCl al 0.9 %)
(D)	221-1106 40 mpk	a	5'-CAG CAG CAG CAG CAG CAG CAG-3'- PEG4COT-clic-K-miniPEG2- Lys(FfFGRGRE)-miniPEG2-VKRKKKP-AC (SEQ ID NO: 154)-PEG4COT-clic-K- miniPEG2-K(SEQ ID NO: 76)-miniPEG2-(SEQ ID NO: 42)-Ac	8	Solución salina (NaCl al 0.9 %)

Los animales se emparejaron por edad y se asignaron a seis grupos de tratamiento. Se usaron dos grupos de control; Grupo 1: ratones FVB/NJ (FVB/NJ; control sin enfermedad) y Grupo 2: ratones HSA-LR (control enfermo), cada uno de los cuales se inyectó con solución salina. Se usaron cuatro grupos de tratamiento (Grupos A-D); ratones HSA-LR inyectados con los compuestos A, B, C o D. Los ratones FVB/NJ tenían 5 semanas y 4 días de edad cuando se inyectaron, mientras que los ratones HSA-LR tenían 6 semanas y 1 o 2 días de edad cuando se inyectaron. Para este experimento se utilizaron cuatro ratones (dos machos y dos hembras) por grupo. Los ratones se sacrificaron 1 semana después del tratamiento. Los tejidos (Gastrocnemio, cuádriceps, tibial anterior (TA)) se cosecharon y se congelaron rápidamente en nitrógeno líquido y se almacenaron a -80 °C para la evaluación posterior del análisis del rescate de corte y empalme.

El ARN total se extrajo de las muestras de tejido y se analizó mediante RT-PCR para evaluar los eventos de rescate de corte y empalme alternativo de ARN inducidos por el AC en (i) el exón 22 de Atp2a1, (ii) el exón 7 de Nfix, (iii) el exón 7a de Clcn1 y (iv) el exón 5 de Mbnl1. Se evaluó el porcentaje de inclusión de exones de interés.

Resultados. Antes del tratamiento, todos los ratones HSA-LR tenían miotonía. Después de la inyección de los compuestos, todos los ratones se desorientaron. Los ratones inyectados con los compuestos A y B (Grupos A y B) se recuperaron en 15 minutos, mientras que los ratones inyectados con los compuestos C y D (Grupos C y D) tardaron varias horas en recuperarse. Todos los ratones tratados se recuperaron completamente al día siguiente. En el momento del sacrificio, los Grupos A y B tenían miotonía; sin embargo, los Grupos C y D evidentemente no tenían miotonía.

Las Figuras 8A-10D muestran las mediciones del corte y empalme del ARN para Atp2a1 (para la

inclusión del exón 22; Figuras 8A, 9A y 10A), Nfix (para inclusión del exón 7; Figura 8B, 9B y 10B), Clcn1 (para la inclusión del exón 7a; Figura 8C, 9C y 10C) y Mbnl1 (para la inclusión del exón 5; Figura 8D, 9D y 10D) en el tejido del músculo gastrocnemio (Figuras 8A-8D), cuádriceps (Figura 9A-9D) y tibial anterior (Figura 10A-10D) de ratones tratados. Los ratones tratados con los compuestos C y D (PMO-EEV) mostraron un rescate de los eventos de corte y empalme de Atp2a1 y Nfix en el tejido gastrocnemio, cuádriceps y tibial anterior, mientras que los grupos con PMO y solución salina no mostraron un rescate de los eventos de corte y empalme (Figura 8A, 8B, 9A, 9B, 10A y 10B). De manera similar, en el tejido del músculo gastrocnemio y cuádriceps, los ratones tratados con los compuestos C y D mostraron un rescate de los eventos de corte y empalme de Clcn1 (Figura 8C y 9C) y Mbnl1 (Figura 8D y 9D) mientras que los grupos de PMO y solución salina no mostraron rescate de los eventos de corte y empalme. Con respecto al corte y empalme de Clcn1 y Mbnl1 en el tejido tibial (Figura 10C-10D), no se detectaron defectos de corte y empalme alternativo en la línea de ratón de control (FVB/NJ) y el modelo de ratón con DM1 (HSA-LR). Como tal, el tratamiento con los compuestos A-D no dio como resultado el rescate del corte y empalme en los genes Clcn1 y Mbnl1 en el tejido tibial.

Estos resultados demuestran los impactos positivos del tratamiento con PMO-EEV en el corte y empalme del estudio in vivo de rescate mediante el uso del modelo de ratón con DM1, así como también el posible uso de los compuestos PMO-EEV para tratar la distrofia miotónica (DM).

Ejemplo 3. Evaluación de diversos compuestos PMO-EEV para corregir eventos de corte y empalme erróneos en mioblastos inmortalizados de pacientes con DM1

El efecto de dos PMO-EEV dirigidos a CUG en DMPK (197-777 y 221-1106; Tabla 14) sobre el rescate del corte y empalme de genes relacionados con DM1 se evaluó in vitro mediante el uso de mioblastos y miotubos musculares derivados de pacientes con DM1.

Experimentación.

Cultivo de células. Se obtuvieron mioblastos inmortalizados de pacientes con DM1 (ASA308DM1) e individuos no afectados (KM1421; AB1190). Los mioblastos de los pacientes con DM1 albergan 2,600 repeticiones CTG en la UTR 3' de la DMPK. Los mioblastos se cultivaron en un medio de cultivo de medio de crecimiento para células musculares esqueléticas (disponible de PromoCell en Heidelberg, Alemania), suero de caballo al 2 % (disponible de Gibco en Bristol, RI), extracto de embrión de pollo al 1 % (disponible de USB Corp en, OH) y 0.5 mg/ml de penicilina/estreptomicina (Gibco). Para la diferenciación miogénica, los cultivos confluentes se cambiaron al medio de diferenciación de DMEM suplementado con suero de caballo al 2 % y se cultivaron durante cuatro

días.

Tabla 14: Compuestos sometidos a prueba en el Ejemplo 3

ID compuesto	Secuencia de EEV (N a C)	Secuencia de PMO (5'-3')	Modificaciones de PMO	Química de conjugación
5	PMO solamente	NA	5'OH	NA
10	197-777 (DM1-1) Ac-PKKKRKV-Lys(ciclo[Ff-Nal-RrRrQ])-PEG ₁₂ -K(N3)-NH ₂ Ac-(SEQ ID NO: 42)--Lys(SEQ ID NO: 78)- PEG ₁₂ -K(N3)-NH ₂	CAGCAGCAGC AGCAGCAGC AG (SEQ ID NO: 154)	5'-sarcosina amida; 3'-C4-ciclooctino	Clic
15	221-1106 (DM1-2) Ac-PKKKRKV-miniPEG2-Lys(ciclo[FfFGRGRQ])-miniPEG2-K(N3)-NH ₂ Ac-(SEQ ID NO: 42)-miniPEG2-Lys(SEQ ID NO: 76)-miniPEG2-K(N3)NH ₂	CAGCAGCAGC AGCAGCAGC AG (SEQ ID NO: 154)	5'-OH; 3'- morfolino en amina secundaria	?Clic con un enlazador PEG4-COT

Tratamiento. Las células musculares de pacientes con DM1 se trataron con 10 µm, 3 µm, 1 µm o 0.3 µm de los compuestos mediante el uso de dos condiciones de tratamiento diferentes. En la primera condición, los mioblastos se sembraron a una confluencia del 75-80 %, los compuestos se diluyeron en serie en medio de crecimiento, y las células se bañaron durante 24 horas para permitir la captación libre del compuesto. El medio que contiene el compuesto se eliminó, los mioblastos se lavaron con DPBS 1X (Gibco), y se diferenciaron durante cuatro días antes de la cosecha. Para la segunda condición de corrida en paralelo, los mioblastos se diferenciaron tres días antes del tratamiento, los compuestos se diluyeron en serie en medio de diferenciación, y los miotubos se cosecharon para su análisis 24 horas después.

Aislamiento de ARN y PCR. El ARN total se aisló con el kit RNEASY Mini (disponible de Qiagen en Germantown, MD) de acuerdo con las instrucciones del fabricante. Para la inclusión de exones, se transcribió de forma inversa 100 ng de ARN y se usó para el PCR (Kit de RT-PCR OneStep, Qiagen). Las muestras se analizaron mediante LabChip (disponible de PerkinElmer en Waltham, MD) con el kit de ensayo de alta sensibilidad de ADN HT.

Resultados. El PMO dirigido a CUG en DMPK (no conjugado con EEV) mejoró el corte y empalme

erróneo del exón 5 de MBNL1 (datos no mostrados). La Figura 11A-11F muestra un rescate mixto del defecto de corte y empalme de MBNL1 (Figura 11A) y sus objetivos (SOS1, IR, DMD, BIN1, LDB3; Figura 11B-11F) en células musculares derivadas de pacientes con DM1 tratadas con diversas concentraciones de EEV-PMO dirigidos a CUG en DMPK (CUG^{exp} 197-777 y CUG^{exp} 221-1106). EEV-PMO 197-777 provocó una corrección moderada de los eventos de corte y empalme erróneo en células musculares de pacientes con DM1. MBNL1 y SOS1 mostraron la mejor respuesta de corrección de errores de corte y empalme. Se eligió EEV-PMO 197-777 como compuesto herramienta para los experimentos de seguimiento descritos más abajo.

Los mioblastos y miotubos derivados de pacientes con DM1 se trataron con 10 µm, 3 µm y 1 µm de EEV-PMO 197-777 dirigidos a CUG en DMPK mediante el uso de métodos similares a los descritos anteriormente. Se evaluó el rescate de los eventos de corte y empalme alternativo de ARN de los objetivos MNBL1 y MNBL1. Se observaron varios grados de corrección del corte y empalme para MNBL1 (exclusión del exón 5; Figura 12A), SOS1 (inclusión del exón 25; Figura 12B), INSR (inclusión del exón 11; Figura 12C), DMD (inclusión del exón 78, Figura 12D), BIN1 (inclusión del exón 11; Figura 12E) y LDB3 (exclusión del exón 11; Figura 12F) después de que los mioblastos y los miotubos se trataron con EEV-PMO.

Las Figuras 44A-44D muestran la inversión de los fenotipos de miotonía en ratones HSA-LR tratados con 20 mpk de 221-1106 cuantificados mediante un ensayo de relajación muscular. La Figura 44A y 44C muestra gráficos del tiempo de relajación al 80 % de la fuerza isométrica máxima y la Figura 44B muestra los datos sin procesar de trazado de fuerza. La Figura 44D muestra la inversión de los fenotipos de miotonía en ratones HSA-LR tratados con 20 mpk de 221-1106 cuantificados mediante trazados representativos de electromiografía (EMG).

Ejemplo 4. Evaluación de PMO-EEV 221-1120 en un modelo de ratón con DM1

Se preparó un modelo de ratón con DM1 para estudiar el efecto de EEV-PMO 221-1120 (también denominado EEV-PMO- DM1-3 o DM1-3; PMO 221 = 5'-CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154; todos los monómeros de PMO); EEV 1120 = Ac-PKKKRKV-AEEA-Lys(ciclo[FGFGRGRQ]-PEG12-OH (Ac-(SEQ ID NO: 42)-AEEA-Lys(SEQ ID NO: 82)-PEG12-OH)) en el corte y empalme y los niveles de ARNm de los genes aguas abajo en el mismo modelo de ratón transgénico HSA-LR como se describió en el Ejemplo 2. El PMO y el EEV se conjugaron mediante el uso de química de amida.

Experimentación. Hubo dos grupos de tratamiento generales: 1) ratones de tipo salvaje; y 2) ratones HSA-LR (modelo de enfermedad DM1). Dentro del grupo de tratamiento HSA-LR hubo dos grupos

de subtratamiento: 1) HAS-LR tratado con solución salina (control); y 2) HSA-LR + EEV-PMO 221-1120. Los ratones se trataron con 15 mpk, 30 mpk, 60 mpk o 90 mpk (basado en el PMO) del PMO-EEV o solución salina mediante inyección intravenosa por la cola. Siete días después del tratamiento, se sacrificaron los ratones y se recolectaron los tejidos para su análisis.

- 5 Ensayos de RT-PCR (corrección del corte y empalme). Los tejidos se homogeneizaron mediante el homogeneizador OMNI BEAD MILL y el ARN se extrajo mediante QIACUBEQ. Los ensayos de RT-PCR se realizaron mediante el uso del kit de RT-PCT de una etapa (Qiagen) de acuerdo con los protocolos del fabricante con 35 ciclos de PCR de 94 °C durante 30 segundos; 60 °C durante 30 segundos y 72 °C durante 30 segundos. La secuencia de cebadores específicos de los genes es la siguiente: Inclusión del exón 7a de Clcn1 cebador directo = 5'-TTCACATCGCCAGCATCTGTGC-3' (SEQ ID NO: 319), cebador inverso = 5'-CACGGAACACAAAGGCACTGAATGT-3' (SEQ ID NO: 320); Inclusión del exón 5 de Mbnl1 cebador directo = 5'-GCTGCCCAATACCAGGTCAAC-3' (SEQ ID NO: 321), cebador inverso = 5'-TGGTGGGAGAAATGCTGTATGC-3' (SEQ ID NO: 322);
- 15 Inclusión del exón 22 de Atp2a1 cebador directo = 5'-GCTCATGGTCCTCAAGATCTCAC-3' (SEQ ID NO: 323), cebador inverso: 5'-GGGTCAGTGCCTCAGCTTTG-3' (SEQ ID NO: 324); Inclusión del exón 7 de Nfix cebador directo = 5'-TCGACGACAGTGAGATGGAG-3' (SEQ ID NO: 325), cebador inverso 5' CAAACTCCTTCAGCGAGTCC-3' (SEQ ID NO: 326). Los cebadores para Clcn1, Mbnl1 y Atp2a1 eran de Klein y otros, The Journal of Clinical Investigation. 2019, 129 (11), pág. 4,739; y los cebadores para Nfix eran de Chen y otros, Scientific Reports. 2016, 6(1), pág. 1. Los productos de ADNc se separaron en E-gel de agarosa al 2 % con colorante SYBR SAFE. El porcentaje de inclusión del exón se calculó mediante la relación de la banda no omitida/(banda no omitida + banda omitida).
- 25 Cálculo del índice de corte y empalme de DM1 en ratón (mDSI). El mDSI se calculó de acuerdo con el protocolo en la literatura de Tanner y otros (Nucleic acid research. 2021, 49 (4), págs. 2240-54). Para cada muestra i, se calcularon los valores de corte y empalme normalizados para cada evento de empalme j como $(PSI_{i,j} - PSI_{\text{tipo salvaje},j}) / (PSI_{\text{HSA-LR},j} - PSI_{\text{tipo salvaje},j})$, donde $PSI_{\text{tipo salvaje},j}$ es el promedio de PSI para el evento j en los ratones de tipo salvaje, y $PSI_{\text{HSA-LR},j}$ es el promedio de PSI para el evento j en los ratones HSA-LR. mDSI se calcula entonces como la media de todos los valores de corte y empalme normalizados, que son Atp2a1, Nfix, Mbnl1 y Clcn1 en los estudios.
- 30 Ensayos de qRT-PCR (atenuación del ARNm de HSA). La transcripción inversa se realizó mediante el uso del kit de transcripción inversa de ADNc de alta capacidad de Life Technologies Corporation

de acuerdo con los protocolos del fabricante. La PCR cuantitativa en tiempo real se realizó mediante el uso de la supermezcla Bio-Rad SyBr Green y la máquina de qPCR QuantStudio3 con cebadores específicos para genes: ARNm de HSA cebador directo = 5'-TTCCATCGTCCACCGCAAAT-3' (SEQ ID NO: 327), cebador inverso = 5'-AGTTTACGATGGCAGCAACG-3' (SEQ ID NO: 328),
5 ambos cebadores de Klein y otros, The Journal of Clinical Investigation. 2019, 129 (11), pág. 4739; y cebador directo de GAPDH de ratón = 5'-AGGTCGGTGTGAACGGATTTG-3' (SEQ ID NO: 329), cebador inverso = 5'-TGTAGACCATGTAGTTGAGGTCA-3' (SEQ ID NO: 330).

ARNseq. Se realizó una ARNseq de poliA mediante el uso de secuenciación de próxima generación para obtener el perfil del transcriptoma. La puntuación Z para cada gen se calculó como el (valor de la muestra – la media)/(la desviación estándar). El análisis del corte y empalme diferencial se realizó en los datos de ARNseq para calcular el por ciento de empalme (PSI) de exones individuales para cada gen. El PSI es una relación de recuentos de lecturas normalizadas que indica la inclusión de un elemento transcrito sobre las lecturas normalizadas totales para ese evento (lecturas de inclusión y exclusión). Por ejemplo, si se incluye un exón en las lecturas el 100 % del tiempo, el PSI es 1.
15 Adicionalmente, si se excluye un exón de las lecturas el 100 % del tiempo, el PSI es 0.

Se analizaron veintidós genes de interés conocidos por ser predictivos de la DM1. Adicionalmente, se analizaron los genes estudiados en Wagner y otros (PLOS Gen 2016 (47)) y Tanner y otros, (NAR 2021 (48), 4, 2,240-2,254). Los exones de ratón se mapearon a la ubicación humana. En algunos casos, el límite de los exones en ratones y/o el genoma humano no se conocía completamente. Como
20 tal, los datos se analizaron mediante el uso de diferentes límites. Los límites correctos se verificaron mediante el uso de los datos de ARNseq.

Análisis de focos ARN CUG. Las secciones de músculo tibial anterior se tiñeron para determinar los focos de CUG (FISH, rojo) y los núcleos (Hoechst, azul). Se obtuvieron imágenes de las secciones del músculo TA y se cuantificó el número de núcleos que tenían un foco de ARN con CUG.

25 Resultados

Atenuación de ARN de HSA. El diafragma solo expresó el 5-10 % de los niveles de ARNm de HSA en comparación con el cuádriceps, tibial anterior y tríceps (Figura 13A). El tratamiento con EEV-PMO no pareció cambiar los niveles de ARNm de HSA en el diafragma (Figura 13B). El nivel de expresión de las 220 repeticiones CUG de HSA puede no ser suficiente para el fenotipo de error de
30 corte y empalme en la DM1 en el diafragma.

El EEV-PMO atenuó el ARNm de HSA de una manera dependiente de la dosis, lo que confirma el acoplamiento al objetivo en tejido de cuádriceps (Figura 14A), gastrocnemio (Figura 14B), tríceps

(Figura 14C) y tibial anterior (Figura 14D). Adicionalmente, el valor de Ct (umbral de ciclo) del ARNm de HSA es similar al nivel de GAPDH de ratón (~15), lo que sugiere una alta expresión del transgén HSA en ratones HSA-LR en el cuádriceps.

mDSI (corrección del corte y empalme). El índice del corte y empalme de DM1 en ratón (mDSI) para el cuádriceps, el gastrocnemio, el tríceps, y tibial anterior se muestra en la Figura 15A-15D. El tratamiento con EEV-PMO corrigió los defectos de corte y empalme relevantes de DM1 (Atp2a1 exón 22, Nfix exón 7, Clcn1 exón 7a, Mbnl1 exón 5) 1 semana después de la inyección en el cuádriceps (Figura 15A), gastrocnemio (Figura 15B), tríceps (Figura 15C) y tibial anterior (Figura 15D) de una manera dependiente de la dosis con dosis más altas que se aproximan o son equivalentes al tipo salvaje (corrección completa). Aproximadamente el 50 %-60 % de la atenuación del ARN de la actina esquelética humana en ratones HSA-LR se logró a concentraciones del fármaco que logran una corrección del corte y empalme casi completa.

Las Figuras 16A-B muestran imágenes de tejido del tibial anterior de ratones HSA-LR (Figura 16A) y ratones HSA-LR tratados con EEV-PMO (Figura 16B) teñidos para determinar los focos de CUG (rojo) y los núcleos (azul). La evaluación cualitativa y cuantitativa (Figura 16C) mostró que el tratamiento con EEV-PMO redujo el número de núcleos que tenían focos de CUG.

Exposición al fármaco. La exposición al fármaco se estudió mediante LC-MS. La Figura 17A-17D muestra una respuesta dependiente de la dosis para la exposición a PMO-EEV en el cuádriceps (Figura 17A), tríceps (Figura 17B), corazón (Figura 17C), gastrocnemio (Figura 17D), tibial anterior (TA; Figura 17F), hígado (Figura 17I) y riñón (Figura 17J). No se observó respuesta dependiente de la dosis en el diafragma (Figura 17G). El EEV-PMO no se detectó en el cerebro excepto en los niveles de dosificación de 60 mpk y 90 mpk. La Figura 17K muestra la exposición al fármaco de diversos tejidos al nivel de dosificación de 60 mpk.

Respuesta de la miotonía: Se observó una reducción de la miotonía dependiente de la dosis en los ratones HSA-LR 7 días después del tratamiento con EEV-PMO-DM1-3 a 15, 30, 60 y 90 mpk (Figura 18A). Es probable que la miotonía mejore una semana después del tratamiento con EEV-PMO-DM1-3. Los ratones HSA-LR tratados con una dosis única de EEV-PMO-DM1-3 de 90 mpk no mostraron signos obvios de miotonía en las extremidades posteriores después de la inducción.

Análisis de los datos de ARNseq. La Figura 19A-19D muestra los resultados de un análisis de componentes principales. El análisis de componentes principales puede usarse para revelar la similitud entre muestras basado en la matriz de distancia. Este tipo de gráfico es útil para visualizar el efecto general de las covariadas experimentales y los efectos de lote. El eje x es la dirección que

explica la mayor varianza y el eje y es la segunda mayor. El porcentaje de la varianza total por dirección se muestra como PCA. Los ratones de tipo salvaje y HSA-LR están en grupos distintos. La expresión génica en el músculo gastrocnemio de ratones HSA-LR tratados con PMO-EEV se desplazó hacia la de los ratones de tipo salvaje.

5 La Figura 20A es un mapa térmico que muestra genes expresados diferencialmente (mediante la puntuación Z) de tres grupos de tratamiento: 1) ratones WT; 2) ratones HSA-LR; y 3) HSA-LR + EEV-PMO 221-1120 (60 mpk). Un total de 956 ($p < 0.5$) genes se expresaron diferencialmente entre los grupos de tratamiento, lo que indica una diferencia entre los ratones de tipo salvaje (WT) y los ratones modelo de la enfermedad (HSA-LR). El tratamiento con EEV-PMO (HSA-LR (+,+)) dio
10 como resultado la corrección global de la expresión génica, que se alejó de un perfil de enfermedad (en rojo, HSA-LR (-,-)) y hacia el de los ratones de tipo salvaje (arriba WT).

La Figura 20B es un mapa térmico usado para visualizar el perfil de expresión de 40 de los 43 genes que se encontró que tienen más de 7 repeticiones CTG a partir de un análisis BLAST. Tres de los genes con repeticiones CTG que se encontraron en el análisis Blast (Crb2, Hsd3b6 e Inhbe) no se
15 incluyeron debido al bajo número de lecturas. Este análisis es útil para identificar genes regulados conjuntamente en las condiciones de tratamiento.

La Figura 21 muestra un gráfico de volcán del cambio transcripcional global en el grupo tratado con EEV-PMO y el grupo HSA-LA. Cada punto de datos en el gráfico de dispersión representa un gen. El cambio en veces de cada gen se representa en el eje x y el logaritmo 10 de su valor de p ajustado
20 está en el eje y. Los genes con un valor de p ajustado menor que 0.05 y un cambio en veces mayor que 2 se indican mediante puntos rojos. Estos representan genes regulados positivamente. Los genes con un valor de p ajustado menor que 0.05 y un cambio en veces menor que -2 se indican mediante puntos azules. Estos representan genes regulados negativamente. Se encontró que tres genes estaban regulados negativamente de forma significativa (Txlnb, Scube2 y Greb1) y que un gen estaba
25 regulado positivamente de forma significativa (Txlnb). La mayoría de los transcritos que contenían al menos (CUG)₇ no se vieron influenciados significativamente. El análisis de PCA de estos genes mostró que Scube2 (Figura 22A), Greb1 (Figura 22B), Ttc7 (Figura 22C), Txlnb (CUG)₉ y Ndr3 (Figura 22E) mostraron corrección cuando se trataron con EEV-PMO. Txlnb se sobre corrigió mediante el tratamiento (Figura 22D).

30 Las Figuras 23A-23D muestran los datos de transcriptoma para diversos genes y diversos grupos de tratamiento. El grupo de tratamiento HSA-LR+ EEV-PMO 221-1120 mostró la corrección de la inclusión del exón 22 de Atp2a1 (Figura 23A), exclusión del exón 7 de Clcn1 (Figura 23B),

exclusión del exón 7 de Nfix (Figura 23C) y exclusión del exón 7 de Mbnl1 (Figura 23D).

La Figura 24 muestra el por ciento de empalme (PSI) de exones individuales para diversos genes. Los genes son biomarcadores de corte y empalme sensibles a MBNL-1 (por ejemplo, genes aguas abajo). La elección de biomarcadores dependientes de MBNL-1 se seleccionó basado en el intervalo dinámico entre los grupos de tipo salvaje y de enfermedad, como se describe en la literatura. El grupo de tratamiento HSA-LR+EEV-PMO 221-1120 mostró corrección para la inclusión/exclusión de exones para todos los 20 genes de interés, lo que incluye Mbnl1, Nfix, Atp2a1, Ldb3, Camk2g, Trim55, Fbox31, Slc8a3, Map3k4, Dctn4, Cacna1s, Ryr1, Slain2, Phka1, Ppp3cc, Ttn, Neb, Irfp2, Rapgef1 y Vsp39.

Ejemplo 5. Evaluación de PMO-EEV 221-1120 en un segundo modelo de ratón con DM1

PMO-EEV DM1-3 (221-1120; ver el Ejemplo 4 para la secuencia) se evaluó en un segundo modelo de ratón con DM1 mediante el uso de un método similar al descrito en el Ejemplo 4.

Experimentación. A ratones HSA-LR de siete semanas de edad se les administró una dosis única de 80 mpk de EEV-PMO-DM1-3 o una dosis de 20 mpk de EEV-PMO-DM1-3 cada dos semanas durante seis semanas (total de 80 mpk en 4 dosis) por vía intravenosa y se cosecharon los tejidos después de 1 semana a 12 semanas después de la dosis única o dos semanas después de la dosis final. Se usó RT-PCR para determinar el corte y empalme alternativo para genes específicos (Atp2a1, Clcn1, Nfix, MBNL1). Se usó Q-PCR para determinar la reducción del nivel de ARNm de actina-HSA después del tratamiento. Se usó LC-masas para determinar el nivel del fármaco en el cuádriceps, gastrocnemio, tibial anterior, tríceps, diafragma, corazón, riñón, hígado, cerebro y plasma. Se registró la reducción de la miotonía 7 días después del tratamiento con el compuesto EEV-PMO-DM1-3.

Resultados. Se observaron tendencias similares a las observadas en el Ejemplo 4 para el rescate del corte y empalme para Atp2a1, Clcn1, Nfix y MBNL1 en los tejidos de tibial anterior, gastrocnemio, tríceps y cuádriceps (datos no mostrados). Adicionalmente, se observaron tendencias similares a las observadas en el Ejemplo 4 para la atenuación del ARNm de HSA en los tejidos tibial anterior, gastrocnemio, tríceps y cuádriceps tanto 1 semana como 4 semanas después del tratamiento (datos no mostrados).

La Figura 25A-25D son gráficos que muestran una disminución en el nivel del fármaco con EEV-PMO-DM1-3 de 80 mpk después de 1 semana a 8 semanas en los tejidos de tibial anterior (Figura 25A), gastrocnemio (Figura 25B), tríceps (Figura 25C) y cuádriceps (Figura 25D). EEV-PMO-DM1-3 (oligómero 60 mpk, fármaco completo 80 mpk) corrige completamente el error de corte y empalme

en gastrocnemio, tríceps, tibial anterior y cuádriceps después de 1 semana de tratamiento. La Figura 26A-26B son gráficos que muestran que se observó una disminución en los niveles del fármaco con la única dosis de 80 mpk de EEV-PMO-DM1-3 después de 1 semana a 4 semanas, a 8 semanas y a 12 semanas en el hígado. Se observó una cantidad relativamente mayor de EEV-PMO-DM1-3 en el hígado 2 semanas después de la última dosis del régimen de dosificación de 6 semanas, en comparación con 4 semanas después del régimen de dosis única.

La Figura 26C-26D muestra una disminución en los niveles del fármaco con la dosis única de 80 mpk de EEV-PMO-DM1-3 después de 1 semana a 4 semanas, a 8 semanas y a 12 semanas en el riñón. También se observó una cantidad relativamente baja de EEV-PMO-DM1-3 en el riñón desde 2 semanas después de la última dosis del régimen de dosificación de 6 semanas, en comparación con 4 semanas después del régimen de dosis única. A las 12 semanas después de la dosis única de EEV-PMO-DM1-3, el fármaco todavía estaba presente en el riñón, pero no en el hígado.

Se realizaron observaciones subjetivas de la miotonía y se muestran en la Tabla 15. El régimen de dosis múltiple (Q2W) no mostró signos de rescate de la miotonía después de dos semanas después del tratamiento. Hubo un efecto mixto sobre la miotonía en ratones tratados con la única dosis de 80 mpk, que desapareció 12 semanas después del tratamiento.

Tabla 15. Observaciones de la miotonía

Grupo	Antes de la administración		1 semana después de la dosis		4 semanas después de la dosis		8 semanas después de la dosis		12 semanas después de la dosis	
	F	M	F	M	F	M	F	M	F	M
FVB	0	0	0	0	0	0	0	0	0	0
HSA-LR	++	++	++	++	++	++	++	++	++	++
221-112 80 mpk	++	++	+	+	+	+	0	+	++	++
221-1120 20 mpk Q2W	++	++	NA	NA	NA	NA	++	++	NA	NA

F = hembra; M = macho; 0 = ningún ratón mostró miotonía; + = miotonía mixta, algunos ratones mostraron reducción de la miotonía; ++ = todos los ratones mostraron miotonía

Se realizó un experimento similar para evaluar la administración intravenosa de EEV-PMO-DM1-3 por una mayor duración y a una dosis más alta. Se trataron ratones HSA-LR de ocho semanas de edad con 40 mpk, 60 mpk, 80 mpk o 120 mpk de EEV-PMO-DM1-3 por vía intravenosa y se cosecharon los tejidos después de 4 semanas a 12 semanas. Se usó RT-PCR para determinar el corte y empalme alternativo para genes específicos (Atp2a1, Clcn1, Nfix, MBNL1). Se registró la

reducción de la miotonía 7 días después del tratamiento con EEV-PMO-DM1-3. Se observaron tendencias similares a las observadas en el Ejemplo 4 para el rescate del corte y empalme de Atp2a1, Clcn1, Nfix y MBNL1 en los tejidos tibial anterior y gastrocnemio (datos no mostrados).

Se realizaron observaciones subjetivas de la miotonía y se muestran en la Tabla 16. Las hembras mostraron más miotonía que los machos. No hay signos de miotonía tanto en ratones machos como hembras a los que se les administró 120 mpk después de 8 semanas después del tratamiento.

Tabla 16. Observaciones de la miotonía

Grupo	Antes de la administración		1 semana después de la dosis		4 semanas después de la dosis		8 semanas después de la dosis		12 semanas después de la dosis	
	F	M	F	M	F	M	F	M	F	M
FVB	0	0	0	0	0	0	0	0	0	0
HSA-LR	++	++	++	++	++	++	++	++	++	++
221-1120 40 mpk	++	++	+	++	++	++	NA	NA	NA	NA
221-1120 60 mpk	++	++	+	0	+	+	NA	NA	NA	NA
221-1120 80 mpk	++	++	0	0	+	+	+	0	+	0
221-1120 120 mpk	++	++	0	0	0	0	0	0	0	0
F = hembra; M = macho; 0 = ningún ratón mostró miotonía; + = miotonía mixta, algunos ratones mostraron reducción de la miotonía; ++ = todos los ratones mostraron miotonía										

Ejemplo 6. Tratamiento de células DM1 derivadas de pacientes con EEV-PMO-DM1-3

PMO-EEV DM1-3 (221-1120; ver el Ejemplo 4 para la secuencia) se evaluó en mioblastos derivados de pacientes con DM1.

Experimentación. Los mioblastos de pacientes se trataron con 30 micromolar de DM1-3 durante cuatro días de diferenciación. La corrección del corte y empalme se evaluó mediante RT-PCR de una etapa y Labpchip (Media gráfica $\pm$ SD; n=4). Se usaron ensayos de detección de proteínas MBNL1 secuestradas y HCR-FISH para detectar los focos de ARN CUG. Resultados: EEV-PMO-DM1-3 promueve una corrección significativa del corte y empalme de biomarcadores y una reducción de los focos nucleares en células musculares derivada de pacientes con DM1.

Resultados. La Figura 27A-27C son gráficos que muestran que EEV-PMO-DM1-3 promueve una corrección significativa del corte y empalme de biomarcadores (MBLN1, SOS1 y NFIX) en células musculares derivadas de pacientes con DM1. Adicionalmente, el tratamiento con DM1-3 dio como resultado la reducción de los focos nucleares en células musculares derivadas de pacientes con DM1 (Figura 28A-28C).

Ejemplo 7. Cribado de citotoxicidad de EEV-PMO-DM1-3 en células renales

PMO-EEV DM1-3 (221-1120; ver el Ejemplo 4 para la secuencia) se evaluó en células renales humanas.

Experimentación. Se expusieron células epiteliales de túbulo proximal primarias humanas (RPTEC) a concentraciones variables (dilución en serie 1:2 en solución salina con un factor de dilución final de 4x de aproximadamente 6 μ M a aproximadamente 800 μ M) de PMO-DM1 y EEV-PMO-DM1-3 durante 24 horas y se cribaron para determinar su viabilidad mediante el uso de un ensayo de viabilidad luminiscente CELLTITER-GLO. Se usó la melitina se usó como control positivo a 16.6 μ M.

Resultados. La Figura 29A-29B muestra que PMO-DM1 o su conjugado EEV-PMO-DM1-3 no mostró ninguna toxicidad incluso con la concentración más alta de 817 μ M o 797 μ M, respectivamente.

Ejemplo 8. Evaluación de PMO-EEV 221-1113 para determinar la capacidad de corregir eventos de errores de corte y empalme y corte y empalme aguas abajo en células HeLa-480 y células inmortalizadas de pacientes con DM1

Células musculares inmortalizadas derivadas de pacientes con DM1 (2,600 repeticiones CUG) y HeLa-480 (línea celular modelo DM1, ver el Ejemplo 1) se trataron con el constructo 221-1113 de EEV-PMO y se analizaron para determinar la corrección del corte y empalme aberrante y la cuantificación de los focos. EEV 1113 es Ac-PKKKRKV-miniPEG-K(*ciclo*(Ff-Nal-GrGrQ)-PEG12-OH (Ac-(SEQ ID NO: 42)-miniPEG-K(*ciclo*(SEQ ID NO: 80)-PEG12-OH). EEV-PMO 221-1113 es el EEV 1113 conjugado a la secuencia 221 de PMO (5'-CAG CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154; todos los monómeros de PMO) mediante química de enlace de amida.

Experimentación. Se usaron métodos similares a los descritos en el Ejemplo 1 y el Ejemplo 3.

Resultados.

Análisis de focos ARN CUG. Las células se tiñeron para detectar los núcleos (Hoeschet, azul) y los focos de ARN con repeticiones CUG (verde) y se obtuvieron imágenes. Se observó una reducción en los focos de ARN CUG entre las células de pacientes con DM1 no tratadas y las células DM1 tratadas con EEV-PMO (Figura 30A-30C). De manera similar, se observó una reducción en los focos de ARN CUG entre las células HeLa-480 no tratadas y las células HeLa-480 con EEV-PMO (Figura 31A-31B).

Corrección del corte y empalme aguas abajo. Las células derivadas de pacientes con DM1 tratadas

con PMO-EEV y las células HeLa-480 se analizaron para determinar el por ciento de inclusión del exón 5 para MBLN1, el por ciento de inclusión del exón 25 para SOS1 y el por ciento de inclusión del exón 7 para NFIX. El tratamiento con EEV-PMO dio como resultado un rescate de los eventos de corte y empalme para Mbnl1 (Figura 32A), Sos1 (Figura 32B) y NFIX (Figura 32C).

- 5 Adicionalmente, las células HeLa-480 tratadas con EEV-PMO mostraron una corrección dependiente de la dosis del corte y empalme de MBNL1 (Figura 33A) y de los errores de corte y empalme aguas abajo de SOS1 (Figura 33B), CLASP1 (Figura 33C), NFIX (Figura 33D) e INSR (Figura 33E) de una manera dependiente de la dosis.

Ejemplo 9. Evaluación de EEV-PMO 221-1106 en un segundo modelo de ratón con DM1

- 10 Se realizó un modelo de ratón con DM1 para estudiar el efecto de EEV-PMO 221-1106 sobre el corte y empalme y los niveles de ARNm de genes aguas abajo.

Experimentación. Se usaron ratones transgénicos de repetición larga de actina esquelética humana (HSA-LR) como el modelo de enfermedad DM1. Se usaron métodos similares a los descritos en el Ejemplo 5.

- 15 Resultados. Las Figuras 34A-34D muestran una corrección dependiente de la dosis de la inclusión del exón 22 en Atp2a1 (Figura 34A), exón 7 en Nfix (Figura 34B), exón 7A en Clcn1 (Figura 34C) y Mbnl1 (Figura 34D) en el gastrocnemio de ratones tratados con diversas concentraciones de EEV-PMO 221-1106. El tratamiento con PMO 221 solamente, no dio como resultado la corrección del corte y empalme.

- 20 Ejemplo 10. Modelo de ratón con DM1 para estudiar el efecto de diferentes longitudes de repeticiones CUG en los PMO

Se preparó un modelo de ratón con DM1 para estudiar el efecto de PMO-EEV 221-1121 (PMO tiene 7 repeticiones CAG, 21-mer) y PMO-EEV 0325-1121 (PMO tiene 8 repeticiones CAG, 24-mer) sobre el corte y empalme y los niveles de ARNm de genes aguas abajo. PMO-EEV 221-1121 es PMO 221 (5'-CAG CAG CAG CAG CAG CAG CAG-3'; SEQ ID NO: 154; todos los monómeros de PMO) conjugado a EEV 1121 (Ac-PKKKRKV-miniPEG2-Lys(ciclo[GfFGrGrQ])-PEG12-OH; Ac-(SEQ ID NO: 42)-miniPEG2-Lys(SEQ ID NO:74)-PEG12-OH) mediante química de amida. PMO-EEV 0325-1121 es PMO 0325 (5'-CAG CAG CAG CAG CAG CAG CAG CAG-3'; SEQ ID NO: 155; todos los monómeros de PMO) conjugado a EEV 1121 mediante química de amida.

- 30 Experimentación. Se usaron ratones transgénicos de repetición larga de actina esquelética humana (HSA-LR) como el modelo de enfermedad DM1. Brevemente, a los ratones HSA-LR se les administró 20 mpk, 40 mpk o 60 mpk de 0221-1121 o 0325-1121 mediante inyección intravenosa

en la vena de la cola. Una semana después de la inyección, se sacrificaron los ratones y se recolectó el tejido. Se usaron otros métodos experimentales que son similares a los descritos en el Ejemplo 5. Resultados. 0221-1121 (21-mer) fue más efectivo en la corrección del corte y empalme de exones en Mbnl1 (Figura 35A), Nfix (Figura 35B) y Atp2a1 (Figura 35C) que 0325-1121 (24-mer) en el tejido del tibial anterior. Este resultado fue inesperado. Se esperaba que el de 24 mer sería más efectivo ya que tendría una mayor eficiencia de hibridación y una temperatura de fusión térmica más alta. En el tejido gastrocnemio, las diferencias fueron menos pronunciadas como se muestra en las Figuras 36A-36C. Se realizaron observaciones subjetivas de la miotonía mediante el uso de los ratones machos (Tabla 17). Se observó miotonía mixta 1 semana después del tratamiento para el 21 mer a 40 mpk, similar a los resultados en la Tabla 11.

Tabla 17: Observaciones de la miotonía

Grupo	Antes de la administración	1 semana después de la dosis	4 semanas después de la dosis
Sexo	M	M	M
FVB	0	0	NA
HSA-LR	++	++	NA
221-112 20 mpk	++	++	NA
221-1120 40 mpk	++	+	
325-1120 20 mpk	++	++	NA
325-1120 40 mpk	++	++	

M= macho; 0 = ningún ratón mostró miotonía; + = miotonía mixta, algunos ratones mostraron reducción de la miotonía; ++ = todos los ratones mostraron miotonía

Ejemplo 11. Estudios farmacocinéticos del EEV-PMO 221-1120 en ratones CD1

Se usó un modelo de ratón CD1 para estudiar la exposición al fármaco (AUC) en plasma, riñón y tibial anterior al constructo EEV-PMO 221-1120 (ver el Ejemplo 4 para la secuencia) y PMO-0221a, también se midió el metabolito principal de 221-1120 (ver la Figura 37).

Experimentación. Ratones CD1 de cinco a siete semanas de edad se trataron con 80 mpk del constructo EEV-PMO 221-1120 mediante inyección intravenosa. Los ratones se desangraron y/o se sacrificaron en diversos puntos temporales.

Resultados. La Tabla 18, la Tabla 19 y la Tabla 20 muestran las propiedades farmacocinéticas observadas en plasma, riñón y tibial anterior, respectivamente. Para las tablas: AUC_{últ} = área bajo la curva desde cero hasta la última concentración cuantificable; D = dosis; C_{máx} = concentración sérica o plasmática máxima; T_{máx} = tiempo hasta alcanzar C_{máx}; CL = aclaramiento total en plasma, suero o sangre; t_{1/2} = semivida de eliminación; V_{ss} = volumen aparente de distribución en equilibrio; Q_h =

flujo sanguíneo hepático (ml/min/kg).

Los valores de AUC para el metabolito son ~1,000 veces menores en el tibial anterior en comparación con el riñón. Los valores del tiempo medio de residencia (MRT) del metabolito en plasma pueden estar directamente relacionados con los valores de MRT en tejido como resultado del movimiento de los tejidos al plasma antes de la excreción urinaria.

Tabla 18: Propiedades farmacocinéticas en plasma

	221-1120	PMO-0221a
AUC _{últ} (nM*h)	9,290	953
AUC _{últ} /D	1,121	115
C _{máx} (nM)	16,217	12
C _{máx} /D	1,956	1.4
T _{máx} (h)	0.1	24
CL (ml/min/kg)	15	-
Q _h (%)	16	-
t _{1/2} (h)	19	68
MRT _{últ} (h)	1.2	60.0
V _{ss} (ml/kg)	1,325	-

Tabla 19: Propiedades farmacocinéticas en riñón

	221-1120	PMO-0221a
AUC _{últ} (pmol/g*h)	51,609	10,113,709
AUC _{últ} /D	6,225	1,219,865
C _{máx} (pmol/g)	8,754	105,053
C _{máx} /D	1,056	12,671
T _{máx} (h)	4	24
t _{1/2} (h)	5	87
MRT _{últ} (h)	8	68

Tabla 20: Propiedades farmacocinéticas en tibial anterior

	221-1120	PMO-0221a
AUC _{últ} (pmol/g*h)	-	9,140
AUC _{últ} /D	-	1,102
C _{máx} (pmol/g)	34	162
C _{máx} /D	4	20
T _{máx} (h)	4	24
t _{1/2} (h)	-	48
MRT _{últ} (h)	-	50

Ejemplo 12. Evaluación de PMO-EEV 221-1120 en un tercer modelo de ratón con DM1

Se realizó un estudio en modelo de ratón con DM1 similar a los Ejemplos 5 y 9 para evaluar el efecto de diversas dosis de PMO-EEV 221-220 (ver el Ejemplo 4 para la secuencia) en ratones HSA-LR.

Experimentación. Se administraron ratones HSA-LR de ocho semanas de edad con 40, 60, 80 o 120
5 mpk de PMO-EEV 221-1120 por vía intravenosa y se cosecharon los tejidos después de 4 a 12 semanas. Se usó RT-PCR para determinar el corte y empalme alternativo para genes específicos (Atp2a1, Clcn1, Nfix, MBNL1). Se usó LC-masas para determinar el nivel del fármaco en cuádriceps, gastro, TA, tríceps, diafragma, corazón, riñón, hígado, cerebro, plasma. Se usó ARNseq para determinar el cambio del nivel de transcripción entre un modelo de enfermedad tratado, un
10 modelo de enfermedad no tratado y el tipo salvaje. Se usó Q-PCR para determinar la reducción del nivel de ARNm de actina-HSA después del tratamiento.

Resultados. Se usó obtención de imágenes por fluorescencia para determinar la reducción de los focos de ARN después del tratamiento con el compuesto EEV-oligo (datos no mostrados). Se registró una reducción de la miotonía 7 días después del tratamiento con el compuesto EEV-oligo (datos no
15 mostrados). Los resultados de estos experimentos muestran tendencias similares a las de los ratones tratados con PMO-EEV 221-1120 en el Ejemplo 5 y el Ejemplo 14.

Se observó la corrección del corte y empalme de MBNL1, NFIX y ATP2A1 en tibial anterior (Figuras 38A, 39A, 40A) y gastrocnemio (Figuras 38B, 39B, 40B) a diversas dosis del EEV-PMO. Se observó corrección del corte y empalme de MBNL1, NFIX y ATP2A1 tanto en tibial anterior
20 como en el gastrocnemio 12 semanas después del tratamiento con EEV-PMO a 120 mpk.

Ejemplo 13. Evaluación de EEV-PMO 221-1120 en células HeLa480

Células HeLa480 se trataron con diversas concentraciones de EEV-PMO 221-1120 (ver el Ejemplo 4 para la secuencia) y se analizaron para detectar los focos de repeticiones CUG, reducción selectiva de r(CUG) y corrección del corte y empalme aguas abajo de MBNL1 y SOS1.

Experimentación. Las células Hela480 se construyeron como se describió en los Ejemplos anteriores. La RT-PCR y la tinción de los focos se realizó de manera similar a otros Ejemplos descritos en la presente descripción.

Resultados. La Figura 41A-41B muestra imágenes de ejemplo de células de control (no tratadas) y células tratadas con 5 μ M, 10 μ M, 20 μ M, 50 μ M y 100 μ M de EEV-PMO 221-1120. Hay una
30 reducción en los focos de CUG (verde) en el grupo de HeLa480 tratadas en comparación con las células HeLa480 no tratadas. La Figura 41B es un gráfico que cuantifica los focos por área nuclear. La Figura 41A-41B muestra que EEV-PMO 221-1120 puede reducir los focos nucleares de ARN

CUG. Se observa una reducción casi completa en la dosis de 5 μ M.

La Figura 42A-42B indica que el tratamiento con EEV-PMO 221-1120 puede atenuar selectivamente el transcrito de DMPK que contiene expansiones repetidas en la línea celular HeLa480.

La Figura 42C-42D muestra que el tratamiento con EEV-PMO 221-1120 corrigió el corte y empalme de MBNL1 (Figura 42C) y SOCS1 (Figura 42D) de una manera dependiente de la dosis.

Se han descrito varias modalidades de la invención. Sin embargo, se debe entender que pueden realizarse diversas modificaciones sin apartarse del espíritu y alcance de la invención. En consecuencia, otras modalidades están dentro del alcance de las siguientes reivindicaciones.

10

15

20

25

30