

(19)



Евразийское  
патентное  
ведомство

(21) 202393234 (13) A1

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки  
2024.07.26(51) Int. Cl. A61K 47/64 (2017.01)  
C07K 7/64 (2006.01)  
C12N 15/113 (2010.01)(22) Дата подачи заявки  
2022.06.22

## (54) АНТИСМЫСЛОВЫЕ СОЕДИНЕНИЯ И СПОСОБЫ ДЛЯ НАЦЕЛИВАНИЯ НА ПОВТОРЫ CUG

(31) 63/213,900; 63/239,671; 63/239,847; 63/290,892;  
63/290,960; 63/298,565; 63/305,071; 63/268,577;  
63/314,369; 63/316,634; 63/317,856; 63/326,201;  
63/362,295; 63/327,179; 63/339,250(71) Заявитель:  
ЭНТРЕДА ТЕРАПЬЮТИКС, ИНК.  
(US)(32) 2021.06.23; 2021.09.01; 2021.09.01; 2021.12.17;  
2021.12.17; 2022.01.11; 2022.01.31; 2022.02.25;  
2022.02.26; 2022.03.04; 2022.03.08; 2022.03.31;  
2022.03.31; 2022.04.04; 2022.05.06(72) Изобретатель:  
Шэнь Сюлун, Цянь Цзыцин, Доэрти  
Патрик, Кейрабади Махбоубех, Ли  
Сян (US)

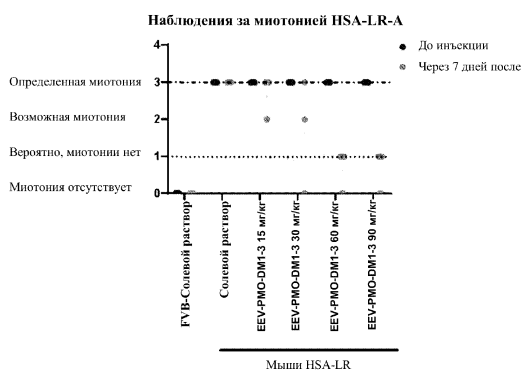
(33) US

(74) Представитель:  
Медведев В.Н. (RU)

(86) PCT/US2022/034517

(87) WO 2022/271818 2022.12.29

(57) Предложены соединения, содержащие циклический пептид, такой как проникающий в клетку циклический пептид, и антисмысловое соединение. Антисмысловое соединение связывается с геном, содержащим экспансированный повтор CTG, или геномным транскриптом, содержащим экспансированный повтор CUG. Соединения могут быть доставлены субъектам для лечения заболеваний, ассоциированных с экспансированными повторами CTG-CUG, таких как миотоническая дистрофия 1 типа (DM1), спиноцеребеллярная атаксия-8 (SCA8) и болезнь, подобная болезни Хантингтона, 2 типа (HDL2).



A1

202393234

202393234

A1

## ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-579948EA/042

### АНТИСМЫСЛОВЫЕ СОЕДИНЕНИЯ И СПОСОБЫ ДЛЯ НАЦЕЛИВАНИЯ НА ПОВТОРЫ CUG

#### ПЕРЕКРЕСТНАЯ ССЫЛКА НА РОДСТВЕННЫЕ ЗАЯВКИ

Настоящая заявка испрашивает приоритет на основании предварительной заявки на патент США с серийным номером: 63/213900, поданной 23 июня 2021 г.; 63/239847, поданной 1 сентября 2021 г.; 63/290892, поданной 17 декабря 2021 г.; 63/305071, поданной 31 января 2022 г.; 63/314369, поданной 26 февраля 2022 г.; 63/316634, поданной 4 марта 2022 г.; 63/317856, поданной 8 марта 2022 г.; 63/326201, поданной 31 марта 2022 г.; 63/327179, поданной 4 апреля 2022 г.; 63/339250, поданной 6 мая 2022 г.; 63/362295, поданной 31 марта 2022 г.; 63/239671, поданной 1 сентября 2021 г.; 63/290960, поданной 17 декабря 2021 г.; 63/298565, поданной 11 января 2022 г.; и 63/268577, поданной 25 февраля 2022 г.

#### ОБЛАСТЬ ТЕХНИКИ

Настоящее изобретение относится к соединениям, композициям и способам для модуляции активности и/или уровней генов, которые содержат экспансированные нуклеотидные повторы, в частности, экспансированные повторы CTG·CUG. Соединения и композиции, содержащие указанные соединения, можно применять для лечения заболеваний, ассоциированных с генами, которые содержат экспансированный нуклеотидный повтор, в частности, экспансированные повторы CTG·CUG.

#### ВВЕДЕНИЕ

Несколько заболеваний ассоциированы с генами, которые содержат экспансированные нуклеотидные повторы, то есть большее число нуклеотидных повторов, чем наблюдается при здоровом фенотипе. Экспансированный повтор может вызывать агрегацию и/или нуклеацию содержащего экспансированный повтор транскрипта и/или вызывать нуклеацию белков, которые связываются с содержащим экспансированный повтор транскриптом. Экспансированные повторы могут приводить к тому, что некоторые белки, такие как белки, осуществляющие процессинг пре-мРНК, секвестрируются на повторе, что ингибирует выполнение белками их нормальных функций, таких как процессинг транскриптов пре-мРНК других генов, которые не содержат экспансированного повтора.

Существует несколько заболеваний, ассоциированных с генами, содержащими экспансированные тринуклеотидные повторы CTG·CUG (CTG относится к повтору ДНК, и CUG относится к соответствующему повтору РНК, который образуется при транскрипции). Заболевания, ассоциированные с генами, содержащими экспансированные тринуклеотидные повторы CTG·CUG, включают, но не ограничиваются перечисленными, миотоническую дистрофию 1 типа (DM1), спиноцеребеллярную атаксию-8 (SCA8), болезнь, подобную болезни Хантингтона, 2 типа (HDL2) и эндотелиальную дистрофию роговицы Фукса (FECD).

Миотоническая дистрофия 1 типа (DM1), наиболее распространенная причина мышечной дистрофии у взрослых, поражающая 1 из 8500 индивидуумов во всем мире, ассоциирована с геном, который содержит экспансированный тринуклеотидный повтор (Lee and Cooper. (2009) «Pathogenic mechanisms of myotonic dystrophy», *Biochem Soc Trans.* 37(06): 10.1042/BST0371281). DM1 представляет собой нарушение, которое поражает скелетные и гладкие мышцы, а также глаз, сердце, эндокринную систему и центральную нервную систему. DM1 вызвана патологической экспансией CTG-тринуклеотидного повтора в некодирующей области гена, кодирующего протеинкиназу мышечной дистрофии (DMPK). Экспансия CTG лежит в пределах области, соответствующей 3'-нетранслируемой области (3'-НТО) мРНК DMPK. В то время как ген *DMPK* у здоровых индивидуумов содержит от 5 до 40 тринуклеотидных повторов CTG, у пациентов, страдающих DM1, имеется от 50 до нескольких тысяч тринуклеотидных повторов CTG. Экспансия CTG-тринуклеотидных повторов приводит к общему нарушению регуляции экспрессии генов у пораженных индивидуумов вследствие нуклеации некоторых регуляторных РНК-связывающих белков в CUG-экспансии в 3'-нетранслируемой области (3'-НТО), что делает РНК-связывающие белки, такие как белок, подобный белкам семейства *muscleblind* (MBNL1-3), неспособными выполнять свою нормальную клеточную функцию. Нуклеированные РНК-связывающие белки недоступны для связывания и влияния на трансляцию других транскриптов мРНК. Агрегаты мРНК с экспансированными CUG и белка образуют отдельные ядерные очаги. Активность дополнительных факторов сплайсинга, таких как представитель семейства подобного CUGBP *Elav 1* (CELF1), также нарушается, что приводит к ошибочному сплайсингу большого числа транскриптов нисходящих генов, ассоциированных с симптомами DM1. С увеличением числа повторов увеличивается тяжесть заболевания и уменьшается возраст начала (Pettersson et al. (2015) «Molecular mechanisms in DM1 - a focus on foci». *Nucleic Acids Res.* 43(4):2433-2441).

CUG-тринуклеотидные повторы в 3'-нетранслируемой области мРНК DMPK образуют дефектные стабильные шпилечные структуры, которые накапливаются в ядре клетки в небольших рибонуклеарных комплексах или микроскопически видимых включениях и нарушают функцию белков, вовлеченных в транскрипцию, сплайсинг или экспорт РНК. Несмотря на то, что гены *DMPK* с повторами CUG транскрибируются в мРНК, мутантные транскрипты секвестрируются в ядре в виде агрегатов (очагов), что приводит к уменьшению цитоплазматических уровней мРНК DMPK. Эти агрегаты приводят к нарушению регуляции альтернативного сплайсинга многих различных транскриптов вследствие секвестрации двух РНК-связывающих белков: MBNL1 (подобный *muscleblind 1*) и CUGBP1 (CUG-связывающий белок 1), что приводит к утрате функции MBNL1 и положительной регуляции CUGBP1 (Lee and Cooper. (2009) «Pathogenic mechanisms of myotonic dystrophy», *Biochem Soc Trans.* 37(06): 10.1042/BST0371281). MBNL1 и CUGBP1-подобный фактор 1 (CELF1) являются связанными с развитием регуляторами событий сплайсинга во время перехода от

эмбриона к взрослому, и модификация их активности при DM1 приводит к экспрессии эмбрионального профиля сплайсинга в тканях взрослых. Последующее влияние уменьшенных уровней MBNL1 и повышенных уровней CELF1 включает нарушение альтернативного сплайсинга, трансляции мРНК и распада мРНК белков, например, транскриптов сердечного тропонина Т (сТNT), инсулинового рецептора (INSR), специфического для мышц канала для ионов хлора (CLCN1) и кальциевой АТФазы 1 саркоплазматического/эндоплазматического ретикулума (ATP2A1), помимо MBNL1 (Konieczny et al. (2017) «Myotonic dystrophy: candidate small molecule therapeutics.» Drug Discov Today. 22(11):1740-174).

Возможные терапевтические подходы к лечению DM1 или других заболеваний, ассоциированных с экспансированными повторами CTG·CUG, включают применение содержащих терапевтический олигонуклеотид соединений. Однако основной проблемой, ассоциированной с применением олигонуклеотидных соединений в терапевтических средствах, является их ограниченная способность получать доступ к внутриклеточному компартменту при системном введении. Внутриклеточная доставка олигонуклеотидных соединений может быть облегчена путем применения систем носителей, таких как полимеры, катионные липосомы, или путем химической модификации конструкции, например, путем ковалентного присоединения молекул холестерина. Однако эффективность внутриклеточной доставки олигонуклеотидных соединений остается низкой. Для повышения эффективности этих соединений все еще требуются улучшенные системы доставки.

Существует неудовлетворенная потребность в эффективных композициях для доставки терапевтических олигонуклеотидных соединений во внутриклеточные компартменты для лечения заболеваний, вызванных экспансированными повторами CTG·CUG, таких как DM1.

#### КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

В настоящем документе описаны соединения, композиции и способы для лечения заболевания, ассоциированного с экспансированным повтором CTG CUG. В вариантах реализации настоящее изобретение относится к соединениям, которые включают антисмысловое соединение (АС) и циклический пептид, такой как проникающий в клетку циклический пептид (сСРР). Согласно вариантам реализации АС связывается с геном или транскриптом гена, содержащим экспансированный повтор CUG. Согласно вариантам реализации циклический пептид облегчает внутриклеточную локализацию АС. Соединения могут содержать носитель для ускользания из эндосом (EEV). EEV может содержать циклический пептид и экзоциклический пептид.

Согласно вариантам реализации настоящего изобретения предложено соединение, содержащее: (а) по меньшей мере один циклический пептид и (b) антисмысловое соединение (АС), которое комплементарно целевому нуклеотиду. Согласно вариантам реализации целевой нуклеотид содержит по меньшей мере один экспансированный повтор CUG или CTG. Согласно вариантам реализации целевой нуклеотид представляет



собой ген, который содержит по меньшей мере один экспансированный повтор CTG. Согласно вариантам реализации целевой нуклеотид представляет собой РНК, которая содержит по меньшей мере один экспансированный повтор CUG. Согласно вариантам реализации РНК, которая содержит по меньшей мере один экспансированный повтор CUG, представляет собой последовательность пре-мРНК. Согласно вариантам реализации экспансированный повтор CUG соответствует экспансированному повтору CTG в гене, из которого транскрибируется пре-мРНК. Согласно вариантам реализации антисмысловое соединение связывается с экспансированным повтором CTG или экспансированным повтором CUG. Согласно вариантам реализации AC содержит 5-40 повторов CAG (например, 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 или 40 повторов). Согласно вариантам реализации AC содержит последовательность нуклеотидов, приведенную в **Таблице 2**, **Таблице 10** или **Таблице 11**. Согласно вариантам реализации AC содержит последовательность нуклеотидов, приведенную в **Таблице 2**.

Согласно вариантам реализации AC содержит по меньшей мере один модифицированный нуклеотид или нуклеиновую кислоту, выбранные из фосфотиоатного (PS) нуклеотида, фосфордиамидат-морфолинового (PMO) нуклеотида, блокированной нуклеиновой кислоты (LNA), пептидной нуклеиновой кислоты (PNA), нуклеотида, содержащего 2'-О-метил (2'-OMe)-модифицированный остов, 2'-О-метоксиэтил (2'-MOE)-нуклеотида, 2',4'-ограниченного этилнуклеотида (cEt), 2'-дезоксид-2'-фтор-бета-D-арабинонуклеиновой кислоты (2'-F-ANA) и их комбинаций. Согласно вариантам реализации AC содержит PMO-нуклеотид.

Согласно вариантам реализации предложены соединения, которые включают циклический пептид, содержащий от 6 до 12 аминокислот, причем по меньшей мере две аминокислоты циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты циклического пептида представляют собой ароматические гидрофобные аминокислоты и по меньшей мере две аминокислоты циклического пептида представляют собой незаряженные неароматические аминокислоты. Согласно вариантам реализации антисмысловое соединение (AC) комплементарно по меньшей мере части экспансированного повтора CUG в последовательности целевой мРНК. Согласно вариантам реализации AC представляет собой фосфордиамидат-морфолиновый (PMO) нуклеотид.

Согласно вариантам реализации по меньшей мере две заряженные аминокислоты циклического пептида представляют собой аргинин. Согласно вариантам реализации по меньшей мере две ароматические гидрофобные аминокислоты циклического пептида представляют собой фенилаланин, нафтилаланин (3-нафт-2-ил-аланин) или их комбинацию. Согласно вариантам реализации по меньшей мере две незаряженные неароматические аминокислоты циклического пептида представляют собой цитруллин, глицин или их комбинацию. Согласно вариантам реализации соединение представляет собой циклический пептид, содержащий от 6 до 12 аминокислот, причем две

аминокислоты циклического пептида представляют собой аргинин, по меньшей мере две аминокислоты представляют собой ароматические гидрофобные аминокислоты, выбранные из фенилаланина, нафтилаланина и их комбинаций, и по меньшей мере две аминокислоты представляют собой незаряженные неароматические аминокислоты, выбранные из цитруллина, глицина и их комбинаций.

Согласно вариантам реализации соединение содержит носитель для ускользания из эндосом, содержащий циклический пептид и экзоциклический пептид (EP). Согласно вариантам реализации EP конъюгирован с линкером в аминогруппе. Линкер может представлять собой линкер, описанный в настоящем документе. Согласно вариантам реализации EP конъюгирован с циклическим пептидом за счет линкера. Согласно вариантам реализации EP конъюгирован с AC за счет линкера. Согласно вариантам реализации EP конъюгирован с линкером, с помощью которого AC конъюгирован с циклическим пептидом.

Согласно вариантам реализации EP содержит от 2 до 10 аминокислот. Согласно вариантам реализации EP содержит от 4 до 8 аминокислотных остатков. Согласно вариантам реализации EP содержит 1 или 2 аминокислоты, содержащие боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму. Согласно вариантам реализации EP содержит 1, 2, 3 или 4 остатка лизина. Согласно вариантам реализации аминокислотная группа на боковой цепи каждого остатка лизина замещена трифторацетильной группой (-COCF<sub>3</sub>), аллилоксикарбонильной (Alloc), 1-(4,4-диметил-2,6-диоксоциклогексиден)этильной (Dde) или (4,4-диметил-2,6-диоксоциклогекс-1-илиден-3)-метилбутильной (ivDde) группой. Согласно вариантам реализации EP содержит по меньшей мере 2 аминокислотных остатка с гидрофобной боковой цепью. Согласно вариантам реализации аминокислотный остаток с гидрофобной боковой цепью выбран из валина, пролина, аланина, лейцина, изолейцина и метионина. Согласно вариантам реализации экзоциклический пептид содержит одну из следующих последовательностей: PKKKRKV; KR; RR; KKK; KGK; KBK; KBR; KRK; KRR; RKK; RRR; KKKK; KKRK; KRKK; KRRK; RKKR; RRRR; KGKK; KKGK; KKKKK; KKKRK; KBKBK; KKKRKV; PGKKRKV; PKGKRKV; PKKGRKV; PKKKGKV; PKKKRGV; или PKKKRKG. Согласно вариантам реализации экзоциклический пептид состоит из одной из следующих последовательностей: PKKKRKV; KR; RR; KKK; KGK; KBK; KBR; KRK; KRR; RKK; RRR; KKKK; KKRK; KRKK; KRRK; RKKR; RRRR; KGKK; KKGK; KKKKK; KKKRK; KBKBK; KKKRKV; PGKKRKV; PKGKRKV; PKKGRKV; PKKKGKV; PKKKRGV; или PKKKRKG. Согласно вариантам реализации экзоциклический пептид имеет структуру: Ac-P-K-K-K-R-K-V-.

Согласно вариантам реализации циклический пептид содержит от 4 до 12 аминокислот. Согласно вариантам реализации циклический пептид содержит от 6 до 12 аминокислот. Согласно вариантам реализации по меньшей мере две аминокислоты циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты циклического пептида представляют собой ароматические

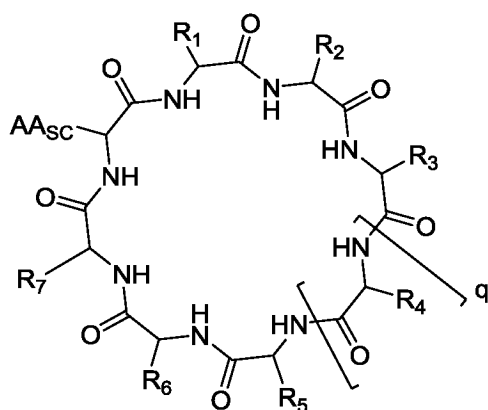
гидрофобные аминокислоты и по меньшей мере две аминокислоты циклического пептида представляют собой незаряженные неароматические аминокислоты. Согласно вариантам реализации по меньшей мере две заряженные аминокислоты циклического пептида представляют собой аргинин, по меньшей мере две ароматические гидрофобные аминокислоты циклического пептида представляют собой фенилаланин, нафтилаланин или их комбинации и по меньшей мере две незаряженные неароматические аминокислоты представляют собой цитруллин, глицин или их комбинации.

Согласно вариантам реализации циклический пептид содержит от 4 до 12 аминокислот, причем по меньшей мере две аминокислоты представляют собой аргинин и по меньшей мере две аминокислоты содержат гидрофобную боковую цепь, при условии, что циклический пептид не является циклическим пептидом, имеющим последовательность SEQ ID NO: 89-117. Согласно вариантам реализации циклический пептид не является циклическим пептидом, имеющим последовательность SEQ ID NO: 89-117.

Последовательности CPP и 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		

где F представляет собой L-фенилаланин, f представляет собой D-фенилаланин, Φ представляет собой L-3-(2-нафтил)-аланин, Ф представляет собой D-3-(2-нафтил)-аланин, R представляет собой L-аргинин, r представляет собой D-аргинин, Q представляет собой L-глутамин, q представляет собой D-глутамин, C представляет собой L-цистеин, U представляет собой L-селеноцистеин, W представляет собой L-триптофан, K представляет собой L-лизин, D представляет собой L-аспарагиновую кислоту, и Ω представляет собой L-норлейцин.

Согласно вариантам реализации циклический пептид имеет следующую структуру:



(A) или

ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или ароматическую или гетероароматическую боковую цепь аминокислоты;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

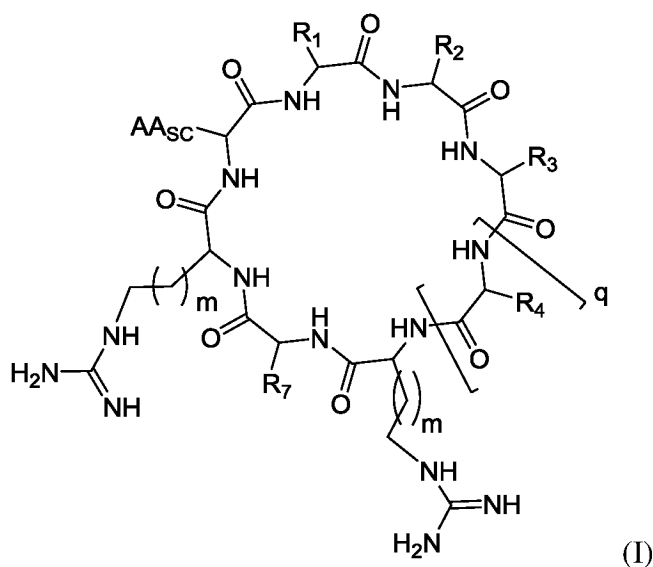
по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  представляет собой боковую цепь 3-гуанидин-2-аминопропионовой кислоты, 4-гуанидин-2-аминобутановой кислоты, аргинина, гомоаргинина, N-метиларгинина, N,N-диметиларгинина, 2,3-диаминопропионовой кислоты, 2,4-диаминобутановой кислоты, лизина, N-метиллизина, N,N-диметиллизина, N-этиллизина, N,N,N-триметиллизина, 4-гуанидинофенилаланина, цитруллина, N,N-диметиллизина,  $\beta$ -гомоаргинина, 3-(1-пиперидинил)аланина;

$AA_{sc}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение; и

$q$  представляет собой 1, 2, 3 или 4.

Согласно вариантам реализации по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляет собой незаряженную неароматическую боковую цепь аминокислоты. Согласно вариантам реализации по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляет собой H или боковую цепь цитруллина.

Согласно вариантам реализации циклический пептид имеет структуру Формулы I:



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

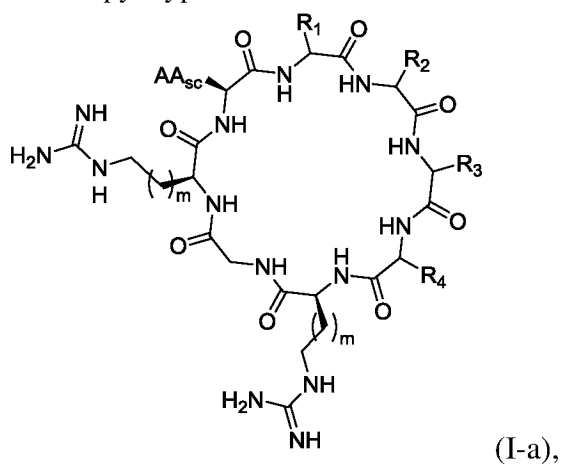
$R_4$  и  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

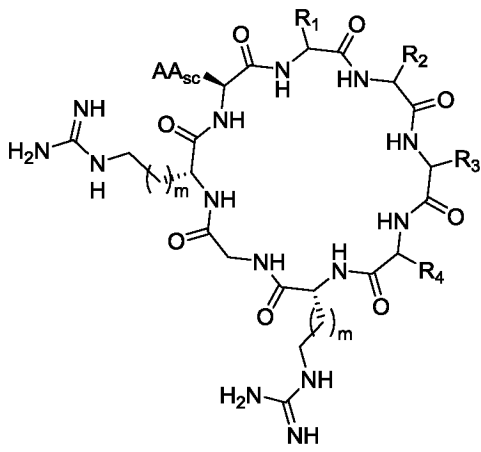
$AA_{sc}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение;

$q$  представляет собой 1, 2, 3 или 4; и

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3.

Согласно вариантам реализации циклический пептид Формулы (I) имеет одну из следующих структур:

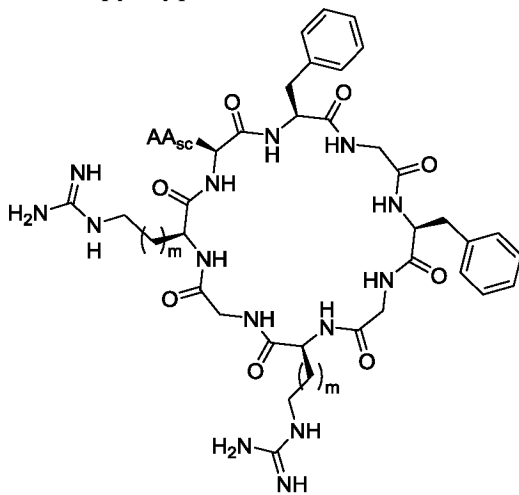




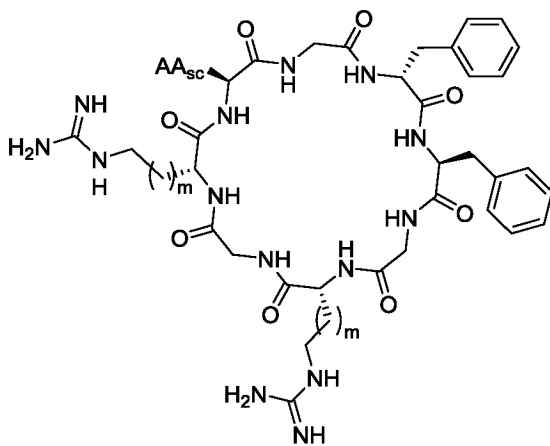
(I-b), или

ее протонированную форму.

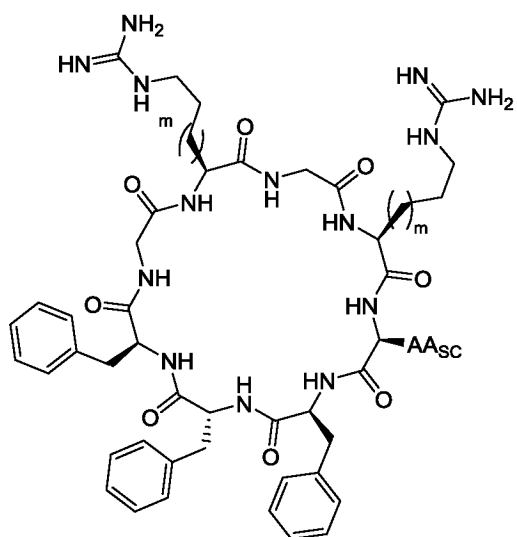
Согласно вариантам реализации циклический пептид Формулы (I) имеет одну из следующих структур:



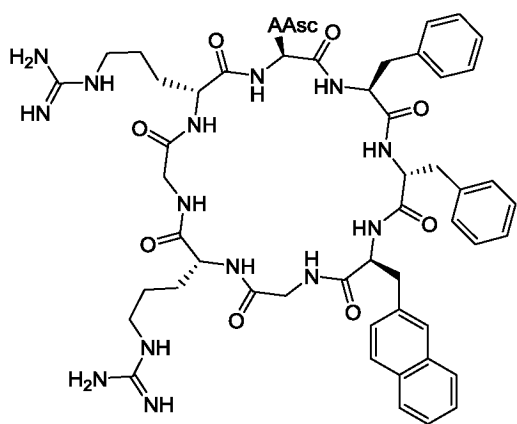
(I-1)



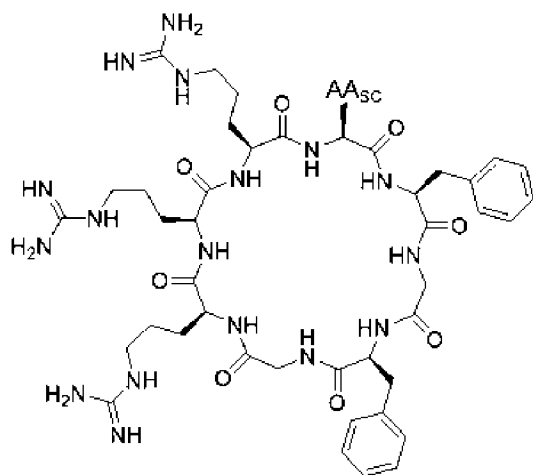
(I-2),



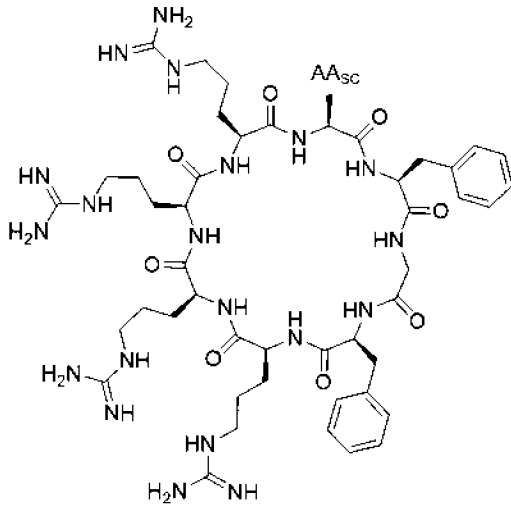
(I-3),



(I-4),



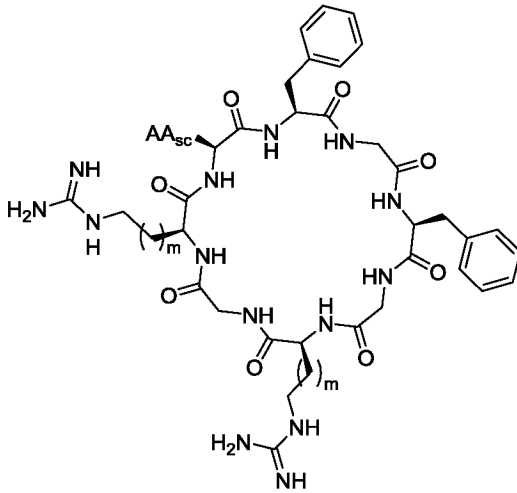
(I-5),



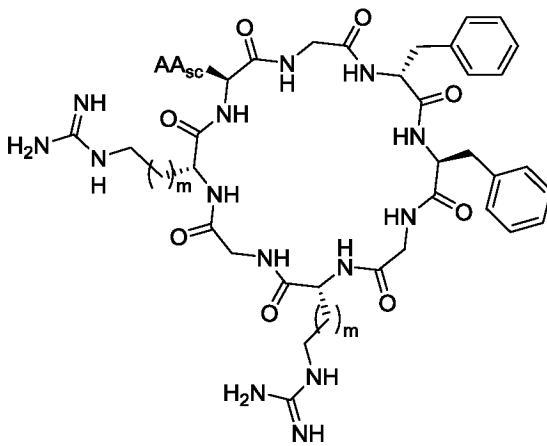
(I-6), или

ее протонированную форму.

Согласно вариантам реализации циклический пептид Формулы (I) имеет одну из следующих структур:



(I-1),

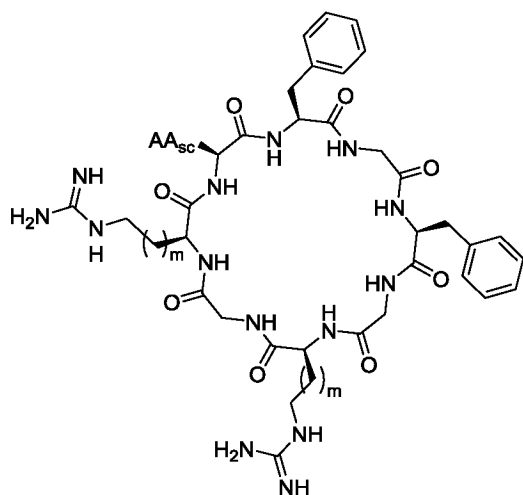


(I-2), или

ее протонированную форму.

Согласно вариантам реализации циклический пептид Формулы (I) имеет следующую структуру:

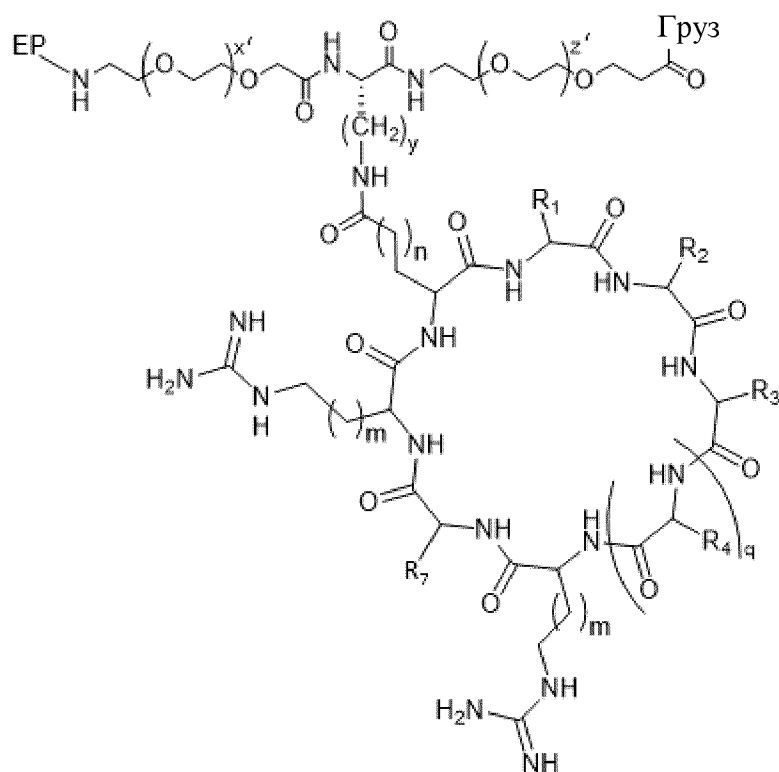




(I-1), или

ее протонированную форму.

Согласно вариантам реализации соединение имеет структуру Формулы С:



(C), или

ее протонированную форму или соль,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или боковую цепь, содержащую арильную или гетероарильную группу, причем по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой боковую цепь, содержащую арильную или гетероарильную группу;

$R_4$  и  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

EP представляет собой экзоциклический пептид;

каждый  $m$  независимо представляет собой целое число от 0 до 3;

n представляет собой целое число от 0 до 2;

x' представляет собой целое число от 1 до 23;

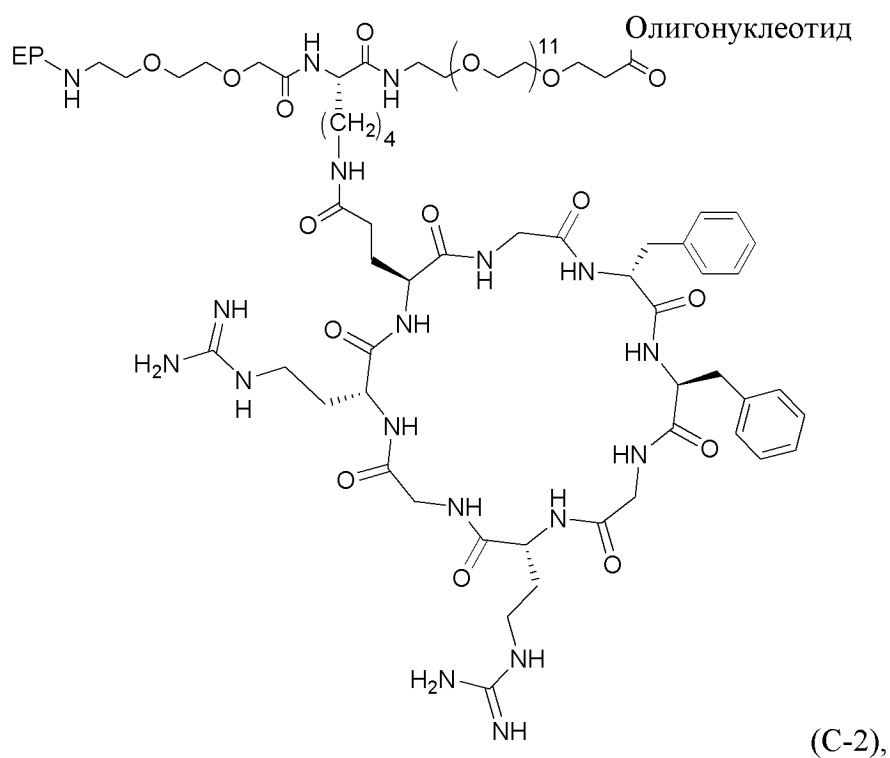
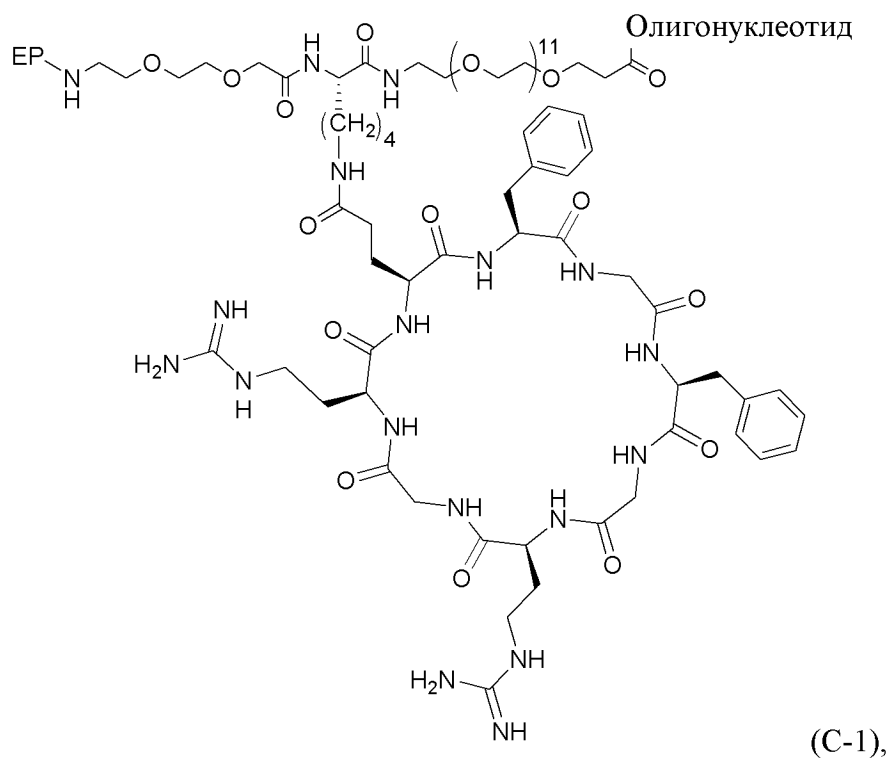
y представляет собой целое число от 1 до 5;

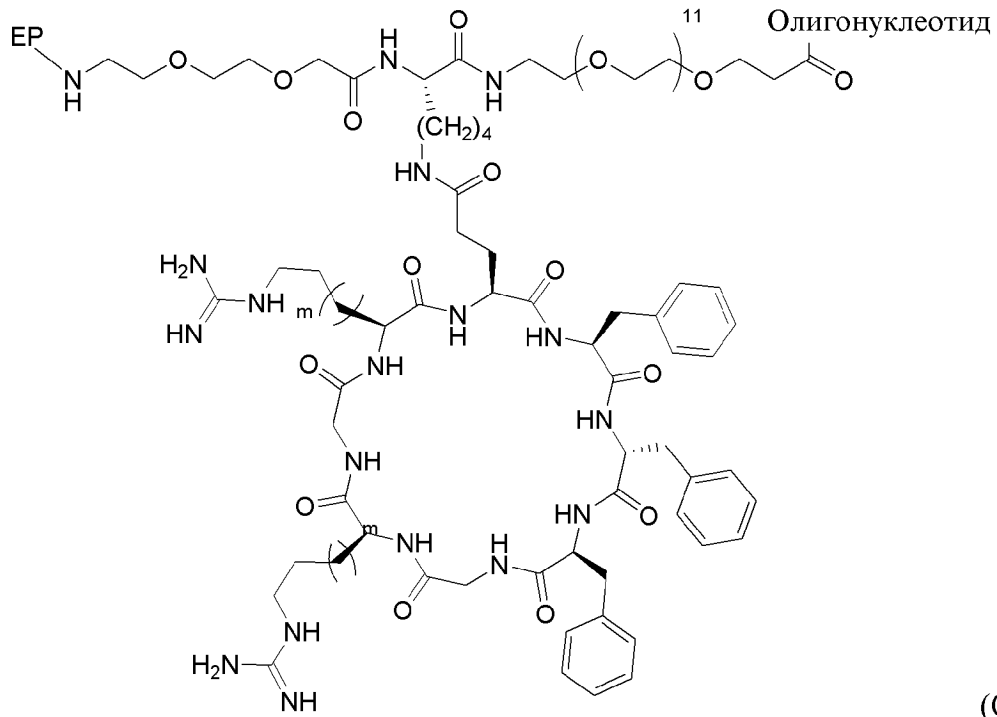
q представляет собой целое число от 1 до 4;

z' представляет собой целое число от 1 до 23, и

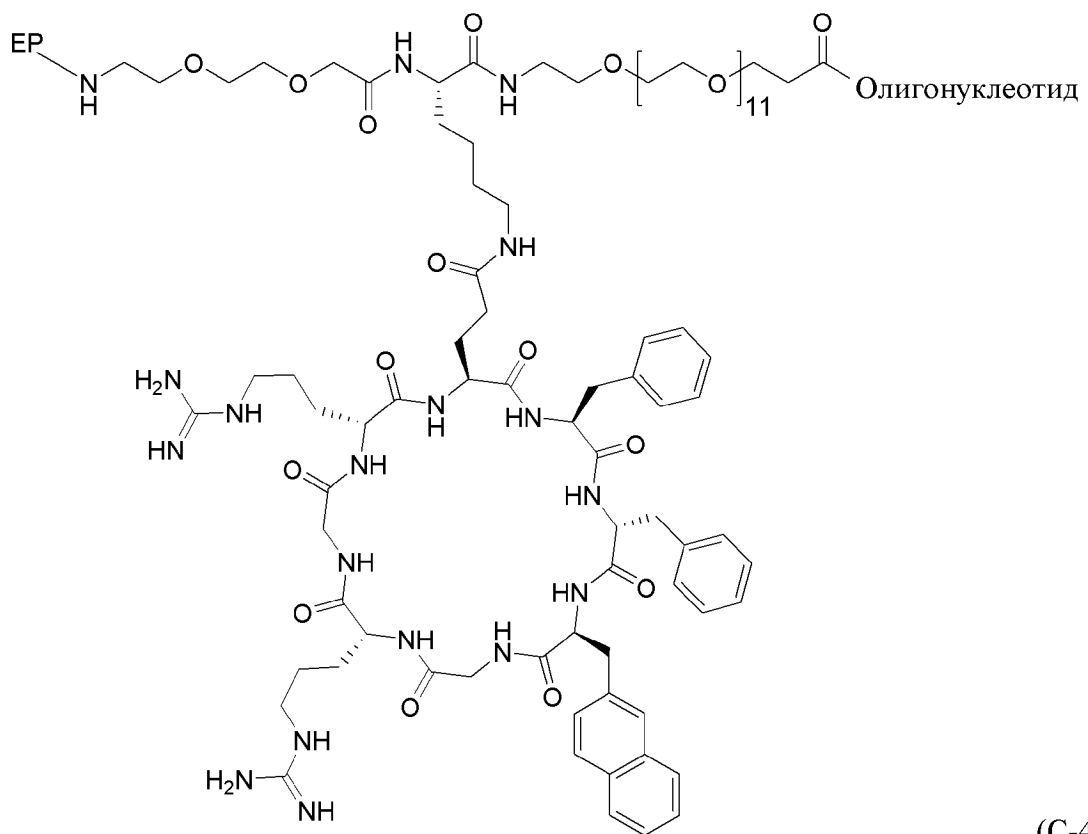
Груз представляет собой указанное антисмысловое соединение.

Согласно вариантам реализации соединение имеет одну или следующие структуры:





(C-3),



(C-4)

или ее протонированную форму или соль,  
 где EP представляет собой экзоциклический пептид, и  
 олигонуклеотид представляет собой антисмысловое соединение.

Согласно вариантам реализации олигонуклеотид соединения Формулы (С-1), (С-2), (С-3) или (С-4) содержит следующую последовательность: 5'-CAG CAG CAG CAG CAG CAG CAG-3'.

Согласно вариантам реализации EP соединения Формулы (С-1), (С-2), (С-3) или (С-4) содержит следующую последовательность: РКККРКV.

### **КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ**

На **ФИГ. 1** приведена схема, показывающая несколько стратегий нацеливания на повторы CUG в мРНК.

На **ФИГ. 2** показаны модифицированные нуклеотиды, использованные в антисмысловых олигонуклеотидах, описанных в настоящем документе. Структуры 1-3 (1=фосфотиоат; 2=(S<sub>C5</sub>-R<sub>p</sub>)- $\alpha,\beta$ -CAN; 3=PMO) представляют собой модификации фосфатного остова; 4 (2-тио-dT) представляет собой модификацию основания; 5-8 (5=2'-OMe-PHK; 6=2'-O-MOE-PHK; 7=2'-F-PHK; 8=2'-F-ANA) представляют собой 2'-модификации сахара; 9-11 представляют собой ограниченные нуклеотиды; 12-14 (9=LNA; 10=(S)-сЕТ; 11=tcДНК; 12=FHNA; 13=(S)5'-С-метил; 14=UNA) представляют собой дополнительную модификацию сахара; и 15-18 (15= E-VP; 16=метилфосфонат; 17=5'-фосфотиоат; 18=(S)-5'-С-метил с фосфатом) представляют собой модификации, стабилизирующие 5'-фосфат; 19 представляет собой морфолиновый сахар. Преобразовано из Khvorova, A., et al., Nat. Biotechnol. 2017 Mar; 35(3): 238-248.

На **ФИГ. 3А-3D** проиллюстрированы химические структуры для конъюгации, чтобы связать АС с проникающим в клетку циклическим пептидом. На **ФИГ. 3А** показано образование амидной связи между пептидами с карбоксильной группой или с TFP-активированным сложным эфиром и остатками первичного амина на 5'-конце АС. На **ФИГ. 3В** показана конъюгация модифицированного вторичным амином или первичным амином АС на 3'-конце и сложного эфира пептида и TFP посредством образования амидной связи. На **ФИГ. 3С** показана конъюгация пептида-азида с 5'-циклооктин-модифицированным АС за счет азид-алкинового циклоприсоединения без медного катализа. На **ФИГ. 3D** продемонстрирована еще одна примерная конъюгация между 3'-циклооктин-модифицированными АС или 3'-азид-модифицированными АС и CPP, содержащими фрагмент линкер-азид или линкер-алкин/циклооктин, за счет азид-алкинового циклоприсоединения без медного катализа или катализируемого медью азид-алкинового циклоприсоединения, соответственно (клик-реакция).

На **ФИГ. 4** показана химическая структура для конъюгации, чтобы связать АС и CPP с дополнительным линкером, содержащим фрагмент полиэтиленгликоля (PEG).

На **ФИГ. 5А-5D** приведены структуры аденинового (**5А**), цитозинового (**5В**), гуанинового (**5С**) и тиминового (**5D**) морфолиновых субъединичных мономеров, которые можно применять для синтеза фосфордиамидат-связанных морфолиновых олигомеров (PMO).

На **ФИГ. 6А-6F** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) MBNL1 (экзон 5; **ФИГ.**

**6A, 6B, 6E**) и CLASP1 (экзон 19; **ФИГ. 6B, 6D, 6F**) через 24 часа (**6A-6B**) и 48 часов (**6C-6D**) после обработки клеток HeLa-48 1 мкМ, 3 мкМ или 10 мкМ различных соединений РМО или РМО-EEV с использованием трансфекционного агента Endo-Porter (**6A-6C**) или без агента Endo-Porter (**6E-6F**). Исходную линию клеток HeLa и линию клеток HeLa-480, обработанную агентом Endo-Porter (**6A-6D**) или без указанного агента (**6E-6F**), включили в качестве контролей.

На **ФИГ. 7A-7B** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) MBNL1 (экзон 5; **ФИГ. 7A**) и CLASP1 (экзон 19; **ФИГ. 7B**) через 48 часов после обработки миобластов DM1 1 мкМ различных соединений РМО или РМО-EEV без реагента для трансфекции Endo-Porter. Два контроля, DM-04 без обработки Endo-Porter и DM-05 без обработки Endo-Porter, включили в качестве контролей.

На **ФИГ. 8A-8D** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) Atp2a1 (экзон 22; **ФИГ. 8A**), Nfix (экзон 7; **ФИГ. 8B**), Clcn1 (экзон 7a; **ФИГ. 8C**) и Mbnl1 (экзон 5; **ФИГ. 8D**) из ткани икроножной мышцы через неделю после обработки мышей HSA-LR (модель DM1 на мышах) РМО, 20 мг/кг РМО-EEV 221-1106 или 40 мг/кг РМО-EEV 221-1106. Мышей FVB/NJ (инбредная мышь дикого типа) и HSA-LR (без обработки) включили в качестве контрольных групп.

На **ФИГ. 9A-9D** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) Atp2a1 (экзон 22; **ФИГ. 9A**), Nfix (экзон 7; **ФИГ. 9B**), Clcn1 (экзон 7a; **ФИГ. 9C**) и Mbnl1 (экзон 5; **ФИГ. 9D**) из ткани четырехглавой мышцы через неделю после обработки мышей HSA-LR (модель DM1 на мышах) РМО, 20 мг/кг РМО-EEV 221-1106 или 40 мг/кг РМО-EEV 221-1106. Мышей FVB/NJ (инбредная мышь дикого типа) и HSA-LR (без обработки) включили в качестве контрольных групп.

На **ФИГ. 10A-10D** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) Atp2a1 (экзон 22; **ФИГ. 10A**), Nfix (экзон 7; **ФИГ. 10B**), Clcn1 (экзон 7a; **ФИГ. 10C**) и Mbnl1 (экзон 5; **ФИГ. 10D**) из ткани передней большеберцовой мышцы через неделю после обработки мышей HSA-LR (модель DM1 на мышах) РМО, 20 мг/кг РМО-EEV 221-1106 или 40 мг/кг РМО-EEV 221-1106. Мышей FVB/NJ (инбредная мышь дикого типа) и HSA-LR (без обработки) включили в качестве контрольных групп.

На **ФИГ. 11A-11F** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) MBNL1 (экзон 5, **ФИГ. 11A**), SOS1 (экзон 25, **ФИГ. 11B**), IR (экзон 11, **ФИГ. 11C**), DMD (экзон 78, **ФИГ. 11D**), BIN1 (экзон 11, **ФИГ. 11E**) и LDB3 (экзон 11, **ФИГ. 11F**) после обработки мышечных клеток, полученных от пациентов, страдающих DM1, различными концентрациями (10 мкМ, 3 мкМ, 1 мкМ, 0,3 мкМ) DMPK CUG-нацеленных EEV-РМО (CUG<sup>exp</sup> 197-777 и CUG<sup>exp</sup> 221-1106). Мышечные клетки из двух групп, от здоровых людей (отрицательный

контроль) и пациентов, страдающих DM1 (положительный контроль), тестировали на события альтернативного сплайсинга РНК в качестве контроля. Все данные собирали из трех отдельных экспериментов (n=3). Использовали Т-критерий для сравнения обработанных и необработанных мышечных трубочек DM1; \*p<0,05; \*\*p<0,01; \*\*\*p<0,001.

На **ФИГ. 12А-12F** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения или исключения экзона) MBL1 (экзон 5, **ФИГ. 12А**), SOS1 (экзон 25, **ФИГ. 12В**), INSR (экзон 11, **ФИГ. 12С**), DMD (экзон 78, **ФИГ. 12D**), BIN1 (экзон 11, **ФИГ. 12Е**) и LDB3 (экзон 11, **ФИГ. 12F**) после обработки миобластов и мышечных трубочек, полученных от пациентов, страдающих DM1, 10 мкМ, 3 мкМ или 1 мкМ DMPK CUG-нацеленного EEV-PMO 197-777. Здоровые клетки пациентов и клетки DM1 использовали в качестве контролей. Все данные собирали из трех отдельных экспериментов (n=3). Использовали Т-критерий для сравнения обработанных и необработанных мышечных трубочек DM1; \*p<0,05; \*\*p<0,01; \*\*\*p<0,001.

На **ФИГ. 13А-13В** показаны относительные уровни мРНК после обработки мышей HSA-LR различными концентрациями PMO-EEV 221-1120. На **ФИГ. 13А** показан относительный уровень мРНК для икроножной мышцы, трехглавой мышцы, передней большеберцовой мышцы и диафрагмы. На **ФИГ. 13В** показаны относительные уровни мРНК в диафрагме.

На **ФИГ. 14А-14С** показаны относительные уровни мРНК в тканях четырехглавой (**14А**), икроножной (**14В**), трехглавой (**14С**) и передней большеберцовой мышцы (**14D**) после обработки мышей HSA-LR различными концентрациями PMO-EEV 221-1120.

На **ФИГ. 15А-15D** показан индекс сплайсинга при DM1 у мышей (mDSI) для различных генов в тканях четырехглавой (**ФИГ. 15А**), икроножной (**ФИГ. 15В**), трехглавой (**ФИГ. 15С**) и передней большеберцовой мышцы (**ФИГ. 15D**) после обработки мышей HSA-LR различными концентрациями PMO-EEV 221-1120.

На **ФИГ. 16А-16С** показана распространенность очагов РНК в передней большеберцовой мышце у мышей HSA-LR после обработки EEV-PMO 221-1120 (EEV-PMO-DM1-3; DM1-3) или без обработки. На **ФИГ. 16А-16В** показаны изображения ткани передней большеберцовой мышцы, окрашенной на очаги РНК CUG (красный) и ядра (синий). На **ФИГ. 16С** показан график количественной оценки процента ядер, которые имеют очаги CUG на основании данных, связанных с изображениями на **ФИГ. 16А-16В**.

На **ФИГ. 17А-17F** показаны графики зависимости ответа от дозы для уровней лекарственного средства в тканях четырехглавой (**17А**), трехглавой (**17В**), сердечной (**17С**), икроножной (**17D**), передней большеберцовой мышцы (**17Е**), диафрагмы (**17F**), в головном мозге (**17H**), печени (**17I**) и почке (**17J**) после обработки мышей HSA-LR различными концентрациями EEV-PMO-DM1-3. На **ФИГ. 17К** показано воздействие лекарственного средства на различные ткани при уровне дозы 60 мг/кг.

На **ФИГ. 18** показано дозозависимое уменьшение миотонии у мышей HSA-LR через 7 дней после обработки EEV-PMO-DM1-3 при 15, 30, 60 и 90 мг/кг.

На **ФИГ. 19А-19D** показаны графики результатов анализа основных компонентов, в котором сравнивали экспрессию генов у непораженных мышей (ДТ), мышей DM1 (HSA-LR) и мышей HSA-LR, обработанных PMO-EEV 221-1120. **ФИГ. 19А** и **19С** представляют собой графики, показывающие три основных компонента, и **ФИГ. 19В** и **19D** представляют собой графики, показывающие два основных компонента.

На **ФИГ. 20А-20В** показаны тепловые карты дифференциально экспрессируемых генов у непораженных мышей (ДТ), мышей DM1 (HSA-LR) и мышей HSA-LR, обработанных 60 мг/кг PMO-EEV 221-1120. **ФИГ. 20А** представляет собой кластерную тепловую карту, показывающую 513 дифференциально экспрессируемых генов. **ФИГ. 20В** представляет собой кластерную тепловую карту, показывающую 40 генов, которые, как известно, содержат повторы CTG·CUG.

**ФИГ. 21** представляет собой вулканический график, показывающий общее изменение транскрипции у необработанных мышей HSA-LR и мышей, обработанных PMO-EEV 221-1120.

**ФИГ. 22А-22Е** представляют собой графики, на которых показан результат анализа основных компонентов для генов Scube2 (**22А**), Greb1 (**22В**), Ttc7 (**22С**), Txlnb(CUG)9 (**22D**) и Ndr3 (**22Е**) у непораженных мышей, мышей HSA-LR и мышей HSA-LR, обработанных PMO-EEV 221-1120.

На **ФИГ. 23А-23D** показаны данные РНК-секвенирования (РНКseq) для Atp2a1 (**23А**; экзон 22 заключен в рамку), Clcn1 (**23В**; экзон 7а заключен в рамку), Nfix (**23С**; экзон 7 заключен в рамку) и Mbn1 (**23D**; экзон 5 заключен в рамку) для непораженных мышей (ДТ-солевой раствор), мышей HSA-LR (HSA-LR солевой раствор) и мышей HSA-LR, обработанных PMO-EEV 221-1120. Для каждой группы лечения показаны два считывания.

На **ФИГ. 24** показано процентное значение индекса сплайсинга (PSI) отдельных экзонов для различных представляющих интерес генов у непораженных мышей (ДТ-солевой раствор), мышей HSA-LR (HSA-LR солевой раствор) и мышей HSA-LR, обработанных PMO-EEV 221-1120.

На **ФИГ. 25А-25D** показаны уровни лекарственного средства у мышей HSA-LR, обработанных 80 мг/кг (60 мг/кг олигонуклеотида, 80 мг/кг полного лекарственного средства) EEV-PMO-DM1-3 через 1-4 недели в тканях передней большеберцовой (**25А**), икроножной (**25В**), трехглавой (**25С**) и четырехглавой (**25D**) мышцы.

На **ФИГ. 26А-26D** показаны уровни лекарственного средства у мышей после обработки мышей HSA-LR однократной дозой 80 мг/кг EEV-PMO-DM1-3. На **ФИГ. 26А-26В** показаны уровни лекарственного средства в печени через 1-12 недель после обработки. На **ФИГ. 26С-26D** показаны уровни лекарственного средства в почке через 1-12 недель после обработки.

**ФИГ. 27А-27С** представляют собой графики, показывающие уровень включения экзона для MBNL1 (экзон 5; **26А**), SOS1 (экзон 25; **26В**) и NFIX (экзон 7; **26С**) после обработки мышечных клеток, полученных от пациентов, страдающих DM1, 30 мкМ EEV-PMO-DM1-3.

На **ФИГ. 28А-28С** показано, что EEV-PMO-DM1-3 уменьшает ядерные очаги CUG (зеленый) в ядре (синий) в мышечных клетках, полученных от пациентов, страдающих DM1. На **ФИГ. 28А-28В** показаны изображения мышечных клеток, полученных от пациентов, страдающих DM1, которые обрабатывали EEV-PMO-DM1-3 или оставляли без обработки. **ФИГ. 28С** представляет собой количественное определение числа очагов CUG на ядро для данных, связанных с изображениями на **ФИГ. 28А**.

На **ФИГ. 29А-29В** показаны исходные данные (**29А**) и нормированные данные (**29В**) люминесцентного анализа жизнеспособности CELLTITER-GLO, в котором клетки RPTEC обрабатывали различными концентрациями PMO-DM1 или EEV-PMO-DM1-3. В качестве положительного контроля использовали мелиттин.

На **ФИГ. 30А-30С** показаны изображения, отображающие очаги РНК с повторами CUG в клетках, полученных от пациентов, страдающих DM1 (**30А**), и клетках, полученных от пациентов, страдающих DM1, обработанных EEV-PMO 221-1113 (**30В**). Клетки окрашивали на ядра (синий; Hoechst) и очаги РНК CUG (зеленый). **ФИГ. 30С** представляет собой график очагов РНК CUG на ядерную область для данных, связанных с изображениями на **ФИГ. 30А-30В**.

На **ФИГ. 31А-31В** показана распространенность очагов РНК CUG7 в HeLa, необработанных клетках HeLa480 и клетках HeLa480, обработанных EEV-PMO 221-1113. На **ФИГ. 31А** показаны изображения клеток, окрашенных на очаги РНК CUG7 (зеленый) и ядра (синий). **ФИГ. 31В** представляет собой график количественного определения очагов CUG7 на ядро для данных, связанных с изображениями на **ФИГ. 31А**.

**ФИГ. 32А-32С** представляют собой графики процента включения экзона 5 в MBNL1 (**32А**), экзона 25 в SOS1 (**32В**) и экзона 7 в NFIX (**32С**) после обработки клеток, полученных от пациентов, страдающих DM1, 30 мкМ EEV-PMO 221-1113.

На **ФИГ. 33А-33Е** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включение экзона) MBNL1 (экзон 5; **33А**), SOS1 (экзон **25**; **33В**), CLASP1 (экзон 19, **33С**), NFIX (экзон 7, **33Д**) и INSR (экзон 11, **33Е**) после обработки мышечных клеток, полученных от пациентов, страдающих DM1, различными концентрациями PMO-EEV 221-1113. Т-критерий использовали для определения значимости; \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ .

На **ФИГ. 34А-34Д** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включения экзона) Atp2a1 (экзон 22, **34А**), Nfix (экзон 7, **34В**), Clcn1 (экзон 7а, **34С**) и Mbnl1 (экзон 5, **34Д**) в ткани икроножной мышцы мышей, обработанных различными концентрациями PMO 221 или EEV-PMO 221-1106.

На **ФИГ. 35А-35С** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включение или исключение экзона) Mbnl1 (экзон 5, **35А**),



Nfix (экзон 7, **35B**) и Atp2a1 (экзон 22, **35C**) в ткани передней большеберцовой мышцы мышцей HSA-LR, обработанных РМО-EEV 0221-1121 (21-мер) или РМО-EEV 0325-1121 (24-мер).

На **ФИГ. 36A-36C** показан анализ методом ОТ-ПЦР событий альтернативного сплайсинга РНК (например, включение экзона) Mbnl1 (экзон 5, **36A**), Nfix (экзон 7, **36B**) и Atp2a1 (экзон 22, **36C**) в ткани икроножной мышцы мышцей HSA-LR, обработанных РМО-EEV 0221-1121 (21-мер) или РМО-EEV 0325-1121 (24-мер).

На **ФИГ. 37** показан РМО-0221a, основной метаболит РМО-EEV 220-1120, детектированный *in vivo*.

На **ФИГ. 38A-38B** показан процент включения экзона в передней большеберцовой мышце (**38A**) и икроножной мышце (**38B**) для MBNL1 (экзон 5) после обработки клеток HeLa480 различными концентрациями EEV-РМО 221-1120.

На **ФИГ. 39A-39B** показан процент включения экзона в передней большеберцовой мышце (**39A**) и икроножной мышце (**39B**) для NFIX (экзон 7) после обработки клеток HeLa480 различными концентрациями EEV-РМО 221-1120.

На **ФИГ. 40A-40B** показан процент включения экзона в передней большеберцовой мышце (**40A**) и икроножной мышце (**40B**) для Atp2a1 (экзон 22) после обработки клеток HeLa480 различными концентрациями EEV-РМО 221-1120.

На **ФИГ. 41** показаны изображения (**41A**), отображающие очаги РНК с повторами CUG в клетках HeLa480 после обработки различными концентрациями EEV-РМО 221-1120. **ФИГ. 41B** представляет собой график очагов РНК на ядерную область для данных, связанных с изображениями на **ФИГ. 41A**.

На **ФИГ. 42A-42D** показаны относительные уровни мРНК повторов r(CUG480), **42A**, относительные уровни мРНК DMPK (**42B**), процент включения экзона 5 MBNL1 (**42C**) и процент включения экзона 25 в SOS1 (**42D**) в клетках HeLa480 после обработки различными концентрациями EEV-РМО 221-1120.

**ФИГ. 43** представляет собой столбчатую диаграмму, показывающую примеры генов, экспрессируемых в мышечной ткани, которые, как известно, содержат повторы CTG-CUG.

На **ФИГ. 44A-44D** показано фенотипическое уменьшение миотонии в модели на мышцах HSA-LR, обработанных 20 мг/кг РМО-EEV 221-1106. На **ФИГ. 44A** и **44C** показаны графики релаксации. На **ФИГ. 44B** показана примерная исходная кривая усилий. На **ФИГ. 44D** показаны типичные электромиографические кривые.

## ПОДРОБНОЕ ОПИСАНИЕ

### Соединения

Согласно вариантам реализации предложены соединения, которые модулируют уровень и/или активность транскрипта гена, содержащего экспансированный тринуклеотидный повтор CUG. Согласно вариантам реализации соединения согласно настоящему изобретению включают по меньшей мере один проникающий в клетку циклический пептид (сСРР) и терапевтический фрагмент (ТМ). сСРР облегчает

проникновение ТМ в клетку. Согласно вариантам реализации соединение включает носитель для ускользания из эндосом (EEV), который содержит сСРР и экзоциклический пептид (EP). сСРР или EEV могут обеспечивать проникновение ТМ в цитозоль или клеточный компартмент для взаимодействия с целевым транскриптом.

### **Терапевтические фрагменты**

Обычно ТМ представляет собой эффекторный фрагмент, который вызывает ответ. Согласно вариантам реализации ТМ вызывает ответ путем модуляции экспрессии, активности и/или уровня целевого транскрипта и/или целевого белка. Согласно вариантам реализации целевой транскрипт включает экспансированный тринуклеотидный повтор CUG. Согласно вариантам реализации ТМ модулирует уровни или целевой транскрипт и/или целевой белок внутри клетки. Согласно вариантам реализации ТМ уменьшает уровень целевого транскрипта и/или целевого белка в клетке.

Согласно вариантам реализации ТМ модулирует активность целевого транскрипта путем снижения аффинности между целевым транскриптом и одним или более белками, которые связываются с целевым транскриптом. Путем снижения аффинности между целевым транскриптом и одним или более белками ТМ может эффективно модулировать активность одного или более белков, которые в противном случае были бы ассоциированы с целевым транскриптом. Например, если один или более белков не связаны с целевым транскриптом, они доступны для выполнения своих функций в отношении других молекул. Например, если один или более белков вовлечены в процессинг пре-мРНК, снижение аффинности одного или более белков в отношении транскрипта, содержащего экспансированный повтор CUG, может позволить одному или более белкам процессировать транскрипты пре-мРНК, которые не содержат экспансированных повторов CUG. Таким образом, ТМ может модулировать активность, экспрессию и/или уровни нисходящих генов (генов, которые не содержат экспансированного повтора CTG), которые регулируются одним или более белками, взаимодействие которых с целевым транскриптом нарушается ТМ.

Согласно вариантам реализации ТМ содержит олигонуклеотид, пептид, антитело и/или малую молекулу. Класс и идентичность ТМ зависят от механизма, используемого для модуляции уровня и/или активности целевого транскрипта, который включает экспансированный тринуклеотидный повтор CUG.

### **Антисмысловое соединение**

Согласно различным вариантам реализации соединения, раскрытые в настоящем документе, содержат проникающий в клетку пептид (СРР), конъюгированный с антисмысловым соединением (АС).

Термин «антисмысловое соединение» относится к олигонуклеотидной последовательности, которая комплементарна или по меньшей мере частично комплементарна целевой нуклеотидной последовательности. АС представляет собой олигонуклеотид, который включает природные ДНК-основания, модифицированные ДНК-основания, природные РНК-основания, модифицированные РНК-основания,

природные РНК-сахара, модифицированные РНК-сахара, природные ДНК-сахара, модифицированные ДНК-сахара, природные межнуклеозидные связи, модифицированные межнуклеозидные связи или любую их комбинацию. АС включают, но не ограничиваются перечисленными, антисмысловые олигонуклеотиды РНКи, микроРНК, антагомиры, аптамеры, рибозимы, иммуностимулирующие олигонуклеотиды, олигонуклеотиды-ловушки, супермир, миметики миРНК, ингибиторы миРНК, адаптеры U1 и их комбинации.

Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна целевому транскрипту, который содержит экспансированный тринуклеотидный повтор CUG. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна экспансированному тринуклеотидному повтору CUG в последовательности целевой мРНК. Несколько заболеваний ассоциированы с экспансированными тринуклеотидными повторами CUG, например, миотоническая дистрофия 1 типа (DM1), эндотелиальная дистрофия роговицы Фукса (FECF), спиноцеребеллярная атаксия-8 (SCA8) и болезнь, подобная болезни Хантингтона (HDL2). В **Таблице 1** приведены примеры нарушений, ассоциированных с нуклеотидными повторами, и характеристики генов с экспансированными нуклеотидными повторами, ассоциированных с такими нарушениями. В следующем документе описаны примерные олигонуклеотиды для лечения заболеваний, ассоциированных с тандемными повторами, и он полностью включен в настоящий документ посредством ссылки: Zain et al. *Neurotherapeutics*. 2019; 16(2): 248-262; Zarouchlioti et al. *Am J Hum Genet*. 2018;102(4):528-539; Fautsch et al. *Prog Retin Eye Res*. 2021; 81:100883.

*Таблица 1:* Заболевания, ассоциированные с экспансированными тринуклеотидными повторами CUG

Заболевание (сокращение)	Ген	Длина нормального повтора	Длина экспансированного повтора	Генный продукт	Последовательность повтора	Расположение повтора
DM1	DMPK	5-35	> 50	Протеинкиназа миотонической дистрофии	CTG•CAG	3'-НТО
FECF	TCF4	< 30	> 40	Фактор транскрипции 4	CTG•CAG	Интрон 3
SCA8	ATXN8O S и/или ATXN8	15-50	> 50	Атаксин 8 и противоположная цепь атаксина 8	CTG•CAG	3'-НТО

HDL2	<i>JPH3</i>	6-27	>40	Джанктофилин 3	CTG•CAG	3'-НТО
------	-------------	------	-----	-------------------	---------	--------

Согласно вариантам реализации АС включает нуклеотидную последовательность, по меньшей мере частично комплементарную нуклеотидной последовательности, которая находится в пределах целевого транскрипта мРНК, который включает экспансированный тринуклеотидный повтор CTG•CUG. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна экспансированному тринуклеотидному повтору CTG•CUG в целевом транскрипте мРНК.

Согласно вариантам реализации АС включает нуклеотидную последовательность, по меньшей мере частично комплементарную нуклеотидной последовательности, которая находится в пределах целевого транскрипта DMPK1, который включает экспансированный тринуклеотидный повтор CTG•CUG. Согласно вариантам реализации АС включает нуклеотидную последовательность, по меньшей мере частично комплементарную нуклеотидной последовательности, которая находится в пределах целевого транскрипта TCF4, который включает экспансированный тринуклеотидный повтор CTG•CUG. Согласно вариантам реализации АС включает нуклеотидную последовательность, по меньшей мере частично комплементарную нуклеотидной последовательности, которая находится в пределах целевого транскрипта ATXN8OS/ATXN8, который включает экспансированный тринуклеотидный повтор CTG•CUG. Согласно вариантам реализации АС включает нуклеотидную последовательность, по меньшей мере частично комплементарную нуклеотидной последовательности, которая находится в пределах целевого транскрипта *JPH3*, который включает экспансированный тринуклеотидный повтор CTG•CUG.

Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна тринуклеотидному повтору в 3'-НТО целевого транскрипта мРНК. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна экспансированному тринуклеотидному повтору CTG•CUG в 3'-НТО целевого транскрипта DMPK1. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна экспансированному тринуклеотидному повтору CTG•CUG в 3'-НТО целевого транскрипта ATXN8OS/ATXN8. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна экспансированному тринуклеотидному повтору CTG•CUG в 3'-НТО целевого транскрипта *JPH3*.

Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна тринуклеотидным повторам, таким как повторы CTG•CUG. Согласно вариантам реализации целевая нуклеотидная последовательность содержит по меньшей мере один экспансированный тринуклеотидный повтор (например, повторы CTG•CUG). Согласно вариантам реализации

целевая нуклеотидная последовательность содержит по меньшей мере 40, по меньшей мере 45, по меньшей мере 50, по меньшей мере 60, по меньшей мере 70, по меньшей мере 80, по меньшей мере 90, по меньшей мере 100, по меньшей мере 150, по меньшей мере 200, по меньшей мере 300, по меньшей мере 400, по меньшей мере 500, по меньшей мере 600, по меньшей мере 700, по меньшей мере 800, по меньшей мере 900, по меньшей мере 1000 или по меньшей мере 2000 тринуклеотидных повторов CTG·CUG. Согласно вариантам реализации экспансированный тринуклеотидный повтор находится в 3'-НТО целевой нуклеотидной последовательности.

Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с по меньшей мере 2, по меньшей мере 3, по меньшей мере 4, по меньшей мере 5, по меньшей мере 6, по меньшей мере 7, по меньшей мере 8, по меньшей мере 9, по меньшей мере 10, по меньшей мере 11, по меньшей мере 12, по меньшей мере 13, по меньшей мере 14, по меньшей мере 15, по меньшей мере 16, по меньшей мере 17, по меньшей мере 18, по меньшей мере 19, по меньшей мере 20 и до 50, до 100, до 150, до 200, до 300, до 400, до 500, до 600, до 700, до 800, до 900, до 1000 или до 2000 тринуклеотидных повторов в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5-10 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5-9 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5-8 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5-7 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5-6 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 5 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации AC включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и

может гибридизоваться с 6 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 7 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 8 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться с 9 тринуклеотидными повторами в целевом транскрипте. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая

Согласно вариантам реализации АС может включать нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в любом местоположении в целевом транскрипте. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в 3'-НТО целевого транскрипта. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в 3'-НТО целевого транскрипта DMPK1, SCA8 и/или HDL2. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в интроне целевого транскрипта. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в интроне 3 целевого транскрипта TCF4. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в локусе STG18.1 транскрипта TCF4. Согласно вариантам реализации АС включает нуклеотидную последовательность, которая по меньшей мере частично комплементарна и может гибридизоваться по меньшей мере с частью смежных экспансированных тринуклеотидных повторов, присутствующих в экзоне целевого транскрипта.

Согласно вариантам реализации АС имеет длину 5 или более, 10 или более, 15 или более, 20 или более, 25 или более, 30 или более, 35 или более, 40 или более или 45 или более нуклеиновых кислот. Согласно вариантам реализации АС имеет длину 50 или менее, 45 или менее, 40 или менее, 35 или менее, 30 или менее, 25 или менее, 20 или

менее, 15 или менее или 10 или менее нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 5 до 50, от 5 до 45, от 5 до 40, от 5 до 35, от 5 до 30, от 5 до 25, от 5 до 20, от 5 до 15 или от 5 до 10 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 10 до 50, от 10 до 45, от 10 до 40, от 10 до 35, от 10 до 30, от 10 до 25, от 10 до 20 или от 10 до 15 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 15 до 50, от 15 до 45, от 15 до 40, от 15 до 35, от 15 до 30, от 15 до 25 или от 15 до 20 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 20 до 50, от 20 до 45, от 20 до 40, от 20 до 35, от 20 до 30 или от 20 до 25 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 25 до 50, от 25 до 45, от 25 до 40, от 25 до 35 или от 25 до 30 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 30 до 50, от 30 до 45, от 30 до 40 или от 30 до 35 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 35 до 50, от 35 до 45 или от 35 до 40 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 40 до 50 или от 40 до 45 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину от 45 до 50 нуклеиновых кислот. Согласно вариантам реализации АС имеет длину 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 или 50 нуклеиновых кислот.

Согласно вариантам реализации АС имеет 100% комплементарность с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС не имеет 100% комплементарности с целевой нуклеотидной последовательностью. В настоящем документе термин «процент комплементарности» относится к числу нуклеиновых оснований (например, природных нуклеиновых оснований или модифицированных нуклеиновых оснований) АС, которые характеризуются комплементарностью нуклеиновых оснований с соответствующим нуклеиновым основанием олигомерного соединения или нуклеиновой кислоты (например, целевой нуклеотидной последовательности), разделенному на общую длину (число нуклеиновых оснований) АС. Специалист в данной области техники установит, что возможно включение несовпадений без устранения активности антисмыслового соединения.

Согласно вариантам реализации АС включает 20% или менее, 15% или менее, 10% или менее, 5% или менее или ноль несовпадений с целевой нуклеотидной последовательностью. Согласно некоторым вариантам реализации АС включает 5% и более, 10% и более или 15% и более несовпадений. Согласно вариантам реализации АС включает от нуля до 5%, от нуля до 10%, от нуля до 15% или от нуля до 20% несовпадений с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС включает от 5 до 10%, от 5% до 15% или от 5% до 20% несовпадений с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС включает от 10% до 15% или от 10% до 20% несовпадений с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС включает от 10% до 20% несовпадений с целевой нуклеотидной последовательностью.

Согласно вариантам реализации АС имеет 80% или более, 85% или более, 90% или более, 95% или более, 96% или более, 97% или более, 98% или более, или 99% или более комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет 100% или менее, 99% или менее, 98% или менее, 97% или менее, 96% или менее, 95% или менее, 90% или менее, 85% или менее комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 80% до 100%, от 80% до 99%, от 80% до 98%, от 80% до 97% от 80% до 96%, от 80% до 95%, от 80% до 90% или от 80% до 85% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 85% до 100%, от 85% до 99%, от 85% до 98%, от 85% до 97%, от 85% до 96%, от 85% до 95% или от 85% до 90% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 90% до 100%, от 90% до 99%, от 90% до 98%, от 90% до 97%, от 90% до 96% или от 90% до 95% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 95% до 100%, от 95% до 99%, от 95% до 98%, от 95% до 97% или от 95% до 96% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 96% до 100%, от 96% до 99%, от 96% до 98% или от 96% до 97% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 97% до 100%, от 97% до 99% или от 97% до 98% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 98% до 100% или от 98% до 99% комплементарности с целевой нуклеотидной последовательностью. Согласно вариантам реализации АС имеет от 99% до 100% комплементарности с целевой нуклеотидной последовательностью.

Согласно вариантам реализации включение нуклеотидных модификаций для изменения аффинности приводит к большему числу несовпадений по сравнению с немодифицированным соединением. Аналогичным образом, некоторые олигонуклеотидные последовательности могут быть более устойчивыми к несовпадениям, чем другие олигонуклеотидные последовательности. Обычный специалист в данной области техники способен определить подходящее число несовпадений между АС и целевой нуклеотидной последовательностью, например, путем определения температуры плавления ( $T_m$ ).  $T_m$  или  $\Delta T_m$  можно рассчитать с помощью методик, которые известны специалисту в данной области техники. Например, методики, описанные в Freier et al. (Nucleic Acids Research, 1997, 25, 22: 4429-4443), позволяют обычному специалисту в данной области техники оценивать нуклеотидные модификации в отношении их способности повышать температуру плавления дуплекса РНК:ДНК.

Согласно вариантам реализации АС включает нуклеотидную последовательность, которая сама по себе представляет собой тринуклеотидный повтор, то есть тринуклеотидный повтор CAG. Обратнoкомплементарная последовательность 5'-CAG-3' имеет 100% комплементарность и может гибридизоваться с тринуклеотидным повтором 5'-CUG-3'. Согласно вариантам реализации АС включает от одного до 50











CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG- CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG	
CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG- CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG- CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG- CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG-CAG- CAG-CAG	197
AGC	H/Π
AGCAGC	H/Π
AGCAGCAGC	H/Π
AGCAGCAGCAGC	198
AGCAGCAGCAGCAGC	199
AGCAGCAGCAGCAGCAGC	200
AGCAGCAGCAGCAGCAGCAGC	201
AGCAGCAGCAGCAGCAGCAGCAGC	202
AGCAGCAGCAGCAGCAGCAGCAGCAGC	203
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	204
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	205
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	206
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	207
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	208
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	209
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGC	210
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGC	211
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGC	212
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGC	213
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGCAGC	214
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGCAGCAGC	215













GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	286
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCA	287
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCA	288
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCA	289
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCA	290
GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCA	291

Согласно вариантам реализации АС, который имеет нуклеотид, который включает повторы СAG, может включать дополнительные нуклеотидные последовательности повтора СAG на 5'-конце, 3'-конце или на обоих концах. Согласно вариантам реализации дополнительные нуклеотидные последовательности могут иметь от 80% до 100% или от 95% до 100% комплементарности с частями целевого транскрипта, с которым они гибридизуются. Дополнительные нуклеотидные последовательности можно добавлять к нуклеотидной последовательности повтора СAG для повышения селективности гибридизации АС с конкретным целевым транскриптом.

Согласно вариантам реализации АС включает нуклеотидную последовательность, которая включает от 1 до 50 повторов СAG и которая представляет собой гэлмер. Гэлмеры являются олигонуклеотидами, которые представляют собой гибрид ДНК/РНК и

индуцируют механизм распада под действием РНКазы. Например, гэдмеры могут иметь центральный сегмент ДНК или миметика ДНК, фланкированный сегментом РНК или миметика РНК, как на 5'-конце, так и на 3'-конце сегмента ДНК или миметика ДНК. Согласно вариантам реализации АС включает гэдмер, который включает нуклеотидную последовательность, гибридизирующуюся с целевой последовательностью нуклеиновой кислоты целевого транскрипта, которая отделена от экспансированного повтора CUG целевого транскрипта.

Согласно вариантам реализации АС согласно настоящему изобретению представляет собой гэдмерный олигонуклеотид, раскрытый в патенте США № 9550988, содержание которого включено в настоящий документ посредством ссылки.

Согласно вариантам реализации АС согласно настоящему изобретению содержит последовательность и/или структуру любого из АС, нацеленных на DMPK, раскрытых в публикации патента США № 2017/0260524, содержание которой включено в настоящий документ посредством ссылки.

Согласно вариантам реализации АС согласно настоящему изобретению содержит последовательность и/или структуру любого из АС или олигонуклеотидов, раскрытых в публикациях патентов США 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 и US20120059042A1, содержание каждого из которых полностью включено в настоящий документ для всех целей.

При использовании АС для нацеливания и/или гибридизации с экспансированным повтором CTG·CUG следует проявлять осторожность, чтобы избежать нецелевых эффектов, когда АС непреднамеренно связывается с нецелевыми транскриптами, которые включают повторы CTG·CUG (например, транскрипт, который включает повторы CTG·CUG, которые не являются экспансированным повтором CTG·CUG). Анализ генома человека *in silico* выявил, что в общей сложности 63 гена человека имеют повторы CTG·CUG (Uhlen, et. al., *Science* 2015 347(6220):1260419)). 63 гена можно ранжировать по экспрессии мРНК плюс количество белка, экспрессируемого в общей мышечной ткани (сердечной, скелетной и гладкой мышцах). Уровень экспрессии может быть количественно определен с использованием RPM (считывания на миллион), экспрессия мРНК представляет собой FPKM (фрагменты на тысячу пар оснований транскрипта на миллион картированных фрагментов), и экспрессия белка представляет собой рTPM (транскрипты на миллион генов, кодирующих белок), используя более 10 RPM в качестве значения отсечки для значимой экспрессии. На **ФИГ. 43** показаны результаты такого анализа *in silico*. Тридцать шесть генов имеют уровень экспрессии >10 RPM. Из 36 генов только три гена (помимо DMPK) имели >10 повторов CTG·CUG. Гены с ≤10 повторами CTG·CUG представляют самый низкий риск нецелевого связывания и токсичности. Число повторов CTG·CUG (11-24) в этих 3 генах (TCF4, CASK, MAP3k4), тем не менее, значительно ниже, чем у пациентов с классической и врожденной DM1. Например, у

пациентов, страдающих DM1 с поздним началом, имеется 100-600 повторов CTG·CUG в DMPK, у пациентов, страдающих классической DM1, имеется 250-750 повторов CTG·CUG в DMPK, и у пациентов, страдающих врожденной DM1, имеется 750-1400 повторов CTG·CUG в DMPK. Тот же анализ *in silico* можно выполнить для печени и почек. CASK является единственным значимым геном с >10 повторами CTG·CUG в почке. Ни один ген с >10 повторами CTG·CUG не был значимым в печени.

АС, описанные в настоящем документе, могут содержать один или более асимметричных центров и, таким образом, образовывать энантиомеры, диастереомеры и другие стереоизомерные конфигурации, которые могут быть определены с точки зрения абсолютной стереохимии как (R) или (S);  $\alpha$  или  $\beta$ ; или как (D) или (L). В антисмысловые соединения, предложенные в настоящем документе, включены все такие возможные изомеры, а также их рацемические и оптически чистые формы.

Эффективность АС можно оценить путем исследования антисмысловой активности, которая осуществляется при их введении. В настоящем документе термин «антисмысловая активность» относится к любой детектируемой и/или измеримой активности, относимой к гибридизации антисмыслового соединения с его целевой нуклеотидной последовательностью. Такое детектирование и/или измерение может быть прямым или косвенным. Согласно вариантам реализации антисмысловую активность оценивают путем детектирования и/или измерения количества белка, экспрессируемого из представляющего интерес транскрипта. Согласно вариантам реализации антисмысловую активность оценивают путем детектирования и/или измерения количества представляющего интерес транскрипта. Согласно вариантам реализации антисмысловую активность оценивают путем детектирования и/или измерения количества альтернативно сплайсированной РНК и/или количества белковых изоформ, транслируемых из целевого транскрипта.

#### *Механизмы модуляции АС*

Согласно вариантам реализации АС может модулировать активность и/или уровень целевого транскрипта внутри клетки. На **ФИГ. 1** показаны примерные механизмы того, как АС может модулировать уровень и/или активность целевого транскрипта.

Согласно вариантам реализации АС может модулировать уровень целевого транскрипта внутри клетки. Например, в вариантах реализации, в которых АС представляет собой гэпмер, связывание АС с целевым транскриптом индуцирует разрушение целевого транскрипта за счет путей РНКазы Н (**ФИГ. 1**, стрелки А и В). Согласно вариантам реализации гэпмер гибридизуется с частью целевого транскрипта, которая отличается от экспансированного тринуклеотидного повтора CUG, и тем самым индуцирует разрушение целевого транскрипта за счет путей РНКазы Н (**ФИГ. 1**, стрелка А). Согласно вариантам реализации гэпмер гибридизуется по меньшей мере с частью экспансированного повтора CUG в пределах целевого транскрипта и тем самым индуцирует разрушение целевого транскрипта за счет путей РНКазы Н (**ФИГ. 1**, стрелка В).

Согласно вариантам реализации АС может модулировать активность целевого транскрипта. Модуляция активности может включать увеличение или уменьшение способности целевого транскрипта связываться с партнером по связыванию. Согласно вариантам реализации АС может модулировать активность целевого транскрипта путем уменьшения способности целевого транскрипта связываться с одним или более белками, которые могут ассоциироваться с транскриптом, в частности, белками, которые ассоциируются по меньшей мере с частью экспансированного повтора CUG целевого транскрипта (ФИГ. 1, стрелка С). Согласно вариантам реализации уменьшение способности целевого транскрипта связываться с одним или более белками включает уменьшение аффинности целевого транскрипта в отношении одного или более белков. Согласно вариантам реализации уменьшение способности целевого транскрипта связываться с одним или более белками включает частичное или полное стерическое блокирование связывания целевого транскрипта с одним или более белками. Например, АС может занимать по меньшей мере часть сайта связывания, которая может быть занята одним или более белками, если она не заблокирована стерически. Согласно вариантам реализации сайт связывания одного или более белков с целевым транскриптом включает по меньшей мере часть экспансированного тринуклеотидного повтора целевого транскрипта. Таким образом, в вариантах реализации АС может занимать по меньшей мере часть экспансированного тринуклеотидного повтора (например, экспансированного повтора CUG), которая может быть занята одним или более белками, если она не заблокирована стерически. Частичное стерическое блокирование целевого транскрипта может привести к уменьшению аффинности между целевым транскриптом и одним или более белками. Например, в вариантах реализации АС связывается по меньшей мере с частью экспансированного повтора CUG целевого транскрипта, что стерически блокирует и/или уменьшает аффинность целевого транскрипта в отношении белка, который может связываться с экспансированным повтором CUG (ФИГ. 1, стрелка С). В следующей обзорной статье описаны дополнительные приложения для стерически блокирующих антисмысловых олигонуклеотидов, и она полностью включена в настоящий документ посредством ссылки: Roberts et al. *Nature Reviews Drug Discovery* (2020) 19: 673-694.

Повторы CUG экспансированного повтора CUG могут образовывать двухцепочечную шпилечную структуру. При патологическом состоянии белки связываются с двухцепочечной шпилечной структурой и становятся секвестрированными и неспособными выполнять другие функции. Согласно вариантам реализации АС связывается по меньшей мере с частью двухцепочечной шпилечной структуры, что стерически блокирует и/или уменьшает аффинность двухцепочечной шпилечной структуры в отношении белкового партнера по связыванию. Согласно вариантам реализации АС связывается по меньшей мере с частью одноцепочечного экспансированного повтора CUG, что ингибирует образование двухцепочечной шпилечной структуры и, таким образом, ингибирует связывание одного или более белков с двухцепочечной шпилечной структурой. Согласно вариантам реализации гибридизация

АС с двойной шпилечной структурой стерически блокирует и/или уменьшает аффинность связывания одного или более белков с двойной шпилечной структурой. Согласно вариантам реализации гибридизация АС по меньшей мере с частью одноцепочечной области экспансированного тринуклеотидного повтора ингибирует образование двухцепочечной шпилечной структуры.

Уменьшение способности целевого транскрипта связываться с одним или более белками может позволить одному или более белкам выполнять другие функции, такие как, например, регуляция сплайсинга нисходящих транскриптов (транскриптов, которые не содержат экспансированного повтора CUG). Таким образом, в вариантах реализации уменьшение способности целевого транскрипта связываться с одним или более белками может повышать уровень одного или более белков в клетке, которые доступны для обеспечения других функций или функции на других транскриптах. Согласно вариантам реализации уменьшение способности целевого транскрипта связываться с одним или более белками может повышать цитозольный уровень одного или более белков в клетке, которые доступны для обеспечения других функций или функции на других транскриптах. Таким образом, в вариантах реализации связывание АС с целевым транскриптом может привести к модуляции уровня и/или активности одного или более белков, которые взаимодействуют с целевым транскриптом.

Согласно вариантам реализации гибридизация АС по меньшей мере с частью экспансированного повтора CUG уменьшает аффинность и/или стерически блокирует связывание MNBL1 с целевым транскриптом. MNBL1 представляет собой фактор сплайсинга, который регулирует сплайсинг транскриптов нисходящих генов. При патологическом фенотипе DM1 MNBL1 связывается с экспансированным повтором CUG целевого транскрипта. Когда MNBL1 связан с целевым транскриптом, он секвестрируется в ядре и не может регулировать сплайсинг транскриптов нисходящих генов (транскриптов, которые не содержат экспансированного повтора CUG). Согласно вариантам реализации гибридизация АС по меньшей мере с частью экспансированного повтора CUG в целевом транскрипте стерически блокирует и/или уменьшает аффинность MNBL1 в отношении целевого транскрипта, что позволяет ему регулировать сплайсинг транскриптов нисходящих генов. Согласно вариантам реализации гибридизация АС по меньшей мере с частью экспансированного повтора CUG в целевом транскрипте стерически блокирует и/или уменьшает аффинность MNBL1 в отношении целевого транскрипта, что увеличивает количество свободного (например, не связанного с транскриптом, содержащим повтор CUG) MNBL1. Согласно вариантам реализации гибридизация АС по меньшей мере с частью экспансированного повтора CUG в целевом транскрипте стерически блокирует и/или уменьшает аффинность MNBL1 в отношении целевого транскрипта, что уменьшает количество MBNL1, связанного с целевым транскриптом и секвестрированного им.

Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация АС по меньшей мере с частью экспансированного повтора

CUG уменьшает аффинность и/или стерически блокирует связывание MNBL1 с целевым транскриптом. MNBL1 представляет собой фактор сплайсинга, который регулирует сплайсинг транскриптов нисходящих генов. При патологическом фенотипе DM1 MNBL1 связывается с экспансированным повтором CUG DMPK1. Когда MNBL1 связан с DMPK1, он секвестрируется в ядре и не может регулировать сплайсинг транскриптов нисходящих генов (транскриптов, которые не содержат экспансированного повтора CUG). Согласно вариантам реализации гибридизация AC по меньшей мере с частью экспансированного повтора CUG в DMPK стерически блокирует и/или уменьшает аффинность MNBL1 в отношении транскрипта DMPK, что позволяет ему регулировать сплайсинг транскриптов нисходящих генов. Согласно вариантам реализации гибридизация AC по меньшей мере с частью экспансированного повтора CUG в DMPK стерически блокирует и/или уменьшает аффинность MNBL1 в отношении транскрипта DMPK, что увеличивает количество свободного (например, не связанного с транскриптом, содержащим повтор CUG) MNBL1. Согласно вариантам реализации гибридизация AC по меньшей мере с частью экспансированного повтора CUG в DMPK стерически блокирует и/или уменьшает аффинность MNBL1 в отношении транскрипта DMPK, что уменьшает количество MNBL1, связанного с транскриптом DMPK1 и секвестрированного им.

Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация AC по меньшей мере с частью экспансированного повтора CUG приводит к уменьшению уровней CUGBP1. При патологическом состоянии DM1 уровень свободного MBNL1 (способного функционировать) уменьшается, тогда как уровень свободного (способного функционировать) CUGBP1 повышается. Повышение уровней CUGBP1 ассоциировано с патологическим состоянием. Таким образом, в вариантах реализации гибридизация AC по меньшей мере с частью экспансированного повтора CUG приводит к повышению уровней свободного (способного функционировать) MBNL1 и/или уменьшению уровней свободного (способного функционировать) CUGBP1.

Уменьшение способности целевого транскрипта связываться с одним или более белками может снижать или ингибировать образование очагов повторов CUG. Транскрипты, которые включают экспансированные нуклеотидные повторы (например, экспансированные повторы CUG), могут быть транскрибированы и затем секвестрированы в ядре. Внутри ядра секвестрированные транскрипты могут образовывать агрегаты. Белки, которые связываются с транскриптом, затем могут нуклеироваться на секвестрированном транскрипте и/или секвестрировать агрегат транскрипта с образованием очагов экспансированного нуклеотидного повтора (например, повтора CUG). Очаги повторов CUG могут быть видны с помощью микроскопии. Согласно вариантам реализации уменьшение способности целевого транскрипта связываться с одним или более белками может снижать или ингибировать образование агрегатов, которые включают целевой транскрипт. Согласно вариантам реализации уменьшение способности целевого транскрипта связываться с одним или более белками может снижать или ингибировать нуклеацию одного или более белков на целевом



транскрипте, на двухцепочечной шпилечной области транскрипта или на агрегате целевых транскриптов. Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, уменьшение способности целевого транскрипта DMPK1 связываться с MNBL1 может снижать или ингибировать нуклеацию MNBL1 на целевом транскрипте DMPK1 или агрегате целевого транскрипта DMPK1. Согласно вариантам реализации гибридизация AC с целевым транскриптом может привести к ингибированию или уменьшению образования ядерных очагов повторов CUG. Согласно вариантам реализации гибридизация AC с целевым транскриптом может привести к ингибированию или уменьшению образования ядерных очагов повторов CUG, образованных из целевого транскрипта DMPK, TCF4, JPH3 и/или ATXN8OS/ATXN8.

Согласно вариантам реализации гибридизация AC с целевым транскриптом может привести к модуляции уровня, экспрессии и/или активности одного или более нисходящих генов. Например, гибридизацию AC с целевым транскриптом можно применять для индукции разрушения целевого транскрипта или стерической блокировки или уменьшения аффинности целевого транскрипта в отношении одного или более белков, что позволяет одному или более белкам, которые были секвестрированы целевым транскриптом, регулировать экспрессию, уровень и/или активность нисходящих генов. Например, в вариантах реализации один или более белков могут включать белок, который участвует в регуляции сплайсинга одного или более нисходящих транскриптов (транскриптов, которые не содержат экспансированного повтора CUG). Согласно вариантам реализации сплайсинг нисходящих транскриптов изменяется, когда белок, участвующий в сплайсинге, связан и секвестрирован на целевом транскрипте. Например, изменение сплайсинга может включать исключение одного или более экзонов или включение одного или более интронов в транскрипт, что приводит к экспрессии различных изоформ белка. Согласно вариантам реализации изменение сплайсинга может привести к включению экзона и/или интрона, который включает преждевременный стоп-кодон, что приводит к получению усеченной изоформы, которая может не обладать активностью или может иметь вредную активность. Изменение сплайсинга транскрипта нисходящего гена может привести к изменению уровня, укладки и/или активности продукта нисходящего гена, который может быть ассоциирован с патологическим фенотипом. При отсутствии связывания с целевым транскриптом, содержащим экспансированный повтор CUG, белок, вовлеченный в сплайсинг, может регулировать сплайсинг, что может привести к коррекции (или восстановлению) сплайсинга транскрипта нисходящего гена, тем самым по меньшей мере частично восстанавливая уровень, укладку и/или активность белкового продукта нисходящего гена, ассоциированного со здоровым фенотипом.

Согласно вариантам реализации гибридизация AC с целевым транскриптом может привести к модуляции сплайсинга транскриптов нисходящих генов, которые регулируются белками, которые секвестрированы целевым транскриптом при патологическом состоянии, ассоциированном с экспансированными нуклеотидными

повторами (например, экспансированными тринуклеотидными повторами). При заболевании, ассоциированном с экспансированными тринуклеотидными повторами, транскрипты нисходящих генов часто подвергаются ошибочному процессингу, например, ошибочному сплайсингу. Ошибочный сплайсинг транскриптов нисходящих генов может привести к разрушению генных продуктов до трансляции или трансляции в белки, которые имеют aberrantную структуру и/или функцию. Например, секвестрация белков, которые регулируют процессинг транскриптов нисходящих генов, может привести к включению экзонов и/или интронов с преждевременными стоп-кодонами, включению интронов, исключению экзонов и/или включению альтернативных экзонов, что может привести к разрушению транскрипта и/или генного продукта до трансляции или трансляции в генный продукт с aberrantной функцией. Изменение уровней транскрипта нисходящего гена и/или генного продукта, и/или aberrantной структуры, и/или функции продуктов нисходящих генов ассоциировано с патологическими фенотипами, характеризующимися экспансированными тринуклеотидными повторами. Согласно вариантам реализации гибридизация АС с целевым транскриптом может привести к модуляции включения экзона, исключения экзона, включения интрона и/или исключения интрона в нисходящих транскриптах, сплайсинг которых регулируется белками, которые секвестрируются на целевом транскрипте при патологическом состоянии. Таким образом, гибридизация АС с целевым транскриптом может привести к положительной регуляции изомеров нисходящих белков и/или транскриптов, ассоциированных со здоровым фенотипом. Аналогичным образом, гибридизация АС с целевым транскриптом может привести к отрицательной регуляции (например, подавлению) нисходящих транскриптов и/или изомеров белков, ассоциированных с патологическим фенотипом.

Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация АС с целевым транскриптом может привести к модуляции сплайсинга транскриптов нисходящих генов, которые регулируются белками, которые секвестрированы целевым транскриптом DMPK при патологическом состоянии. При DM1 несколько транскриптов нисходящих генов являются ошибочно сплайсированными, что ведет. Ошибочно сплайсированные гены ассоциированы с патологическим фенотипом. Таким образом, модуляция сплайсинга генов может включать коррекцию (например, восстановление) сплайсинга генов с получением генных продуктов нисходящих генов, которые ассоциированы со здоровым фенотипом. Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация АС с целевым транскриптом может привести к модуляции сплайсинга транскриптов нисходящих генов, которые регулируются MNBL1, регулятором сплайсинга, который секвестрируется целевым транскриптом DMPK при патологическом состоянии. Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация АС с целевым транскриптом может привести к модуляции сплайсинга транскриптов нисходящих генов, которые регулируются CUGBP1, белком, на активность которого влияют экспансированные повторы CUG. Согласно вариантам реализации, в которых

целевой транскрипт представляет собой DMPK, гибридизация АС с целевым транскриптом может привести к правильному процессингу (например, сплайсингу) нисходящих генов, которые регулируются MNBL1 и/или CUGBP1. Согласно вариантам реализации, в которых целевой транскрипт представляет собой DMPK, гибридизация АС с целевым транскриптом может привести к модуляции сплайсинга нисходящих генов, включая, но не ограничиваясь перечисленными, 4833439L19Rik, Abcc9, Atp2a1, Arhgef10, Arhgap28, Armcx6, Angel1, Best3, Bin1, Brd2, Cacna1s, Cacna2d1, Cpd, Creb3, Ccpg1, Clasp1, CIC-1, Clcn1, Clk4, Creb2, Camk2g, Capzb, Copz2, Coch, cTNT, Ctu2, Cyp2s1, Dctn4, Dnm11, Eya4, Efna3, Efna2, Fbxo31, Fbxo21, Frem2, Fgd4, Fuca1, Fn1, Gogla4, Gpr3711, 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, 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, Vsp39 или любую их комбинацию.

Ошибочный сплайсинг многих вышеупомянутых транскриптов нисходящих генов приводит к определенным патологическим фенотипам DM1. Например, MNBL1 является фактором сплайсинга с потерей функции при DM1 вследствие включения экзона 5. MNBL1 секвестрируется экспансией CUG DMPK и образует ядерные очаги РНК. Кроме того, SOS1 способствует активации Ras для положительной регуляции сигнального пути RAS/MAPK. При DM1 экзон 25 SOS1 исключен, что приводит к ингибированию путей мышечной гипертрофии. При DM1 IR/INSR характеризуется исключением экзона 11, что приводит к более высоким уровням изоформы, отличной от мышечной, с низким уровнем передачи сигнала, и уменьшению метаболического ответа на инсулин при DM1 (резистентность к инсулину). Аналогичным образом, при DM1 наблюдается исключение экзона 78 DMD. Исключение экзона 78 приводит к выходу транскрипта за пределы рамки считывания в С-концевом домене. Этот мутантный белок экспрессируется у пациентов, страдающих DM1, и ассоциирован с механизмом, ответственным за истощение мышц у пациентов. BIN1 необходим для правильного образования мышечных Т-трубочек (ЕС-связывание). Исключение экзона 11 приводит к образованию неактивной изоформы и обнаруживается у пациентов, страдающих DM1. LDB3 взаимодействует с  $\alpha$ -актином в Z-диске в поперечнополосатой мышце и сохраняет мышечную структуру. Включение экзона 11 LDB3, детектированное при DM1, приводит к уменьшению аффинности в отношении протеинкиназы С (PKC). Следовательно, PKC становится гиперактивной при DM1. Согласно вариантам реализации модуляция одного или более нисходящих генов приводит к коррекции или восстановлению сплайсинга транскриптов, который ассоциирован со здоровым фенотипом. Таким образом, в вариантах реализации гибридизация АС с целевым транскриптом DMPK приводит к восстановлению ошибочного сплайсинга нисходящих генов/транскриптов, что снижает уровень нисходящих генов/транскриптов, ассоциированных с патологическим фенотипом. Таким образом, в вариантах реализации

гибридизация АС с целевым транскриптом DMPK приводит к восстановлению ошибочного сплайсинга нисходящих генов/транскриптов, что повышает уровень нисходящих генов/транскриптов, ассоциированных со здоровым фенотипом.

Согласно вариантам реализации АС ингибирует экспрессию целевого транскрипта. Согласно вариантам реализации АС ингибирует экспрессию целевого транскрипта путем блокирования доступа механизма процессинга пре-мРНК и/или механизма трансляции и/или его способности завершать трансляцию и/или процессинг пре-мРНК. Согласно вариантам реализации АС ингибирует экспрессию целевого транскрипта путем индукции разрушения целевого транскрипта, например, посредством путей РНКазы Н.

### **Структура АС**

АС включает олигонуклеотид и/или олигонуклеозид. Олигонуклеотиды и/или олигонуклеозиды представляют собой нуклеотиды или нуклеозиды, связанные межнуклеозидными связями. Нуклеозиды включают пентозный сахар (например, рибозу или дезоксирибозу) и азотистое основание, ковалентно присоединенное к сахару. Встречающиеся в природе (или традиционные основания) основания, обнаруживаемые в ДНК и/или РНК, представляют собой аденин (А), гуанин (G), тимин (Т), цитозин (С) и урацил (U). Встречающиеся в природе сахара (или традиционные сахара), обнаруживаемые в ДНК и/или РНК, представляют собой дезоксирибозу (ДНК) и рибозу (РНК). Встречающаяся в природе нуклеозидная связь (или традиционная межнуклеозидная связь) представляет собой фосфодиэфирную связь. Согласно вариантам реализации АС согласно настоящему изобретению могут иметь все природные сахара, основания и межнуклеозидные связи.

Химически модифицированные нуклеозиды рутинно применяют для включения в антисмысловые соединения, чтобы улучшить одно или более свойств, таких как устойчивость к нуклеазам, фармакокинетика или аффинность в отношении целевой РНК. Согласно вариантам реализации АС согласно настоящему изобретению могут содержать один или более модифицированных нуклеозидов. Согласно вариантам реализации АС согласно настоящему изобретению могут содержать один или более модифицированных сахаров. Согласно вариантам реализации АС согласно настоящему изобретению могут содержать одно или более модифицированных оснований. Согласно вариантам реализации АС согласно настоящему изобретению могут содержать одну или более модифицированных межнуклеозидных связей.

В целом, нуклеиновое основание представляет собой любую группу, которая содержит один или более атомов или групп атомов, способных образовывать водородные связи с основанием другой нуклеиновой кислоты. В дополнение к «немодифицированным» или «природным» нуклеиновым основаниям (А, G, Т, С и U) специалистам в данной области техники известны многие модифицированные нуклеиновые основания или миметики нуклеиновых оснований, которые пригодны для соединений, описанных в настоящем документе. Обычно модифицированное нуклеиновое основание относится к нуклеиновому основанию, которое является достаточно сходным

по структуре с исходным нуклеиновым основанием, такому как, например, 7-диазапурин, 5-метилцитозин, 2-тио-dT (**ФИГ. 2**) или G-фиксирующее основание. Обычно миметик нуклеинового основания представляет собой нуклеиновое основание, которое включает структуру, которая является более сложной, чем модифицированное нуклеиновое основание, такое как, например, трициклический миметик нуклеинового основания на основе феноксазина. Способы получения вышеупомянутых модифицированных нуклеиновых оснований хорошо известны специалистам в данной области техники.

Согласно вариантам реализации АС может включать один или более нуклеозидов, имеющих модифицированный фрагмент сахара. Согласно вариантам реализации фуранозильный сахар природного нуклеозиды может иметь 2'-модификацию, модификации для получения ограниченного нуклеозиды и другие (см. **ФИГ. 2**). Например, в вариантах реализации фуранозильное сахарное кольцо природного нуклеозиды может быть модифицировано рядом способов, включая, но не ограничиваясь перечисленными, добавление группы заместителя, образование мостика между двумя негеминальными атомами кольца с образованием бициклической нуклеиновой кислоты (BNA) или заблокированной нуклеиновой кислоты; обмен кислорода фуранозильного кольца на С или N; и/или замену атома или группы таким образом (см. **ФИГ. 2**). Модифицированные сахара хорошо известны и могут применяться для повышения или уменьшения аффинности АС в отношении его целевой нуклеотидной последовательности. Модифицированные сахара также можно применять для повышения устойчивости АС к нуклеазам. Сахара также могут быть заменены группами, имитирующими сахара, помимо прочего. Согласно вариантам реализации один или более сахаров нуклеозидов АС заменены метиленморфолиновым кольцом, показанным как 19 на **ФИГ. 2**.

Согласно вариантам реализации АС включает один или более нуклеозидов, которые включают бициклический модифицированный сахар (BNA; иногда называемые мостиковыми нуклеиновыми кислотами). Примеры BNA включают, но не ограничиваются перечисленными, LNA (мостик 4'-(CH<sub>2</sub>)-O-2'), 2'-тио-LNA (мостик 4'-(CH<sub>2</sub>)-S-2'), 2'-амино-LNA (мостик 4'-(CH<sub>2</sub>)-NR-2'), ENA (мостик 4'-(CH<sub>2</sub>)<sub>2</sub>-O-2'), 4'-(CH<sub>2</sub>)<sub>3</sub>-2'-мостиковую BNA, 4'-(CH<sub>2</sub>CH(CH<sub>3</sub>))-2'-мостиковую BNA" cEt (мостик 4'-(CH(CH<sub>3</sub>))-O-2') и cMOE BNA (мостик 4'-(CH(CH<sub>2</sub>OCH<sub>3</sub>))-O-2'). BNA получены и раскрыты в патентной литературе, а также в научной литературе (Srivastava, et al. J. Am. Chem. Soc. (2007), ACS Advanced online publication, 10.1021/ja071106y; Albaek et al. J. Org. Chem. (2006), 71, 7731 - 7740; Fluiter, et al. Chembiochem (2005), 6, 1104-1109; Singh et al., Chem. Commun. (1998), 4, 455-456; Koshkin et al., Tetrahedron (1998), 54, 3607-3630; Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A. (2000), 97, 5633-5638; Kumar et al., Bioorg. Med. Chem. Lett. (1998), 8, 2219-2222; WO 94/14226; WO 2005/021570; Singh et al., J. Org. Chem. (1998), 63, 10035-10039, WO 2007/090071; патенты США №№ 7053207; 6268490; 6770748; 6794499; 7034133; и 6525191; и заявка на патент США, опубликованная через 18 месяцев с даты приоритета 2004-0171570; 2004-0219565; 2004-0014959; 2003-0207841; 2004-0143114; и 20030082807).

Согласно вариантам реализации АС включает один или более нуклеозидов, которые включают блокированную нуклеиновую кислоту (LNA). В LNA 2'-гидроксильная группа рибозильного сахарного кольца связана с 4'-атомом углерода сахарного кольца, образуя тем самым 2'-С,4'-С-оксиметиленовую связь, с образованием бициклического сахарного фрагмента (рассмотрено в Elayadi et al., *Curr. Opinion Invens. Drugs* (2001), 2, 558-561; Braasch et al., *Chem. Biol.* (2001), 8 1-7; и Orum et al., *Curr. Opinion Mol. Ther.* (2001), 3, 239-243; см. также патенты США: 6268490 и 6670461). Связь может представлять собой метиленовую (-CH<sub>2</sub>-) группу, образующую мостик между 2'-атомом кислорода и 4'-атомом углерода, в этом случае термин LNA используется для бициклического фрагмента; в случае этиленовой группы в этом положении используется термин ENA<sup>TM</sup> (Singh et al., *Chem. Commun.* (1998), 4, 455-456; ENA<sup>TM</sup>; Morita et al., *Bioorganic Medicinal Chemistry* (2003), 11, 2211-2226). LNA и другие бициклические аналоги сахаров проявляют очень высокую термическую стабильность дуплекса с комплементарными ДНК и РНК (T<sub>m</sub>=от +3 до +10°C), стабильность к 3'-экзонуклеолитическому разрушению и хорошую растворимость. Описаны сильные и нетоксичные антисмысловые олигонуклеотиды, содержащие LNA (Wahlestedt et al., *Proc. Natl. Acad. Sci. U.S.A.* (2000), 97, 5633-5638).

Изомер LNA, который также исследован, представляет собой альфа-L-LNA, который, как было показано, обладает превосходной стабильностью против 3'-экзонуклеазы. Альфа-L-LNA были включены в антисмысловые гэтмеры и химеры, которые показали сильную антисмысловую активность (Frieden et al., *Nucleic Acids Research* (2003), 21, 6365-6372).

Описан синтез и получение мономеров LNA аденина, цитозина, гуанина, 5-метилцитозина, тимина и урацила, а также их олигомеризация и свойства распознавания нуклеиновых кислот (Koshkin et al., *Tetrahedron*, 1998, 54, 3607-3630). LNA и их получение также описано в WO 98/39352 и WO 99/14226.

Также получены аналоги LNA, фосфотиоат-LNA и 2'-тио-LNA (Kumar et al., *Bioorg. Med. Chem. Lett.*, 1998, 8, 2219-2222). Также описано получение аналогов LNA, содержащих дуплексы олигодезоксирибонуклеотидов в качестве субстратов для полимераз нуклеиновых кислот (Wengel et al., WO 99/14226). Кроме того, описан синтез 2'-амин-LNA, конформационно ограниченного высокоаффинного аналога олигонуклеотида (Singh et al., *J. Org. Chem.* (1998), 63, 10035-10039). Кроме того, получены 2'-амин- и 2'-метиламин-LNA, и ранее сообщалось о термической стабильности их дуплексов с комплементарными цепями РНК и ДНК.

Способы получения модифицированных сахаров хорошо известны специалистам в данной области техники. Некоторые типичные патенты и публикации, в которых описано получение таких модифицированных сахаров, включают, но не ограничиваются перечисленными, патенты США: 4981957; 5118800; 5319080; 5359044; 5393878; 5446137; 5466786; 5514785; 5519134; 5567811; 5576427; 5591722; 5597909; 5610300; 5627053; 5639873; 5646265; 5658873; 5670633; 5792747; 5700920; и 6600032; и WO 2005/121371.

### *Межнуклеозидные связи*

В настоящем документе описаны межнуклеозидные линкерные группы, которые связывают нуклеозиды или иным образом модифицированные нуклеозидные мономерные звенья друг с другом с образованием олигонуклеотида и/или олигонуклеотида, содержащего АС. АС могут включать встречающиеся в природе межнуклеозидные связи, не встречающиеся в природе межнуклеозидные связи или и то, и другое.

Во встречающихся в природе ДНК и РНК межнуклеозидная линкерная группа представляет собой фосфодиэфир, который ковалентно связывает соседние нуклеозиды друг с другом с образованием линейного полимерного соединения. Во встречающихся в природе ДНК и РНК фосфодиэфир связан с 2'-, 3'- или 5'-гидроксильным фрагментом сахара. В олигонуклеотидах фосфатные группы обычно называют «образующими межнуклеозидный остов олигонуклеотида». Во встречающихся в природе ДНК и РНК связь или остов РНК и ДНК представляет собой фосфодиэфирную связь в направлении от 3' к 5'. Согласно вариантам реализации межнуклеозидные линкерные группы АС представляют собой фосфодиэфиры. Согласно вариантам реализации межнуклеозидные линкерные группы АС представляют собой 3'-5'-фосфодиэфирные связи.

Два основных класса неприродных межнуклеозидных линкерных групп определяются наличием или отсутствием атома фосфора. Типичные фосфоросодержащие межнуклеозидные связи, включают, но не ограничиваются перечисленными, фосфортриэфиры, метилфосфонаты, фосфорамидат и фосфотиоаты. Типичные не содержащие фосфора межнуклеозидные линкерные группы включают, но не ограничиваются перечисленными, метиленметилямино (-CH<sub>2</sub>-N(CH<sub>3</sub>)-O-CH<sub>2</sub>-), тиодиэфир (-O-C(O)-S-), тионокарбамат (-O-C(O)(NH)-S-); силоксан (-O-Si(H<sub>2</sub>-O-) и N,N'-диметилгидразин (-CH<sub>2</sub>-N(CH<sub>3</sub>)-N(CH<sub>3</sub>)-). АС, имеющие фосфорные межнуклеозидные линкерные группы, называют олигонуклеотидами. Антисмысловые соединения, имеющие не содержащие фосфора межнуклеозидные линкерные группы, называются олигонуклеотидами. Модифицированные межнуклеозидные связи, по сравнению с природными фосфодиэфирными связями, можно применять для изменения, как правило, повышения, устойчивости антисмыслового соединения к нуклеазам. Межнуклеозидные связи, содержащие хиральный атом, могут быть получены в виде рацемических, хиральных или в виде смеси. Типичные хиральные межнуклеозидные связи включают, но не ограничиваются перечисленными, алкилфосфонаты и фосфотиоаты. Способы получения фосфоросодержащих и не содержащих фосфора связей хорошо известны специалистам в данной области техники.

Согласно вариантам реализации два или более нуклеозидов, содержащих модифицированные сахара и/или модифицированные нуклеиновые основания, могут быть соединены с использованием фосфорамидатов. Согласно вариантам реализации два или более нуклеозидов, содержащих метиленморфолиновое кольцо, могут быть соединены фосфорамидатной межнуклеозидной связью.

Антисмысловые соединения, которые включают нуклеиновые основания с метиленморфолиновым кольцом, которые связаны фосфорамидатной межнуклеозидной связью, могут называться фосфорамидат-морфолиновыми олигомерами (РМО).

#### *Конъюгатные группы*

Согласно вариантам реализации АС модифицированы путем ковалентного присоединения одной или более конъюгатных групп. В целом, конъюгатные группы модифицируют одно или более свойств присоединенного АС, включая, но не ограничиваясь перечисленными, фармакодинамику, фармакокинетику, связывание, всасывание, клеточное распределение, клеточное поглощение, заряд и клиренс. Конъюгатные группы рутинно используют в области химии и связывают непосредственно или посредством необязательного линкерного фрагмента или линкерной группы с исходным соединением, таким как АС. Конъюгатные группы включают, но не ограничиваются перечисленными, интеркаляторы, репортерные молекулы, полиамины, полиамиды, полиэтиленгликоли, тиоэфиры, полиэфиры, холестерин, тиохолестерин, фрагменты холевой кислоты, фолат, липиды, фосфолипиды, биотин, феназин, фенантридин, антрахинон, адамантан, акридин, флуоресцеины, родамины, кумарины и красители. Согласно вариантам реализации конъюгатная группа представляет собой полиэтиленгликоль (PEG), и PEG конъюгирован либо с АС, либо с CPP (CPP обсуждается в другом месте настоящего документа).

Согласно вариантам реализации конъюгатные группы включают липидные фрагменты, такие как холестеринный фрагмент (Letsinger et al., Proc. Natl. Acad. Sci. USA (1989), 86, 6553); холевую кислоту (Manoharan et al., Bioorg. Med. Chem. Lett. (1994), 4, 1053); простой тиоэфир, например, гексил-S-тримитилтиол (Manoharan et al., Ann. N.Y. Acad. Sci. (1992), 660, 306; Manoharan et al., Bioorg. Med. Chem. Lett. (1993), 3, 2765); тиохолестерин (Oberhauser et al., Nucl. Acids Res. (1992), 20, 533); алифатическую цепь, например, додекандиоловый или ундецильный остатки (Saison-Behmoaras et al., EMBO J. (1991), 10, 111; Kabanov et al., FEBS Lett. (1990), 259, 327; Svinarchuk et al., Biochimie (1993), 75, 49); фосфолипид, например, дигексадецил-рац-глицерин или триэтиламмоний 1,2-ди-О-гексадецил-рац-глицеро-3-Н-фосфонат (Manoharan et al., Tetrahedron Lett. (1995), 36, 3651; Shea et al., Nucl. Acids Res. (1990), 18, 3777); цепь полиамина или полиэтиленгликоля (Manoharan et al., Nucleosides & Nucleotides (1995), 14, 969); адамантануксусную кислоту (Manoharan et al., Tetrahedron Lett. (1995), 36, 3651); пальмитильный фрагмент (Mishra et al., Biochim. Biophys. Acta. (1995), 1264, 229); или фрагмент октадециламина или гексиламинокарбонилхолестерина (Crooke et al., J. Pharmacol. Exp. Ther. (1996), 277, 923).

#### *Типы антисмысловых соединений*

Можно применять различные типы АС, например, включая антисмысловый олигонуклеотид, кРНК, мРНК, антагомиР, аптамер, рибозим, супермиР, миметик миРНК, ингибитор миРНК или их комбинации.

#### *Антисмысловые олигонуклеотиды*



Согласно различным вариантам реализации антисмысловое соединение (АС) представляет собой антисмысловой олигонуклеотид (ASO), который комплементарен целевой нуклеотидной последовательности. Подразумевается, что термин «антисмысловой олигонуклеотид (ASO)» или просто «антисмысловой» включает олигонуклеотиды, которые комплементарны целевой нуклеотидной последовательности. Термин также охватывает ASO, которые могут не быть полностью комплементарными желаемой целевой нуклеотидной последовательности. ASO включают отдельные цепи ДНК и/или РНК, которые комплементарны выбранной целевой нуклеотидной последовательности или целевому гену. ASO могут включать одно или более модифицированных оснований ДНК и/или РНК, модифицированных сахаров и/или неприродных межнуклеозидных связей. Согласно вариантам реализации ASO могут включать одну или более фосфорамидатных межнуклеозидных связей. Согласно вариантам реализации ASO представляет собой фосфорамидат-морфолиновые олигомеры (PMO). ASO могут иметь любую характеристику, любую длину, могут связываться с любой целевой нуклеотидной последовательностью и/или элементом последовательности и осуществлять любой механизм, описанный касательно АС.

Было продемонстрировано, что антисмысловые олигонуклеотиды являются эффективными в качестве нацеленных ингибиторов синтеза белка и, следовательно, могут применяться для специфичного ингибирования синтеза белка геном, который является мишенью. Эффективность ASO в отношении ингибирования синтеза белка хорошо установлена. На сегодняшний день эти соединения были перспективными в нескольких моделях *in vitro* и *in vivo*, включая модели воспалительного заболевания, рака и ВИЧ (Agrawal, Trends in Biotech. (1996), 14:376-387). Антисмысловое соединение также может влиять на клеточную активность путем специфичной гибридизации с хромосомной ДНК.

Способы получения антисмысловых олигонуклеотидов известны в данной области техники и могут быть легко адаптированы для получения антисмыслового олигонуклеотида, который нацелен на любую полинуклеотидную последовательность. Отбор антисмысловых олигонуклеотидных последовательностей, специфичных в отношении конкретной целевой последовательности, основан на анализе выбранной целевой последовательности и определении вторичной структуры,  $T_m$ , энергии связывания и относительной стабильности. Антисмысловые олигонуклеотиды могут быть отобраны на основании их относительной неспособности образовывать димеры, шпилечные или другие вторичные структуры, которые будут снижать или препятствовать специфичному связыванию с целевой мРНК в клетке-хозяине. Целевые области мРНК включают области, которые находятся у AUG-кодона инициации трансляции или вблизи него, и те последовательности, которые по существу комплементарны 5'-областям мРНК. Эти анализы вторичной структуры и соображения по выбору целевого сайта могут быть реализованы, например, с использованием версии 4 программного обеспечения для анализа праймеров OLIGO (Molecular Biology Insights) и/или программного обеспечения алгоритма BLASTN 2.0.5 (Altschul et al, Nucleic Acids Res. 1997, 25(17):3389-402).

### *РНК-интерференция*

Согласно вариантам реализации АС включает молекулу, которая опосредует РНК-интерференцию (РНКи). В настоящем документе выражение «опосредует РНКи» относится к способности подавлять целевой транскрипт специфичным для последовательности образом. Не ограничиваясь какой-либо теорией, полагают, что при сайленсинге используется механизм или процесс РНКи и направляющая РНК, например, соединение киРНК, содержащее от примерно 21 до примерно 23 нуклеотидов. Согласно вариантам реализации АС нацелен на целевой транскрипт для разрушения. Таким образом, в вариантах реализации молекулу РНКи можно применять для нарушения экспрессии представляющего интерес гена или полинуклеотида. Согласно вариантам реализации молекулу РНКи используют для индукции разрушения целевого транскрипта, такого как пре-мРНК или зрелая мРНК.

Согласно вариантам реализации АС включает короткую интерферирующую РНК (киРНК), которая вызывает ответ РНКи.

Короткие интерферирующие РНК (киРНК) представляют собой дуплексы нуклеиновых кислот, длина которых обычно составляет от примерно 16 до примерно 30 нуклеотидов, которые могут ассоциироваться с цитоплазматическим мультибелковым комплексом, известным как РНКи-индуцированный комплекс сайленсинга (RISC). RISC, загруженный киРНК, опосредует разрушение гомологичных транскриптов, поэтому киРНК может быть разработана для нокдауна экспрессии белка с высокой специфичностью. В отличие от других антисмысловых технологий киРНК функционирует посредством природного механизма, развившегося для контроля экспрессии генов посредством некодирующей РНК. Различные реагенты для РНКи, включая киРНК, нацеленные на клинически значимые мишени, в настоящее время находятся в фармацевтической разработке, как описано, например, в de Fougerolles, A. et al., *Nature Reviews* (2007) 6:443-453.

Хотя первыми описанными молекулами РНКи были гибриды РНК:РНК, которые включают как смысловую цепь РНК, так и антисмысловую цепь РНК, было продемонстрировано, что гибриды смысловая цепь ДНК:антисмысловая цепь РНК, гибриды смысловая цепь РНК:антисмысловая цепь ДНК и гибриды ДНК:ДНК способны опосредовать РНКи (Lamberton, J.S. and Christian, A.T., *Molecular Biotechnology* (2003), 24:111-119). Согласно вариантам реализации применяют молекулы РНКи, которые включают любой из этих разных типов двухцепочечных молекул. Кроме того, понятно, что молекулы РНКи можно применять и вводить в клетки в различных формах. Соответственно, в настоящем документе молекулы РНКи охватывают любые и все молекулы, способные опосредовать РНКи в клетках, включая, но не ограничиваясь перечисленными, двухцепочечные олигонуклеотиды, которые включают две отдельные цепи, т. е. смысловую цепь и антисмысловую цепь, например, короткую интерферирующую РНК (киРНК); двухцепочечный олигонуклеотид, который включает две отдельные цепи, которые связаны друг с другом ненуклеотидильным линкером;

олигонуклеотиды, которые включают петлю шпильки из комплементарных последовательностей, которая образует двухцепочечную область, например, молекулы кшРНКи, и векторы экспрессии, которые экспрессируют один или более полинуклеотидов, способных образовывать двухцепочечный полинуклеотид отдельно или в комбинации с другим полинуклеотидом.

«Соединение одноцепочечной киРНК» в контексте настоящего документа представляет собой соединение киРНК, которое состоит из отдельной молекулы. Оно может содержать дуплексную область, образованную путем внутрицепочечного спаривания, например, оно может представлять собой или включать шпильку или структуру «ручки». Соединения одноцепочечной киРНК могут быть антисмысловыми по отношению к молекуле-мишени.

Соединение одноцепочечной киРНК может быть достаточно длинным, чтобы оно могло поступать в RISC и участвовать в RISC-опосредуемом расщеплении целевой мРНК. Соединение одноцепочечной киРНК имеет длину по меньшей мере примерно 14, по меньшей мере примерно 15, по меньшей мере примерно 20, по меньшей мере примерно 25, по меньшей мере примерно 30, по меньшей мере примерно 35, по меньшей мере примерно 40 или до примерно 50 нуклеотидов. Согласно определенным вариантам реализации длина одноцепочечной киРНК составляет менее примерно 200, примерно 100 или примерно 60 нуклеотидов.

Соединения шпилечной киРНК могут иметь дуплексную область, которая равна или содержит по меньшей мере примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24 или примерно 25 пар нуклеотидов. Длина дуплексной области может быть равна или меньше примерно 200, примерно 100 или примерно 50 пар нуклеотидов. Согласно определенным вариантам реализации диапазоны длины дуплексной области составляют от примерно 15 до примерно 30, от примерно 17 до примерно 23, от примерно 19 до примерно 23 и от примерно 19 до примерно 21 пары нуклеотидов. Шпилька может иметь одноцепочечный липкий конец или концевую неспаренную область. Согласно определенным вариантам реализации липкие концы имеют длину от примерно 2 до примерно 3 нуклеотидов. Согласно вариантам реализации липкий конец находится на смысловой стороне шпильки и в вариантах реализации он находится на антисмысловой стороне шпильки.

«Соединение двухцепочечной киРНК» в контексте настоящего документа представляет собой соединение киРНК, которое включает более одной, а в некоторых случаях две, цепи, в которых межцепочечная гибридизация может приводить к образованию области дуплексной структуры.

Антисмысловая цепь соединения двухцепочечной киРНК может иметь длину, равную или по меньшей мере примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 25, примерно 30, примерно 40 или примерно 60 нуклеотидов. Ее длина может быть равна или меньше примерно 200, примерно 100 или примерно 50 нуклеотидов. Диапазоны длины могут включать от

примерно 17 до примерно 25, от примерно 19 до примерно 23 и от примерно 19 до примерно 21 нуклеотида. В настоящем документе термин «антисмысловая цепь» означает цепь соединения киРНК, которая достаточно комплементарна молекуле-мишени, например, целевой нуклеотидной последовательности целевого транскрипта.

Смысловая цепь соединения двухцепочечной киРНК может иметь длину, равную или по меньшей мере примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 25, примерно 30, примерно 40 или примерно 60 нуклеотидов. Ее длина может быть равна или меньше примерно 200, примерно 100 или примерно 50 нуклеотидов. Диапазоны длины могут включать от примерно 17 до примерно 25, от примерно 19 до примерно 23 и от примерно 19 до примерно 21 нуклеотида.

Длина двухцепочечной части соединения двухцепочечной киРНК может быть равна или по меньшей мере примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 30, примерно 40 или примерно 60 пар нуклеотидов. Ее длина может быть равна или меньше примерно 200, примерно 100 или примерно 50 пар нуклеотидов. Диапазоны длины могут включать от примерно 15 до примерно 30, от примерно 17 до примерно 23, от примерно 19 до примерно 23 и от примерно 19 до примерно 21 пары нуклеотидов.

Согласно вариантам реализации соединение киРНК является достаточно большим, чтобы его можно было расщепить эндогенной молекулой, например, Dicer, с получением меньших соединений киРНК, например, агентов на основе киРНК.

Смысловая и антисмысловая цепи могут быть выбраны таким образом, чтобы соединение двухцепочечной киРНК содержало одиночную цепь или неспаренную область на одном или обоих концах молекулы. Таким образом, соединение двухцепочечной киРНК может содержать смысловую и антисмысловую цепи, спаренные таким образом, чтобы содержать липкий конец, например, один или два 5'- или 3'-липких конца или 3'-липкий конец из 1-3 нуклеотидов. Липкие концы могут быть получены в результате того, что одна цепь длиннее другой, или в результате того, что две цепи одинаковой длины расположены в шахматном порядке. Некоторые варианты реализации будут иметь по меньшей мере один 3'-липкий конец. Согласно вариантам реализации оба конца молекулы киРНК будут иметь 3'-липкий конец. Согласно вариантам реализации липкий конец представляет собой 2 нуклеотида.

Согласно вариантам реализации длина дуплексной области составляет от примерно 15 до примерно 30 или примерно 18, примерно 19, примерно 20, примерно 21, примерно 22 или примерно 23 нуклеотида, например, в обсуждаемом выше диапазоне для соединений окиРНК (киРНК с липкими концами). Соединения окиРНК могут напоминать по длине и структуре природные продукты процессинга Dicer из длинных дкиРНК. Также включены варианты реализации, в которых две цепи соединения окиРНК связаны, например, ковалентно связаны. Согласно вариантам реализации включены шпилечные

или другие одноцепочечные структуры, которые обеспечивают двухцепочечную область и 3'-липкие концы.

Соединения киРНК, описанные в настоящем документе, включая соединения двухцепочечной киРНК и соединения одноцепочечной киРНК, могут опосредовать сайленсинг целевой РНК, например, мРНК, например, транскрипта гена, который кодирует белок. Для удобства такая мРНК также называется в настоящем документе мРНК, подлежащей сайленсингу. Такой ген также называется целевым геном. В целом, РНК, подлежащая сайленсингу, представляет собой эндогенный ген.

Согласно вариантам реализации соединение киРНК «достаточно комплементарно» целевому транскрипту так, что соединение киРНК подавляет продукцию белка, кодируемого целевой мРНК. Согласно вариантам реализации соединение киРНК «достаточно комплементарно» по меньшей мере части целевого транскрипта так, что соединение киРНК подавляет продукцию генного продукта, кодируемого целевым транскриптом. Согласно другому варианту реализации соединение киРНК «точно комплементарно» целевой нуклеотидной последовательности (например, части целевого транскрипта) так, что целевая нуклеотидная последовательность и соединение киРНК ренатурируются, например, с образованием гибрида, состоящего исключительно из пар оснований по Уотсону-Крику в области точной комплементарности. «Достаточно комплементарный» целевой нуклеотидной последовательности может включать внутреннюю область (например, по меньшей мере примерно 10 нуклеотидов), которая точно комплементарна целевой нуклеотидной последовательности. Кроме того, в определенных вариантах реализации соединение киРНК специфично различает однонуклеотидное различие. В этом случае соединение киРНК опосредует РНКи только в том случае, если в области (например, в пределах 7 нуклеотидов) однонуклеотидного различия обнаружена точная комплементарность.

Способы терапевтического применения РНКи являются чрезвычайно широкими, поскольку конструкции киРНК и миРНК можно синтезировать с любой нуклеотидной последовательностью, направленной против целевого белка. На сегодняшний день конструкции киРНК продемонстрировали способность специфично отрицательно регулировать целевые белки в моделях как *in vitro*, так и *in vivo*, а также в клинических исследованиях.

### *МикроРНК*

Согласно вариантам реализации АС включает молекулу микроРНК. МикроРНК (миРНК) представляют собой высококонсервативный класс малых молекул РНК, которые транскрибируются из ДНК в геномах растений и животных, но не транслируются в белок. Процессированные миРНК представляют собой одноцепочечные молекулы РНК из 17-25 нуклеотидов, которые встраиваются в РНК-индуцированный комплекс сайленсинга (RISC) и идентифицированы как ключевые регуляторы развития, пролиферации клеток, апоптоза и дифференцировки. Считается, что они играют роль в регуляции экспрессии генов путем связывания с 3'-нетранслируемой областью специфичных мРНК. RISC

опосредует отрицательную регуляцию экспрессии генов посредством трансляционного ингибирования, расщепления транскриптов или и того, и другого. RISC также участвует в транскрипционном сайленсинге в ядре широкого диапазона эукариот.

#### *Антагомиры*

Согласно вариантам реализации АС представляет собой антагомир. Антагомиры представляют собой РНК-подобные олигонуклеотиды, которые несут различные модификации для защиты от РНКазы и придания таких фармакологических свойств как улучшенное поглощение тканями и клетками. Они отличаются от нормальной РНК, например, полным 2'-О-метилованием сахара, фосфотиоатным остовом и, например, холестеринным фрагментом на 3'-конце. Антагомиры можно применять для эффективного сайленсинга эндогенных миРНК путем образования дуплексов, которые включают антагомир и эндогенную миРНК, что предотвращает миРНК-индуцированный сайленсинг генов. Примером антагомир-опосредуемого сайленсинга миРНК является сайленсинг miR-122, описанный в Krutzfeldt et al., Nature (2005), 438: 685-689, которая явным образом полностью включена в настоящий документ посредством ссылки. РНК антагомира можно синтезировать с использованием стандартных протоколов твердофазного синтеза олигонуклеотидов (заявка на патент США с серийным номером 11/502158 и 11/657341; раскрытие каждой из которых включено в настоящий документ посредством ссылки).

Антагомир может включать конъюгированные с лигандом мономерные субъединицы и мономеры для синтеза олигонуклеотидов. Мономеры описаны в заявке на патент США № 10/916185. Антагомир может иметь структуру ZXY, такую как описанная в РСТ заявке № РСТ/US2004/07070. Антагомир может образовывать комплекс с амфипатическим фрагментом. Амфипатические фрагменты для применения с олигонуклеотидными агентами описаны в РСТ заявке № РСТ/US2004/07070.

#### *Аптамеры*

Согласно вариантам реализации АС включает аптамер. Аптамеры представляют собой молекулы нуклеиновой кислоты или пептида, которые связываются с конкретной представляющей интерес молекулой с высокой аффинностью и специфичностью (Tuerk and Gold, Science 249:505 (1990); Ellington and Szostak, Nature 346:818 (1990)). Были успешно получены аптамеры ДНК или РНК, которые связывают множество различных частиц от больших белков до малых органических молекул (Eaton, Curr. Opin. Chem. Biol. (1997), 1: 10-16; Famulok, Curr. Opin. Struct. Biol. (1999), 9:324-9; и Hermann and Patel, Science (2000), 287:820-5). Аптамеры могут быть на основе РНК или ДНК и могут содержать рибопереключател. Рибопереключател представляет собой часть молекулы мРНК, которая может непосредственно связывать малую молекулу-мишень, и связывание которой с мишенью влияет на активность гена. Таким образом, мРНК, которая содержит рибопереключател, непосредственно участвует в регуляции своей собственной активности в зависимости от наличия или отсутствия ее молекулы-мишени. Обычно аптамеры конструируют с помощью повторяющихся циклов отбора *in vitro* или,

эквивалентно, SELEX (систематической эволюции лигандов путем экспоненциального обогащения) для связывания с различными молекулярными мишенями, такими как малые молекулы, белки, нуклеиновые кислоты, и даже с клетками, тканями и организмами. Аптамер может быть получен любым известным способом, включая синтетические, рекомбинантные способы и способы очистки, и его можно применять по отдельности или в комбинации с другими аптамерами, специфичными в отношении той же мишени. Кроме того, термин «аптамер» также включает «вторичные аптамеры», содержащие консенсусную последовательность, полученную путем сравнения двух или более известных аптамеров к конкретной мишени. Согласно вариантам реализации аптамер представляет собой «внутриклеточный аптамер» или «интрамер», который специфично распознает внутриклеточные мишени (Famulok et al., Chem Biol. (2001),8(10):931-939; Yoon and Rossi, Adv. Drug Deliv. Rev. (2018), 134:22-35; каждая из которых включена в настоящий документ посредством ссылки).

### *Рибозимы*

Согласно вариантам реализации АС представляет собой рибозим. Рибозимы представляют собой комплексы молекул РНК, имеющие специфичные каталитические домены, которые обладают эндонуклеазной активностью (Kim and Cech, Proc. Natl. Acad. Sci. USA (1987),84(24):8788-92; Forster and Symons, Cell (1987) 24, 49(2):211-20). Например, большое число рибозимов ускоряют реакции переноса фосфоэфиров с высокой степенью специфичности, часто расщепляя только один из нескольких фосфоэфиров в олигонуклеотидном субстрате (Cech et al., Cell (1981), 27(3 Pt 2):487-96; Michel and Westhof, J. Mol. Biol. (1990), 5, 216(3):585-610; Reinhold-Hurek and Shub, Nature (1992), 14, 357(6374): 173-6). Эта специфичность была отнесена к требованию к связыванию субстрата за счет специфичных взаимодействий спаривания оснований с внутренней направляющей последовательностью (IGS) рибозима перед химическим взаимодействием.

В настоящее время известны по меньшей мере шесть основных разновидностей встречающихся в природе ферментативных РНК. Каждая из них может катализировать транс-гидролиз фосфодиэфирных связей РНК (и, таким образом, может расщеплять другие молекулы РНК) в физиологических условиях. В целом, ферментативные нуклеиновые кислоты действуют путем первоначального связывания с целевой РНК. Такое связывание происходит посредством связывающей мишень части ферментативной нуклеиновой кислоты, которая удерживается в непосредственной близости от ферментативной части молекулы, которая расщепляет целевую РНК. Таким образом, ферментативная нуклеиновая кислота сначала распознает и затем связывает целевую РНК посредством комплементарного спаривания оснований и после связывания с правильным сайтом действует ферментативно для разрезания целевой РНК. Стратегическое расщепление такой целевой РНК нарушит ее способность управлять синтезом кодируемого белка. После связывания ферментативной нуклеиновой кислоты и расщепления ее целевой РНК она высвобождается от этой РНК для поиска другой мишени и может повторно связывать и расщеплять новые мишени.

Молекула ферментативной нуклеиновой кислоты может быть образована, например, в РНК «головки молотка», шпильки, вируса гепатита  $\delta$ , интрона группы I или РНКазы Р (в сочетании с направляющей последовательностью РНК) или, например, мотиве РНК *Neurospora VS*. Конкретные примеры мотивов «головки молотка» описаны Rossi et al. *Nucleic Acids Res.* (1992), 20(17):4559-65. Примеры шпильчных мотивов описаны Hampel et al. (европейская публикация заявки на патент EP 0360257), Hampel and Tritz, *Biochemistry* (1989), 28(12):4929-33; Hampel et al, *Nucleic Acids Res.* (1990),18(2):299-304 и в патенте США 5631359. Пример мотива вируса гепатита описан Perrotta and Been, *Biochemistry* (1992), 31(47): 11843-52; пример мотива РНКазы Р описан Guerrier-Takada et al., *Cell* (1983), 35(3 Pt 2):849-57; мотив рибозима РНК *Neurospora VS* описан Collins (Saville and Collins, *Cell* (1990), 61(4):685-96; Saville and Collins, *Proc. Natl. Acad. Sci. USA* (1991),88(19):8826-30; Collins and Olive, *Biochemistry* (1993),32(11):2795-9); и пример интрона группы I описан в патенте США № 4987071. Согласно вариантам реализации молекулы ферментативной нуклеиновой кислоты имеют сайт специфического связывания субстрата, который комплементарен одной или более областям ДНК или РНК целевого гена, и они имеют нуклеотидные последовательности внутри или вокруг этого сайта связывания субстрата, которые придают молекуле РНК-расщепляющую активность. Таким образом, конструкции рибозимов не обязательно должны ограничиваться конкретными мотивами, упомянутыми в настоящем документе.

Рибозимы могут быть сконструированы, как описано в международной публикации заявки на патент № WO 93/23569 и международной публикации заявки на патент № WO 94/02595, каждая из которых намерено включена в настоящий документ посредством ссылки, и синтезированы для тестирования *in vitro* и *in vivo*, как описано в указанных документах. Согласно вариантам реализации рибозим нацелен на целевую нуклеотидную последовательность целевого транскрипта.

Активность рибозима может быть повышена путем изменения длины связывающих плеч рибозима или химического синтеза рибозимов с модификациями, которые предотвращают их разрушение сывороточными рибонуклеазами (см., например, международную публикацию заявки на патент № WO 92/07065; международную публикацию заявки на патент № WO 93/15187; международную публикацию заявки на патент № WO 91/03162, европейскую публикацию заявки на патент № 92110298.4; патент США № 5334711; и международную публикацию заявки на патент № WO 94/13688, в которых описаны различные химические модификации, которые могут быть внесены в сахарные фрагменты ферментативных молекул РНК), модификациями, которые повышают их эффективность в клетках, и путем удаления П-оснований стебля для сокращения времени синтеза РНК и снижения химических требований.

### *Супермир*

Согласно вариантам реализации АС представляет собой супермир. Супермир относится к одноцепочечному, двухцепочечному или частично двухцепочечному олигомеру или полимеру РНК, полимеру ДНК, или и тому, и другому, или их



модификациям, который имеет нуклеотидную последовательность, которая по существу идентична miРНК и которая является бессмысленной относительно своей мишени. Этот термин включает олигонуклеотиды, состоящие из встречающихся в природе нуклеиновых оснований, сахаров и ковалентных межнуклеозидных (остовных) связей и которые содержат по меньшей мере одну не встречающуюся в природе часть, которая функционирует подобным образом. Такие модифицированные или замещенные олигонуклеотиды обладают желаемыми свойствами, такими как, например, повышенное клеточное поглощение, повышенная аффинность в отношении целевой нуклеиновой кислоты и повышенная стабильность в присутствии нуклеаз. Согласно вариантам реализации супермира не включает смысловую цепь, и в другом варианте реализации супермир не является самогибридирующимся в значительной степени. Супермир может иметь вторичную структуру, но он является по существу одноцепочечным при физиологических условиях. Супермир, который является по существу одноцепочечным, является одноцепочечным до такой степени, что менее примерно 50% (например, менее примерно 40%, примерно 30%, примерно 20%, примерно 10% или примерно 5%) супермира могут образовывать дуплекс с самим собой. Супермир может включать шпилечный сегмент, например, последовательность, например, на 3'-конце может самогибридизоваться и образовывать дуплексную область, например, дуплексную область, состоящую из по меньшей мере примерно 1, примерно 2, примерно 3 или примерно 4 или менее примерно 8, примерно 7, примерно 6 или примерно 5 нуклеотидов, или примерно 5 нуклеотидов. Дуплексная область может быть соединена линкером, например, нуклеотидным линкером, например, примерно 3, примерно 4, примерно 5 или примерно 6 dT, например, модифицированными dT. Согласно другому варианту реализации супермир образует дуплекс с более коротким олигонуклеотидом, например, длиной примерно 5, примерно 6, примерно 7, примерно 8, примерно 9 или примерно 10 нуклеотидов, например, на одном или обоих из 3'- и 5'-конца или на одном конце и в неконцевой области или в середине супермира.

#### *Миметики miРНК*

Согласно вариантам реализации АС представляет собой миметик miРНК. Миметики miРНК представляют собой класс молекул, которые можно применять для имитации способности одной или более miРНК к сайленсингу генов. Таким образом, термин «миметик микроРНК» относится к синтетическим некодирующим РНК (т. е. miРНК получена не путем очистки из источника эндогенной miРНК), которые способны вступать в путь РНКи и регулировать экспрессию генов. Миметики miРНК могут быть разработаны в виде зрелых молекул (например, одноцепочечных) или предшественников миметиков (например, при- или пре-miРНК). Миметки miРНК могут включать нуклеиновую кислоту (модифицированные или модифицированные нуклеиновые кислоты), включая олигонуклеотиды, которые включают, но не ограничиваются перечисленными, РНК, модифицированную РНК, ДНК, модифицированную ДНК, заблокированные нуклеиновые кислоты или 2'-O,4'-C-этилен-мостиговые нуклеиновые

кислоты (ENA), или любую комбинацию перечисленных выше (включая гибриды ДНК-РНК). Кроме того, миметики миРНК могут включать конъюгаты, которые могут влиять на доставку, внутриклеточную компартиментализацию, стабильность, специфичность, функциональность, использование цепи и/или эффективность. В одной конструкции миметики миРНК представляют собой двухцепочечные молекулы (например, с дуплексной областью длиной от примерно 16 до примерно 31 нуклеотида) и содержат одну или более последовательностей, которые идентичны зрелой цепи конкретной миРНК. Модификации могут включать 2'-модификации (включая 2'-О-метил-модификации и 2'-F-модификации) на одной или обеих цепях молекулы и межнуклеозидные модификации (например, фосфотиоатные модификации), которые повышают стабильность и/или специфичность нуклеиновой кислоты. Кроме того, миметики миРНК могут включать липкие концы. Липкие концы могут включать от примерно 1 до примерно 6 нуклеотидов на 3'- или 5'-конце любой цепи и могут быть модифицированы для повышения стабильности или функциональности. Согласно вариантам реализации миметик миРНК включает дуплексную область из примерно 16-31 нуклеотида и один или более из следующих профилей химической модификации: смысловая цепь содержит 2'-О-метил-модификации нуклеотидов 1 и 2 (считая от 5'-конца смыслового олигонуклеотида) и всех из Cs и Us; модификации антисмысловой цепи могут включать 2'-F-модификацию всех Cs и Us, фосфорилирование 5'-конца олигонуклеотида и стабилизированные межнуклеозидные связи, ассоциированные с 2-нуклеотидным 3'-липким концом.

#### *Ингибитор миРНК*

Согласно вариантам реализации АС представляет собой ингибитор миРНК. Термины «антимир», «ингибитор микроРНК», «ингибитор миР» или «ингибитор миРНК» являются синонимами и относятся к олигонуклеотидам или модифицированным олигонуклеотидам, которые препятствуют способности специфичных миРНК. В целом, ингибиторы представляют собой нуклеиновую кислоту или модифицированные нуклеиновые кислоты по своей природе, включая олигонуклеотиды, которые включают РНК, модифицированную РНК, ДНК, модифицированную ДНК, заблокированные нуклеиновые кислоты (LNA) или любую комбинацию перечисленных выше.

Модификации включают 2'-модификации (включая 2'-О-алкильные модификации и 2'-F-модификации) и межнуклеозидные модификации (например, фосфотиоатные модификации), которые могут влиять на доставку, стабильность, специфичность, внутриклеточную компартиментализацию или эффективность. Кроме того, ингибиторы миРНК могут включать конъюгаты, которые могут влиять на доставку, внутриклеточную компартиментализацию, стабильность и/или эффективность. Ингибиторы могут принимать различные конфигурации, включая одноцепочечные, двухцепочечные (дуплексы РНК/РНК или РНК/ДНК) и шпилечные конструкции, в целом, ингибиторы микроРНК содержат одну или более последовательностей или частей последовательностей, которые комплементарны или частично комплементарны зрелой цепи (или цепям) миРНК, на

которую требуется нацелиться. Кроме того, ингибитор миРНК также может включать дополнительные последовательности, расположенные в направлении 5'- и 3'-конца от последовательности, которая является обратнoкомплементарной зрелой миРНК. Дополнительные последовательности могут быть обратнoкомплементарными последовательностям, которые являются смежными со зрелой миРНК в при-миРНК, из которой происходит зрелая миРНК, или дополнительные последовательности могут представлять собой произвольные последовательности (имеющие смесь А, G, C или U). Согласно вариантам реализации одна или обе дополнительные последовательности представляют собой произвольные последовательности, способные образовывать шпильки. Таким образом, в вариантах реализации последовательность, которая является обратнoкомплементарной миРНК, фланкирована на 5'-стороне и на 3'-стороне шпилечными структурами. Ингибиторы микроРНК, если они являются двухцепочечными, могут включать несовпадения между нуклеотидами на противоположных цепях. Кроме того, ингибиторы микроРНК могут быть связаны с конъюгатными фрагментами, чтобы облегчить поглощение ингибитора в клетку. Например, ингибитор микроРНК может быть связан с холестерил-5-(бис(4-метоксифенил)(фенил)метокси)-3-гидроксипентилкарбаматом), что обеспечивает пассивное поглощение ингибитора микроРНК в клетку. Ингибиторы микроРНК, включая шпилечные ингибиторы микроРНК, подробно описаны в Vermeulen et al., RNA 13: 723- 730 (2007) и в WO2007/095387 и WO 2008/036825, каждый из которых полностью включен в настоящий документ посредством ссылки. Обычный специалист в данной области техники может выбрать последовательность из базы данных для требуемой миРНК и разработать ингибитор, пригодный для способов, раскрытых в настоящем документе.

Линкерные группы или бифункциональные линкерные фрагменты, такие как те, которые известны в данной области техники, пригодны для соединений, предложенных в настоящем документе. Линкерные группы можно применять для присоединения химических функциональных групп, конъюгатных групп, репортерных групп и других групп к селективным сайтам в исходном соединении, таком как, например, АС. В целом, бифункциональный линкерный фрагмент включает гидрокарбильный фрагмент, имеющий две функциональные группы. Одна из функциональных групп выбрана для связывания с исходной молекулой или представляющим интерес соединением, а другая выбрана для связывания по существу любой выбранной группы, такой как химическая функциональная группа или конъюгатная группа. Можно применять любой из линкеров, описанных в настоящем документе. Согласно вариантам реализации линкер включает цепную структуру или олигомер из повторяющихся звеньев, таких как этиленгликолевые или аминокислотные звенья. Примеры функциональных групп, которые рутинно применяют в бифункциональном линкерном фрагменте, включают, но не ограничиваются перечисленными, электрофилы для взаимодействия с нуклеофильными группами и нуклеофилы для взаимодействия с электрофильными группами. Согласно вариантам реализации бифункциональные линкерные фрагменты включают амино, гидроксил,

карбоновую кислоту, тиол, ненасыщенность (например, двойные или тройные связи) и т. п. Некоторые неограничивающие примеры бифункциональных линкерных фрагментов включают 8-амино-3,6-диоксооктановую кислоту (ADO), сукцинимидил-4-(N-малеимидометил)циклогексан-1-карбоксилат (SMCC) и 6-аминогексановую кислоту (АНЕХ или АНА). Другие линкерные группы включают, но не ограничиваются перечисленными, замещенный C1-C10 алкил, замещенный или незамещенный C2-C10 алкенил или замещенный или незамещенный C2-C10 алкинил, причем неограничивающий перечень замещающих групп включает гидроксил, амино, алкокси, карбокси, бензил, фенил, нитро, тиол, тиалкокси, галоген, алкил, арил, алкенил и алкинил.

Согласно вариантам реализации АС включает нуклеотидную модификацию, разработанную для того, чтобы не поддерживать активность РНКазы Н. Нуклеотидные модификации антисмысловых соединений, которые не поддерживают активность РНКазы Н, известны и включают, но не ограничиваются перечисленными, 2'-О-метоксиэтил/фосфотиоатные (МОЕ) модификации. Предпочтительно АС с МОЕ-модификациями имеют повышенную аффинность в отношении РНК-мишени и повышенную стабильность против нуклеаз.

#### *Иммуностимулирующие олигонуклеотиды*

Согласно вариантам реализации терапевтический фрагмент представляет собой иммуностимулирующий олигонуклеотид. Иммуностимулирующие олигонуклеотиды (ISS; одно- или двухцепочечные) способны индуцировать иммунный ответ при введении пациенту, который может представлять собой млекопитающее или другого пациента. ISS включают, например, определенные палиндромы, приводящие к образованию шпилечных вторичных структур (см. Yamamoto S., et al. (1992) J. Immunol. 148: 4072-4076), или мотивы CpG, а также другие известные признаки ISS (такие как домены мульти-G, см. WO 96/11266).

Иммунный ответ может представлять собой врожденный или адаптивный иммунный ответ. Иммунная система разделена на более врожденную иммунную систему и приобретенную адаптивную иммунную систему позвоночных, последняя из которых дополнительно разделена на гуморальные клеточные компоненты. Согласно конкретным вариантам реализации иммунный ответ может происходить в слизистой оболочке.

Иммуностимулирующие нуклеиновые кислоты считаются неспецифичными в отношении последовательности, когда не требуется, чтобы они специфично связывались с целевым полинуклеотидом и снижали его экспрессию, чтобы вызвать иммунный ответ. Таким образом, определенные иммуностимулирующие нуклеиновые кислоты могут включать последовательность, соответствующую области встречающегося в природе гена или мРНК, но они все еще могут считаться неспецифичными в отношении последовательности иммуностимулирующими нуклеиновыми кислотами.

Согласно вариантам реализации иммуностимулирующая нуклеиновая кислота или олигонуклеотид включает по меньшей мере один динуклеотид CpG. Олигонуклеотид или динуклеотид CpG могут быть неметилованными или метилированными. Согласно

другому варианту реализации иммуностимулирующая нуклеиновая кислота включает по меньшей мере один динуклеотид CpG, имеющий метилированный цитозин. Согласно вариантам реализации нуклеиновая кислота включает один динуклеотид CpG, причем цитозин в указанном динуклеотиде CpG является метилированным. Согласно конкретному варианту реализации нуклеиновая кислота включает последовательность 5' TAACGTTGAGGG'CAT 3' (SEQ ID NO: 369). Согласно альтернативному варианту реализации нуклеиновая кислота включает по меньшей мере два динуклеотида CpG, причем по меньшей мере один цитозин в динуклеотидах CpG является метилированным. Согласно дополнительному варианту реализации каждый цитозин в динуклеотидах CpG, присутствующих в последовательности, является метилированным. Согласно другому варианту реализации нуклеиновая кислота включает множество динуклеотидов CpG, причем по меньшей мере один из указанных динуклеотидов CpG включает метилированный цитозин.

Дополнительные специфичные последовательности нуклеиновых кислот олигонуклеотидов (ODN), подходящие для применения в композициях и способах, описаны в Raney et al, Journal of Pharmacology and Experimental Therapeutics, 298:1185-1192 (2001). Согласно определенным вариантам реализации ODN, применяемые в композициях и способах, имеют фосфодизфирный («PO») остов или фосфотиоатный («PS») остов и/или по меньшей мере один метилированный остаток цитозина в мотиве CpG.

#### *Олигонуклеотиды-ловушки*

Согласно вариантам реализации терапевтический фрагмент представляет собой олигонуклеотид-ловушку. Поскольку факторы транскрипции распознают их относительно короткие последовательности связывания, даже в отсутствие окружающей геномной ДНК, короткие олигонуклеотиды, несущие консенсусную связывающую последовательность конкретного фактора транскрипции, можно применять в качестве инструментов для манипулирования экспрессией генов в живых клетках. Эта стратегия включает внутриклеточную доставку таких «олигонуклеотидов-ловушек», которые затем распознаются и связываются целевым фактором. Занятость ДНК-связывающего сайта фактора транскрипции ловушкой делает фактор транскрипции неспособным впоследствии связываться с промоторными областями целевых генов. Ловушки можно применять в качестве терапевтических агентов, либо для ингибирования экспрессии генов, которые активируются фактором транскрипции, либо для положительной регуляции генов, которые подавляются связыванием фактора транскрипции. Примеры использования олигонуклеотидов-ловушек можно найти в Mann et al., J. Clin. Invest, 2000, 106: 1071-1075, которая явным образом полностью включена в настоящий документ посредством ссылки.

#### *Адаптер U1*

Согласно некоторым вариантам реализации терапевтический фрагмент представляет собой адаптер U1. Адаптеры U1 ингибируют сайты полиА и представляют собой бифункциональные олигонуклеотиды с целевым доменом, комплементарным сайту

в концевом экзоне целевого гена, и «доменом U1», который связывается с компонентом малой ядерной РНК U1 из U1 мяРНП (Goraczniak, et al., 2008, Nature Biotechnology, 27(3), 257-263, которая явным образом полностью включена в настоящий документ посредством ссылки). U1 мяРНП представляет собой рибонуклеопротеиновый комплекс, функция которого заключается главным образом в управлении ранними стадиями образования сплайсосомы путем связывания с границей экзон-интрон пре-мРНК (Brown and Simpson, 1998, Annu Rev Plant Physiol Plant Mol Biol 49:77-95). Нуклеотиды 2-11 5'-конца пары оснований U1 мяРНК связываются с 5'-ss пре-РНК. Согласно одному варианту реализации олигонуклеотиды представляют собой адаптеры U1. Согласно одному варианту реализации адаптер U1 можно вводить в комбинации по меньшей мере с одним другим агентом иРНК.

#### *Механизм редактирования генов (CRISPR)*

Согласно вариантам реализации соединения, раскрытые в настоящем документе, включают один или более CPP (или сCPP), конъюгированных с механизмом редактирования генов CRISPR. В настоящем документе «механизм редактирования генов CRISPR» относится к белку, нуклеиновым кислотам или их комбинациям, которые можно применять для редактирования генома. Неограничивающие примеры механизма редактирования генов включают нРНК, нуклеазы, ингибиторы нуклеаз и их комбинации и комплексы. В следующих патентных документах описан механизм редактирования генов CRISPR: патент США № 8697359, патент США № 8771945, патент США № 8795965, патент США № 8865406, патент США № 8871445, патент США № 8889356, патент США № 8895308, патент США № 8906616, патент США № 8932814, патент США № 8945839, патент США № 8993233, патент США № 8999641, заявка на патент США № 14/704551 и заявка на патент США № 13/842859. Каждый из вышеупомянутых патентных документов полностью включен в настоящий документ посредством ссылки.

Согласно вариантам реализации с помощью линкера конъюгируют сCPP с механизмом редактирования генов CRISPR. Можно использовать любой линкер, описанный в настоящем раскрытии или известный специалисту в данной области техники.

#### *нРНК*

Согласно вариантам реализации соединения включают CPP (или сCPP), конъюгированный с нРНК. нРНК нацелена на геномные локусы в прокариотической или эукариотической клетке.

Согласно вариантам реализации нРНК представляет собой одномолекулярную направляющую РНК (онРНК). онРНК включает спейсерную последовательность и каркасную последовательность. Спейсерная последовательность представляет собой короткую последовательность нуклеиновой кислоты, используемую для нацеливания нуклеазы (например, нуклеазы Cas9) на конкретную представляющую интерес нуклеотидную область (например, последовательность геномной ДНК, подлежащую расщеплению). Согласно вариантам реализации спейсер может иметь длину примерно 17-24 основания, например, примерно 20 оснований. Согласно вариантам реализации спейсер

может иметь длину примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29 или примерно 30 оснований. Согласно вариантам реализации спейсер может иметь длину по меньшей мере 15, по меньшей мере 16, по меньшей мере 17, по меньшей мере 18, по меньшей мере 19, по меньшей мере 20, по меньшей мере 21, по меньшей мере 22, по меньшей мере 23, по меньшей мере 24, по меньшей мере 25, по меньшей мере 26, по меньшей мере 27, по меньшей мере 28, по меньшей мере 29 или по меньшей мере 30 оснований. Согласно вариантам реализации спейсер может иметь длину примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29 или примерно 30 оснований. Согласно вариантам реализации спейсерная последовательность имеет содержание GC от примерно 40% до примерно 80%.

Согласно вариантам реализации спейсер нацелен на сайт, который непосредственно предшествует 5'-мотиву, прилегающему к протоспейсеру (PAM). Последовательность PAM может быть выбрана, исходя из желаемой нуклеазы. Например, последовательность PAM может представлять собой любую из последовательностей PAM, показанных в **Таблице 3** ниже, где N относится к любой нуклеиновой кислоте, R относится к A или G, Y относится к C или T, W относится к A или T, и V относится к A или C, или G.

Таблица 3: Примерные нуклеазы и последовательности PAM

Последовательность PAM (от 5' к 3')	Нуклеаза	Выделена из
NGG	SpCas9	<i>Streptococcus pyogenes</i>
NGRRT или NGRRN	SaCas9	<i>Staphylococcus aureus</i>
NNNGATT	NmeCas9	<i>Neisseria meningitidis</i>
NNNNRYAC	CjCas9	<i>Campylobacter jejuni</i>
NNAGAAW	StCas9	<i>Streptococcus thermophiles</i>
TTTV	LbCpf1	<i>Lachnospiraceae bacterium</i>
TTTV	AsCpf1	<i>Acidaminococcus sp.</i>

Согласно вариантам реализации спейсер может быть нацелен на последовательность гена млекопитающего, например, гена человека. Согласно вариантам реализации спейсер может быть нацелен на мутантный ген. Согласно вариантам реализации спейсер может быть нацелен на кодирующую последовательность. Согласно вариантам реализации спейсер может быть нацелен на экзонную последовательность. Согласно вариантам реализации спейсер может быть нацелен на сайт полиаденилирования (PS). Согласно вариантам реализации спейсер может быть нацелен на элемент последовательности PS. Согласно вариантам реализации спейсер может быть нацелен на сигнал полиаденилирования (PAS), промежуточную последовательность (IS), сайт расщепления (CS), нисходящий элемент (DES) или их часть или комбинацию. Согласно

вариантам реализации спейсер может быть нацелен на элемент сплайсинга (SE) или цис-регуляторный элемент сплайсинга (SRE).

Каркасная последовательность представляет собой последовательность в онРНК, которая отвечает за связывание с нуклеазой (например, Cas9). Каркасная последовательность не включает спейсер/нацеливающую последовательность. Согласно вариантам реализации длина каркаса может составлять от примерно 1 до примерно 10, от примерно 10 до примерно 20, от примерно 20 до примерно 30, от примерно 30 до примерно 40, от примерно 40 до примерно 50, от примерно 50 до примерно 60, от примерно 60 до примерно 70, от примерно 70 до примерно 80, от примерно 80 до примерно 90, от примерно 90 до примерно 100, от примерно 100 до примерно 110, от примерно 110 до примерно 120 или от примерно 120 до примерно 130 нуклеотидов. Согласно вариантам реализации длина каркаса может составлять примерно 1, примерно 2, примерно 3, примерно 4, примерно 5, примерно 6, примерно 7, примерно 8, примерно 9, примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19, примерно 20, примерно 21, примерно 22, примерно 23, примерно 24, примерно 25, примерно 26, примерно 27, примерно 28, примерно 29, примерно 30, примерно 31, примерно 32, примерно 33, примерно 34, примерно 35, примерно 36, примерно 37, примерно 38, примерно 39, примерно 40, примерно 41, примерно 42, примерно 43, примерно 44, примерно 45, примерно 46, примерно 47, примерно 48, примерно 49, примерно 50, примерно 51, примерно 52, примерно 53, примерно 54, примерно 55, примерно 56, примерно 57, примерно 58, примерно 59, примерно 60, примерно 60, примерно 61, примерно 62, примерно 63, примерно 64, примерно 65, примерно 66, примерно 67, примерно 68, примерно 69, примерно 70, примерно 71, примерно 72, примерно 73, примерно 74, примерно 75, примерно 76, примерно 77, примерно 78, примерно 79, примерно 80, примерно 81, примерно 82, примерно 83, примерно 84, примерно 85, примерно 86, примерно 87, примерно 88, примерно 89, примерно 90, примерно 91, примерно 92, примерно 93, примерно 94, примерно 95, примерно 96, примерно 97, примерно 98, примерно 99, примерно 100, примерно 101, примерно 102, примерно 103, примерно 104, примерно 105, примерно 106, примерно 107, примерно 108, примерно 109, примерно 110, примерно 111, примерно 112, примерно 113, примерно 114, примерно 115, примерно 116, примерно 117, примерно 118, примерно 119, примерно 120, примерно 121, примерно 122, примерно 123, примерно 124 или примерно 125 нуклеотидов. Согласно вариантам реализации каркас может иметь длину по меньшей мере 10, по меньшей мере 20, по меньшей мере 30, по меньшей мере 40, по меньшей мере 50, по меньшей мере 60, по меньшей мере 70, по меньшей мере 80, по меньшей мере 90, по меньшей мере 100, по меньшей мере 110, по меньшей мере 120 или по меньшей мере 125 нуклеотидов.

Согласно вариантам реализации нРНК представляет собой двухмолекулярную направляющую РНК, например, crРНК и tracrРНК. Согласно вариантам реализации нРНК может дополнительно включать поли(А)-хвост.



Согласно вариантам реализации соединение, которое включает СРР, конъюгировано с нуклеиновой кислотой, которая включает нРНК. Согласно вариантам реализации нуклеиновая кислота включает примерно 1, примерно 2, примерно 3, примерно 4, примерно 5, примерно 6, примерно 7, примерно 8, примерно 9, примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18, примерно 19 или примерно 20 нРНК. Согласно вариантам реализации нРНК распознают одну и ту же мишень. Согласно вариантам реализации нРНК распознают различные мишени. Согласно вариантам реализации нуклеиновая кислота, которая включает нРНК, включает последовательность, кодирующую промотор, причем указанный промотор управляет экспрессией нРНК.

#### *Нуклеаза*

Согласно вариантам реализации соединения включают проникающий в клетку пептид, конъюгированный с нуклеазой. Согласно вариантам реализации нуклеаза представляет собой нуклеазу типа II, типа V-A, типа V-B, типа VC, типа V-U, типа VI-B. Согласно вариантам реализации нуклеаза представляет собой нуклеазу на основе эффектора, подобного активаторам транскрипции (TALEN), мегануклеазу или нуклеазу с «цинковым пальцем». Согласно вариантам реализации нуклеаза представляет собой нуклеазу Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B-подобную, Cas13a (C2c2), Cas13b или Cas14. Например, в некоторых вариантах реализации нуклеаза представляет собой нуклеазу Cas9 или нуклеазу Cpf1.

Согласно вариантам реализации нуклеаза представляет собой модифицированную форму или вариант нуклеазы Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B-подобной, Cas13a (C2c2), Cas13b или Cas14. Согласно вариантам реализации нуклеаза представляет собой модифицированную форму или вариант нуклеазы TAL, мегануклеазы или нуклеазы с «цинковым пальцем». «Модифицированная» или «вариантная» нуклеаза представляет собой нуклеазу, которая является, например, усеченной, слитой с другим белком (таким как другая нуклеаза), каталитически инактивированной и т. д. Согласно вариантам реализации нуклеаза может иметь по меньшей мере примерно 80%, по меньшей мере примерно 85%, по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 98%, по меньшей мере примерно 99% или примерно 100% идентичности последовательности с встречающейся в природе нуклеазой Cas9, Cas12a (Cpf1), Cas12b, Cas12c, Tnp-B-подобной, Cas13a (C2c2), Cas13b, Cas14 или TALEN, мегануклеазой или нуклеазой с «цинковым пальцем». Согласно вариантам реализации нуклеаза представляет собой нуклеазу Cas9, полученную из *S. pyogenes* (SpCas9). Согласно вариантам реализации нуклеаза имеет по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичности последовательности с нуклеазой Cas9, полученной из *S. pyogenes* (SpCas9). Согласно вариантам реализации нуклеаза представляет собой Cas9, полученную из *S. aureus* (SaCas9). Согласно вариантам реализации нуклеаза имеет по меньшей мере

примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичности последовательности с Cas9, полученной из *S. aureus* (SaCas9). Согласно вариантам реализации Cpf1 представляет собой фермент Cpf1 из *Acidaminococcus* (вид BV3L6, номер доступа UniProt U2UMQ6). Согласно вариантам реализации нуклеаза имеет по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичности последовательности с ферментом Cpf1 из *Acidaminococcus* (вид BV3L6, номер доступа UniProt U2UMQ6).

Согласно вариантам реализации Cpf1 представляет собой фермент Cpf1 из *Lachnospiraceae* (вид ND2006, номер доступа UniProt A0A182DWE3). Согласно вариантам реализации нуклеаза имеет по меньшей мере примерно 90%, по меньшей мере примерно 95%, по меньшей мере примерно 96%, по меньшей мере примерно 97%, по меньшей мере примерно 98% или по меньшей мере примерно 99% идентичности последовательности с ферментом Cpf1 из *Lachnospiraceae*. Согласно вариантам реализации последовательность, кодирующая нуклеазу, оптимизирована по кодонам для экспрессии в клетках млекопитающих. Согласно вариантам реализации последовательность, кодирующая нуклеазу, оптимизирована по кодонам для экспрессии в клетках человека или клетках мыши.

Согласно вариантам реализации соединение, которое включает CPP, конъюгировано с нуклеазой. Согласно вариантам реализации нуклеаза представляет собой растворимый белок.

Согласно вариантам реализации соединение, которое включает CPP, конъюгировано с нуклеиновой кислотой, кодирующей нуклеазу. Согласно вариантам реализации нуклеиновая кислота, кодирующая нуклеазу, включает последовательность, кодирующую промотор, причем указанный промотор управляет экспрессией нуклеазы.

#### *Комбинации нРНК и нуклеазы*

Согласно вариантам реализации соединения включают один или более CPP (или сCPP), конъюгированных с нРНК и нуклеазой. Согласно вариантам реализации один или более CPP (или сCPP) конъюгированы с нуклеиновой кислотой, кодирующей нРНК и/или нуклеазу. Согласно вариантам реализации нуклеиновая кислота, кодирующая нуклеазу и нРНК, включает последовательность, кодирующую промотор, причем указанный промотор управляет экспрессией нуклеазы и нРНК. Согласно вариантам реализации нуклеиновая кислота, кодирующая нуклеазу и нРНК, включает два промотора, причем первый промотор контролирует экспрессию нуклеазы, и второй промотор контролирует экспрессию нРНК. Согласно вариантам реализации нуклеиновая кислота, кодирующая нРНК и нуклеазу, кодирует от примерно 1 до примерно 20 нРНК или от примерно 1, примерно 2, примерно 3, примерно 4, примерно 5, примерно 6, примерно 7, примерно 8, примерно 9, примерно 10, примерно 11, примерно 12, примерно 13, примерно 14, примерно 15, примерно 16, примерно 17, примерно 18 или примерно 19 и до примерно 20

нРНК. Согласно вариантам реализации нРНК распознают различные мишени. Согласно вариантам реализации нРНК распознают одну и ту же мишень.

Согласно вариантам реализации соединения включают проникающий в клетку пептид (или сСРР), конъюгированный с рибонуклеопротеином (РНП), который включает нРНК и нуклеазу.

Согласно вариантам реализации композицию, которая включает: (а) СРР, конъюгированный с нРНК, и (b) нуклеазу, доставляют в клетку. Согласно вариантам реализации композицию, которая включает: (а) СРР, конъюгированный с нуклеазой, и (b) нРНК, доставляют в клетку.

Согласно вариантам реализации композицию, которая включает: (а) первый СРР, конъюгированный с нРНК, и (b) второй СРР, конъюгированный с нуклеазой, доставляют в клетку. Согласно вариантам реализации первый СРР и второй СРР являются одинаковыми. Согласно вариантам реализации первый СРР и второй СРР отличаются.

#### *Представляющий интерес генетический элемент*

Согласно вариантам реализации соединения, раскрытые в настоящем документе, включают проникающий в клетку пептид, конъюгированный с представляющим интерес генетическим элементом. Согласно вариантам реализации представляющий интерес генетический элемент заменяет последовательность геномной ДНК, расщепленную нуклеазой. Неограничивающие примеры представляющих интерес генетических элементов включают гены, однонуклеотидный полиморфизм, промотор или терминаторы.

#### *Ингибиторы нуклеаз*

Согласно вариантам реализации соединения, раскрытые в настоящем документе, включают проникающий в клетку пептид, конъюгированный с ингибитором нуклеазы (например, Cas9). Возможное нецелевое редактирование является ограничением редактирования генов. Доставка ингибитора нуклеазы будет ограничивать нецелевое редактирование. Согласно вариантам реализации ингибитор нуклеазы представляет собой полипептид, полинуклеотид или малую молекулу. Примерные ингибиторы нуклеаз описаны в публикации США № 2020/087354, международной публикации № 2018/085288, публикации США № 2018/0382741, международной публикации № 2019/089761, международной публикации № 2020/068304, международной публикации № 2020/041384 и международной публикации № 2019/076651, каждая из которых полностью включена в настоящий документ посредством ссылки.

#### *Терапевтические полипептиды*

Согласно вариантам реализации терапевтический фрагмент включает полипептид. Согласно вариантам реализации терапевтический фрагмент включает белок или его фрагмент. Согласно вариантам реализации терапевтический фрагмент включает РНК-связывающий белок или его РНК-связывающий фрагмент. Согласно вариантам реализации терапевтический фрагмент включает фермент. Согласно вариантам реализации терапевтический фрагмент включает РНК-расщепляющий фермент или его активный фрагмент.

### *Конъюгатные группы*

Согласно вариантам реализации АС модифицированы путем ковалентного присоединения одной или более конъюгатных групп. В целом, конъюгатные группы модифицируют одно или более свойств присоединенного АС, включая, но не ограничиваясь перечисленными, фармакодинамику, фармакокинетику, связывание, всасывание, клеточное распределение, клеточное поглощение, заряд и клиренс. Конъюгатные группы рутинно используют в области химии и связывают непосредственно или посредством необязательного линкерного фрагмента или линкерной группы с исходным соединением, таким как АС. Конъюгатные группы включают, но не ограничиваются перечисленными, интеркаляторы, репортерные молекулы, полиамины, полиамиды, полиэтиленгликоли, тиоэфиры, полиэфиры, холестерин, тиохолестерин, фрагменты холевой кислоты, фолат, липиды, фосфолипиды, биотин, феназин, фенантридин, антрахинон, адамантан, акридин, флуоресцеины, родамины, кумарины и красители. Согласно вариантам реализации конъюгатная группа представляет собой полиэтиленгликоль (PEG), и PEG конъюгирован либо с АС, либо с CPP.

Конъюгатные группы включают липидные фрагменты, такие как холестерин (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553); холевую кислоту (Manoharan et al., Bioorg. Med. Chem. Lett., 1994, 4, 1053); простой тиоэфир, например, гексил-S-тримитилтиол (Manoharan et al., Ann. N.Y. Acad. Sci., 1992, 660, 306; Manoharan et al., Bioorg. Med. Chem. Lett., 1993, 3, 2765); тиохолестерин (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533); алифатическую цепь, например, додекандиоловый или ундецильный остатки (Saison-Behmoaras et al., EMBO J., 1991, 10, 111; Kabanov et al., FEBS Lett., 1990, 259, 327; Svinarchuk et al., Biochimie, 1993, 75, 49); фосфолипид, например, дигексадецил-рац-глицерин или триэтиламмоний 1,2-ди-О-гексадецил-рац-глицеро-3-Н-фосфонат (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651; Shea et al., Nucl. Acids Res., 1990, 18, 3777); цепь полиамина или полиэтиленгликоля (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969); адамантануксусную кислоту (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651); пальмитильный фрагмент (Mishra et al., Biochim. Biophys. Acta., 1995, 1264, 229); или фрагмент октадециламина или гексиламинокарбонилхолестерина (Crooke et al., J. Pharmacol. Exp. Ther., 1996, 277, 923).

Линкерные группы или бифункциональные линкерные фрагменты, такие как те, которые известны в данной области техники, пригодны для соединений, предложенных в настоящем документе. Линкерные группы можно применять для присоединения химических функциональных групп, конъюгатных групп, репортерных групп и других групп к селективным сайтам в исходном соединении, таком как, например, АС. В целом, бифункциональный линкерный фрагмент содержит гидрокарбильный фрагмент, имеющий две функциональные группы. Одна из функциональных групп выбрана для связывания с исходной молекулой или представляющим интерес соединением, а другая выбрана для связывания по существу любой выбранной группы, такой как химическая функциональная группа или конъюгатная группа. Можно применять любой из линкеров,

описанных в настоящем документе. Согласно вариантам реализации линкер содержит цепную структуру или олигомер из повторяющихся звеньев, таких как этиленгликолевые или аминокислотные звенья. Примеры функциональных групп, которые рутинно применяют в бифункциональном линкерном фрагменте, включают, но не ограничиваются перечисленными, электрофилы для взаимодействия с нуклеофильными группами и нуклеофилы для взаимодействия с электрофильными группами. Согласно вариантам реализации бифункциональные линкерные фрагменты включают amino, гидроксил, карбоновую кислоту, тиол, ненасыщенность (например, двойные или тройные связи) и т. п. Некоторые неограничивающие примеры бифункциональных линкерных фрагментов включают 8-амино-3,6-диоксооктановую кислоту (ADO), сукцинимидил-4-(N-малеимидометил)циклогексан-1-карбоксилат (SMCC) и 6-аминогексановую кислоту (АНЕХ или АНА). Другие линкерные группы включают, но не ограничиваются перечисленными, замещенный C1-C10 алкил, замещенный или незамещенный C2-C10 алкенил или замещенный или незамещенный C2-C10 алкинил, причем неограничивающий перечень замещающих групп включает гидроксил, amino, алкокси, карбокси, бензил, фенил, нитро, тиол, тиаалкокси, галоген, алкил, арил, алкенил и алкинил.

Согласно вариантам реализации АС может быть связан с 10 аргинин-сериновым дипептидными повторами. АС, связанные с 10 аргинин-сериновыми дипептидными повторами для искусственного привлечения энхансерных факторов сплайсинга, применяли *in vitro* для индукции включения мутантных экзонов BRCA1 и SMN2, которые в противном случае были бы пропущены. См. Cartegni and Krainer 2003, которая включена в настоящий документ посредством ссылки.

#### **Носители для ускользания из эндосом (EEV)**

Носитель для ускользания из эндосом (EEV) можно применять для транспортировки груза через клеточную мембрану, например, для доставки груза в цитозоль или ядро клетки. Груз может включать терапевтический фрагмент (ТМ). EEV может содержать проникающий в клетку пептид (CPP), например, проникающий в клетку циклический пептид (сCPP). Согласно вариантам реализации EEV содержит сCPP, который конъюгирован с экзоциклическим пептидом (EP). EP может взаимозаменяемо называться модулирующим пептидом (MP). EP может содержать последовательность сигнала ядерной локализации (NLS). EP может быть связан с грузом. EP может быть связан с сCPP. EP может быть связан с грузом и сCPP. Связь между EP, грузом, сCPP или их комбинациями может быть нековалентной или ковалентной. EP может быть присоединен пептидной связью к N-концу сCPP. EP может быть присоединен пептидной связью к C-концу сCPP. EP может быть присоединен к сCPP посредством боковой цепи аминокислоты в сCPP. EP может быть присоединен к сCPP посредством боковой цепи лизина, которая может быть конъюгирована с боковой цепью глутамина в сCPP. EP может быть конъюгирован с 5'- или 3'-концом олигонуклеотидного груза. EP может быть связан с линкером. Экзоциклический пептид может быть конъюгирован с аминогруппой линкера. EP может быть связан с линкером через C-конец EP и сCPP через боковую цепь на сCPP

и/или EP. Например, EP может содержать концевой лизин, который затем может быть связан с cSPР, содержащим глутамин, амидной связью. Если EP содержит концевой лизин, и боковую цепь лизина можно использовать для присоединения cSPР, C- или N-конец может быть присоединен к линкеру на грузе.

#### *Экзоциклические пептиды*

Экзоциклический пептид (EP) может содержать от 2 до 10 аминокислотных остатков, например, 2, 3, 4, 5, 6, 7, 8, 9 или 10 аминокислотных остатков, включая все диапазоны и значения между ними. EP может содержать от 6 до 9 аминокислотных остатков. EP может содержать от 4 до 8 аминокислотных остатков.

Каждая аминокислота в экзоциклическом пептиде может представлять собой природную или неприродную аминокислоту. Термин «неприродная аминокислота» относится к органическому соединению, которое является аналогом природной аминокислоты в том смысле, что оно имеет структуру, подобную природной аминокислоте, соответственно, оно имитирует структуру и реакционную способность природной аминокислоты. Неприродная аминокислота может представлять собой модифицированную аминокислоту и/или аналог аминокислоты, который не является одной из 20 распространенных встречающихся в природе аминокислот или редких природных аминокислот селеноцистеина или пирролизина. Неприродные аминокислоты также могут представлять собой D-изомер природных аминокислот. Примеры подходящих аминокислот включают, но не ограничиваются перечисленными, аланин, аллозолейцин, аргинин, цитруллин, аспарагин, аспарагиновую кислоту, цистеин, глутамин, глутаминовую кислоту, глицин, гистидин, изолейцин, лейцин, лизин, метионин, нафтилаланин, фенилаланин, пролин, пироглутаминовую кислоту, серин, треонин, триптофан, тирозин, валин, их производное или их комбинации. Эти и другие аминокислоты перечислены в **Таблице 4** вместе с их сокращениями, используемыми в настоящем документе. Например, аминокислоты могут представлять собой A, G, P, K, R, V, F, H, NaI или цитруллин.

EP может содержать по меньшей мере один положительно заряженный аминокислотный остаток, например, по меньшей мере один лизиновый остаток и/или по меньшей мере один аминокислотный остаток, содержащий боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму. EP может содержать 1 или 2 аминокислотных остатка, содержащих боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму. Аминокислотный остаток, содержащий боковую цепь, содержащую гуанидиновую группу, может представлять собой остаток аргинина. Протонированные формы могут означать их соль по всему раскрытию.

EP может содержать по меньшей мере два, по меньшей мере три или по меньшей мере четыре или более остатков лизина. EP может содержать 2, 3 или 4 остатка лизина. Аминогруппа на боковой цепи каждого остатка лизина может быть замещена защитной группой, включая, например, трифторацетильную (-COCF<sub>3</sub>), аллилоксикарбонильную (Alloc), 1-(4,4-диметил-2,6-диоксоциклогексиден)этильную (Dde) или (4,4-диметил-2,6-

диоксоциклогекс-1-илиден-3)-метилбутильную (ivDde) группу. Аминогруппа на боковой цепи каждого остатка лизина может быть замещена трифторацетильной (-COCF<sub>3</sub>) группой. Защитная группа может быть включена для осуществления амидной конъюгации. Защитную группу можно удалить после конъюгации EP с cCPP.

EP может содержать по меньшей мере 2 аминокислотных остатка с гидрофобной боковой цепью. Аминокислотный остаток с гидрофобной боковой цепью может быть выбран из валина, пролина, аланина, лейцина, изолейцина и метионина. Аминокислотный остаток с гидрофобной боковой цепью может представлять собой валин или пролин.

EP может содержать по меньшей мере один положительно заряженный аминокислотный остаток, например, по меньшей мере один остаток лизина и/или по меньшей мере один остаток аргинина. EP может содержать по меньшей мере два, по меньшей мере три или по меньшей мере четыре или более остатков лизина и/или остатков аргинина.

EP может содержать 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), KBKBK (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), 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) или PKKKRKG (SEQ ID NO:48), где В представляет собой бета-аланин. Аминокислоты в EP могут иметь D- или L-стереохимию.

EP может содержать 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), KBKBK (SEQ ID NO:24), KKKRKV (SEQ 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) или PKKKRKG (SEQ ID NO:48). EP может содержать PKKKRKV (SEQ ID NO:42), RR, RRR, RHR, RBR, RBRBR (SEQ ID NO:49), RBHBR (SEQ

ID NO:50) или HBRBH (SEQ ID NO:51), где В представляет собой бета-аланин. Аминокислоты в EP могут иметь D- или L-стереохимию.

EP может состоять из 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), KBKBK (SEQ ID NO:24), KKKRKV (SEQ ID NO:33), PKKKRKV (SEQ ID NO:42), PGKKRKV (SEQ ID NO:Z43), PKGKRKV (SEQ ID NO:Z44), PKKGRKV (SEQ ID NO:Z45), PKKKGKV (SEQ ID NO:46), PKKRGV (SEQ ID NO:47) или PKKKRKG (SEQ ID NO:48). EP может состоять из PKKKRKV (SEQ ID NO:42), RR, RRR, RHR, RBR, RBRBR (SEQ ID NO:49), RBHBR (SEQ ID NO:50) или HBRBH (SEQ ID NO:51), где В представляет собой бета-аланин. Аминокислоты в EP могут иметь D- или L-стереохимию.

EP может содержать аминокислотную последовательность, идентифицированную в данной области техники как последовательность ядерной локализации (NLS). EP может состоять из аминокислотной последовательности, идентифицированной в данной области техники как последовательность ядерной локализации (NLS). EP может содержать NLS, содержащую аминокислотную последовательность PKKKRKV (SEQ ID NO:42). EP может состоять из NLS, содержащей аминокислотную последовательность PKKKRKV (SEQ ID NO:42). EP может содержать NLS, содержащую аминокислотную последовательность, выбранную из NLSKRPAAIKKAGQAKKKK (SEQ ID NO:52), PAAKRVKLD (SEQ ID NO:53), RQRRNELKRSF (SEQ ID NO:54), RMRKFKNKGKDTAELRRRRVEVSVELR (SEQ ID NO:Z55), KAKKDEQILKRRNV (SEQ ID NO:56), VSRKRPRP (SEQ ID NO:57), PPKKARED (SEQ ID NO:58), PQPKKKPL (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) и RKCLQAGMNLEARKTKK (SEQ ID NO:66). EP может состоять из NLS, содержащей аминокислотную последовательность, выбранную из 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), PQPKKKPL (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) и RKCLQAGMNLEARKTKK (SEQ ID NO:66).

Все экзоциклические последовательности также могут содержать N-концевую ацетильную группу. Следовательно, например, EP может иметь структуру: Ac-PKKKRKV (SEQ ID NO:42).

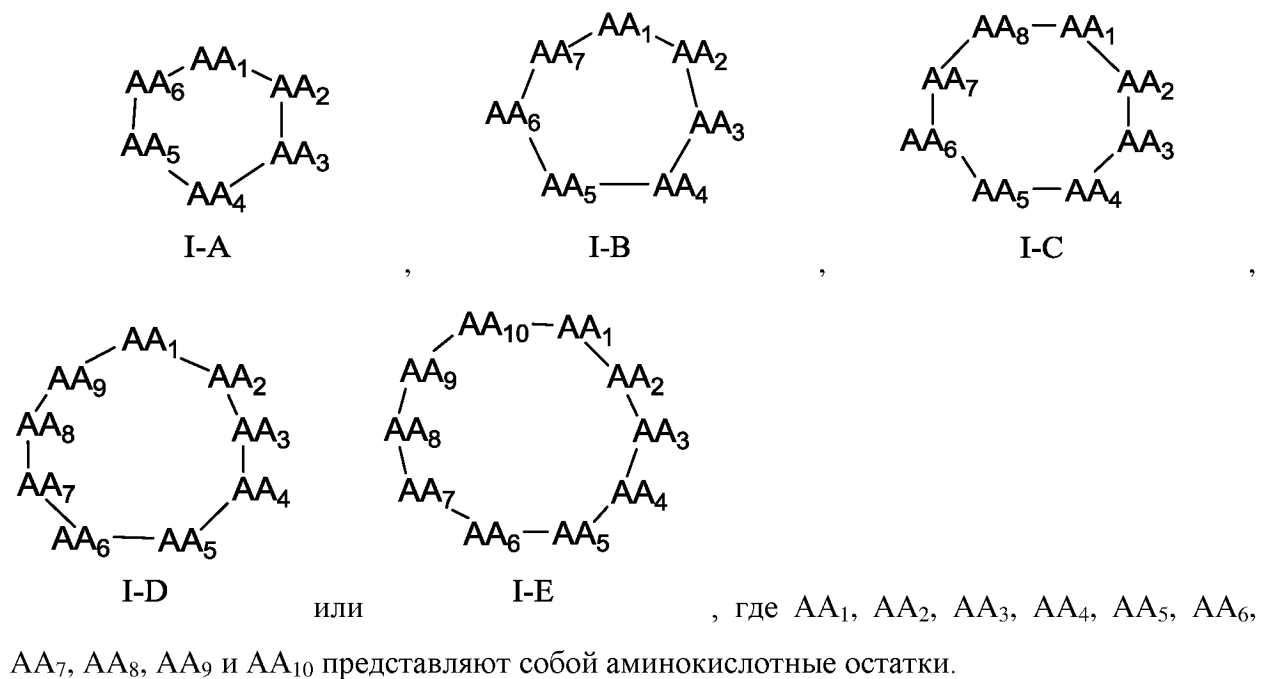
### **Проникающие в клетку пептиды (CPP)**

Проникающий в клетку пептид (CPP) может содержать от 6 до 20 аминокислотных остатков. Проникающий в клетку пептид может представлять собой проникающий в



клетку циклический пептид (сСРР). сСРР способен проникать через клеточную мембрану. Экзоциклический пептид (EP) может быть конъюгирован с сСРР, и полученная конструкция может называться носителем для ускользания из эндосом (EEV). сСРР может направлять груз (например, терапевтический фрагмент (ТМ), такой как олигонуклеотид, пептид или малая молекула) для проникновения через мембрану клетки. сСРР может доставлять груз в цитозоль клетки. сСРР может доставлять груз в клеточное местоположение, где расположена мишень (например, пре-мРНК). Для конъюгации сСРР с грузом (например, пептидом, олигонуклеотидом или малой молекулой) по меньшей мере одну связь или одиночную пару электронов на сСРР можно заменить.

Общее число аминокислотных остатков в сСРР находится в пределах от 6 до 20 аминокислотных остатков, например, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 аминокислотных остатков, включая все диапазоны и поддиапазоны между ними. сСРР может содержать от 6 до 13 аминокислотных остатков. сСРР, раскрытый в настоящем документе, может содержать от 6 до 10 аминокислот. Например, сСРР, содержащий 6-10 аминокислотных остатков, может иметь структуру в соответствии с любой из Формул от I-A до I-E:



сСРР может содержать от 6 до 8 аминокислот. сСРР может содержать 8 аминокислот.

Каждая аминокислота в сСРР может представлять собой природную или неприродную аминокислоту. Термин «неприродная аминокислота» относится к органическому соединению, которое является аналогом природной аминокислоты в том смысле, что оно имеет структуру, подобную природной аминокислоте, соответственно, оно имитирует структуру и реакционную способность природной аминокислоты. Неприродная аминокислота может представлять собой модифицированную аминокислоту и/или аналог аминокислоты, который не является одной из 20 распространенных

встречающихся в природе аминокислот или редких природных аминокислот селеноцистеина или пирролизина. Неприродные аминокислоты также могут представлять собой D-изомер природной аминокислоты. Примеры подходящих аминокислот включают, но не ограничиваются перечисленными, аланин, аллозолейцин, аргинин, цитруллин, аспарагин, аспарагиновую кислоту, цистеин, глутамин, глутаминовую кислоту, глицин, гистидин, изолейцин, лейцин, лизин, метионин, нафтилаланин, фенилаланин, пролин, пироглутаминовую кислоту, серин, треонин, триптофан, тирозин, валин, их производное или их комбинации. Эти и другие аминокислоты перечислены в **Таблице 5** вместе с их сокращениями, используемыми в настоящем документе.

Таблица 4. Сокращения аминокислот

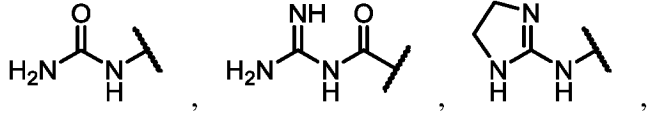
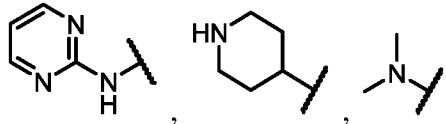
Аминокислота	Сокращения*	
	L-аминокислота	D-аминокислота
2-[2-[2-аминоэтокси]этокси]уксусная кислота	AEEA, миниPEG, PEG2	H/П
Аланин	Ala (A)	ala (a)
Аллоизолейцин	Aile	Aile
Аргинин	Arg (R)	arg (r)
Аспарагин	Asn (N)	asn (n)
аспарагиновая кислота	Asp (D)	asp (d)
Цистеин	Cys (C)	cys (c)
Цитруллин	Cit	Cit
Циклогексилаланин	Cha	cha
2,3-диаминопропионовая кислота	Dap	dap
4-фторфенилаланин	Fpa ( $\Sigma$ )	pfa
глутаминовая кислота	Glu (E)	glu (e)
глутамин	Gln (Q)	gln (q)
глицин	Gly (G)	gly (g)
гистидин	His (H)	his (h)
Гомопролин (также известный как пипеколиновая кислота)	Pip ( $\Theta$ )	pip ( $\theta$ )
изолейцин	Ile (I)	ile (i)
лейцин	Leu (L)	leu (l)
лизин	Lys (K)	lys (k)
метионин	Met (M)	met (m)
3-(2-нафтил)-аланин	Nal ( $\Phi$ )	nal ( $\phi$ )

Аминокислота	Сокращения*	
	L-аминокислота	D-аминокислота
3-(1-нафтил)-аланин	1-Nal	1-nal
норлейцин	Nle ( $\Omega$ )	nle
фенилаланин	Phe (F)	phe (f)
фенилглицин	Phg ( $\Psi$ )	phg
4-(фосфонодифторметил)фенилаланин	F <sub>2</sub> Pmp ( $\Lambda$ )	f <sub>2</sub> pmp
пролин	Pro (P)	pro (p)
саркозин	Sar ( $\Xi$ )	sar
селеноцистеин	Sec (U)	sec (u)
серин	Ser (S)	ser (s)
треонин	Thr (T)	thr (y)
тирозин	Tyr (Y)	tyr (y)
триптофан	Trp (W)	trp (w)
валин	Val (V)	val (v)
Трет-бутилаланин	Tle	tle
Пеницилламин	Pen	Pen
Гомоаргинин	ГомоArg	гомоarg
Никотиниллизин	Lys(NIC)	lys(NIC)
Трифторацетиллизин	Lys(TFA)	lys(TFA)
Метиллейцин	MeLeu	meLeu
3-(3-бензотиенил)-аланин	Bta	bta
*однобуквенные сокращения: заглавные буквы указывают на форму L-аминокислоты, строчная буква указывает на форму D-аминокислоты.		

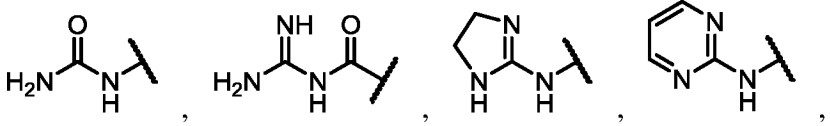
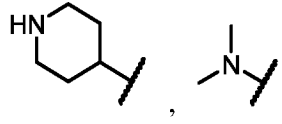
В настоящем документе «полиэтиленгликоль» и «PEG» используются взаимозаменяемо. «PEG<sub>m</sub>» и «PEG<sub>n</sub>» представляют собой или получены из молекулы формулы  $\text{HO}(\text{CO})-(\text{CH}_2)_n-(\text{OCH}_2\text{CH}_2)_m-\text{NH}_2$ , где  $n$  представляет собой любое целое число от 1 до 5, и  $m$  представляет собой любое целое число от 1 до 23. Согласно вариантам реализации  $n$  представляет собой 1 или 2. Согласно вариантам реализации  $n$  представляет собой 1. Согласно вариантам реализации  $n$  представляет собой 2. Согласно вариантам реализации  $n$  представляет собой 1 и  $m$  представляет собой 2. Согласно вариантам реализации  $n$  представляет собой 2 и  $m$  представляет собой 2. Согласно вариантам реализации  $n$  представляет собой 1 и  $m$  представляет собой 4. Согласно вариантам реализации  $n$  представляет собой 2 и  $m$  представляет собой 4. Согласно вариантам реализации  $n$  представляет собой 1 и  $m$  представляет собой 12. Согласно вариантам реализации  $n$  представляет собой 2 и  $m$  представляет собой 12.

В настоящем документе «миниРЕG<sub>m</sub>» или «миниРЕG<sub>m</sub>» представляют собой или получены из молекулы Формулы  $\text{HO}(\text{CO})-(\text{CH}_2)_n-(\text{OCH}_2\text{CH}_2)_m-\text{NH}_2$ , где  $n$  представляет собой 1, и  $m$  представляет собой любое целое число от 1 до 23. Например, «миниРЕG<sub>2</sub>» или «миниРЕG<sub>2</sub>» представляет собой или получен из (2-[2-[2-аминоэтокси]этокси]уксусной кислоты), и «миниРЕG<sub>4</sub>» или «миниРЕG<sub>4</sub>» представляет собой или получен из  $\text{HO}(\text{CO})-(\text{CH}_2)_n-(\text{OCH}_2\text{CH}_2)_m-\text{NH}_2$ , где  $n$  представляет собой 1 и  $m$  представляет собой 4.

сСРР может содержать от 4 до 20 аминокислот, причем: (i) по меньшей мере одна аминокислота имеет боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму; (ii) по меньшей мере одна аминокислота не имеет боковой цепи

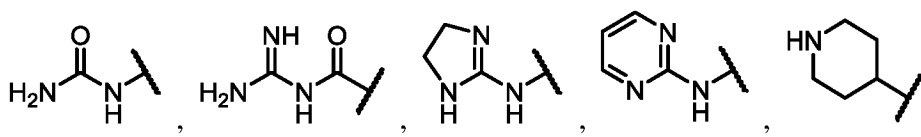
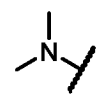
или имеет боковую цепь, содержащую , , или ее протонированную форму; и (iii) по меньшей мере две аминокислоты независимо имеют боковую цепь, содержащую ароматическую или гетероароматическую группу.

По меньшей мере две аминокислоты могут не иметь боковой цепи или могут иметь

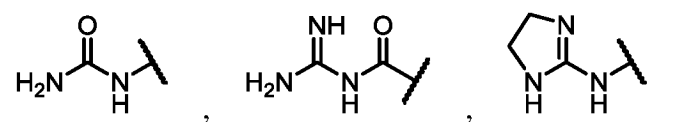

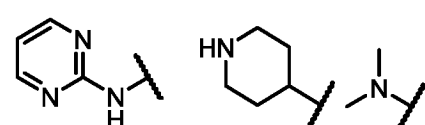
боковую цепь, содержащую , , или ее протонированную форму. В настоящем документе, если боковая цепь отсутствует, аминокислота содержит два атома водорода на атоме(атомах) углерода (например,  $-\text{CH}_2-$ ), связывающих амин и карбоновую кислоту.

Аминокислота, не имеющая боковой цепи, может представлять собой глицин или  $\beta$ -аланин.

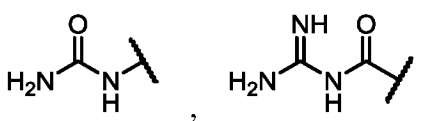
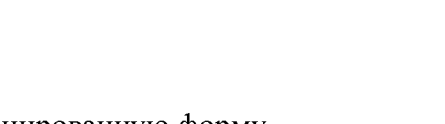
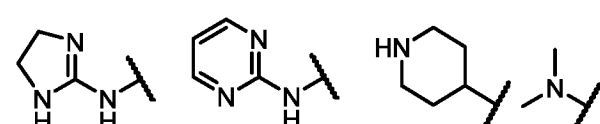
сСРР может содержать от 6 до 20 аминокислотных остатков, которые образуют сСРР, причем: (i) по меньшей мере одна аминокислота может представлять собой остатки глицина,  $\beta$ -аланина или 4-аминомасляной кислоты; (ii) по меньшей мере одна аминокислота может иметь боковую цепь, содержащую арильную или гетероарильную группу; и (iii) по меньшей мере одна аминокислота имеет боковую цепь, содержащую

гуанидиновую группу,  , или ее протонированную форму. 

cCPP может содержать от 6 до 20 аминокислотных остатков, которые образуют cCPP, причем: (i) по меньшей мере две аминокислоты могут независимо представлять собой остатки глицина,  $\beta$ -аланина или 4-аминомасляной кислоты; (ii) по меньшей мере одна аминокислота может иметь боковую цепь, содержащую арильную или гетероарильную группу; и (iii) по меньшей мере одна аминокислота имеет боковую цепь,

содержащую гуанидиновую группу,  ,  , или ее протонированную форму. 

cCPP может содержать от 6 до 20 аминокислотных остатков, которые образуют cCPP, причем: (i) по меньшей мере три аминокислоты могут независимо представлять собой остатки глицина,  $\beta$ -аланина или 4-аминомасляной кислоты; (ii) по меньшей мере одна аминокислота может иметь боковую цепь, содержащую ароматическую или гетероароматическую группу; и (iii) по меньшей мере одна аминокислота может иметь

боковую цепь, содержащую гуанидиновую группу,  ,  ,  , или ее протонированную форму.

#### *Глицин и родственные аминокислотные остатки*

cCPP может содержать (i) 1, 2, 3, 4, 5 или 6 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 2 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 3 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 4 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 5 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 6 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может содержать (i) 3, 4 или 5 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. cCPP может

содержать (i) 3 или 4 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации.

сСРР может содержать (i) 1, 2, 3, 4, 5 или 6 остатков глицина. сСРР может содержать (i) 2 остатка глицина. сСРР может содержать (i) 3 остатка глицина. сСРР может содержать (i) 4 остатка глицина. сСРР может содержать (i) 5 остатков глицина. сСРР может содержать (i) 6 остатков глицина. сСРР может содержать (i) 3, 4 или 5 остатков глицина. сСРР может содержать (i) 3 или 4 остатка глицина. сСРР может содержать (i) 2 или 3 остатка глицина. сСРР может содержать (i) 1 или 2 остатка глицина.

сСРР может содержать (i) 3, 4, 5 или 6 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 3 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 4 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 5 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 6 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 3, 4 или 5 остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации. сСРР может содержать (i) 3 или 4 остатка глицина,  $\beta$ -аланина, 4-аминомасляной кислоты или их комбинации.

сСРР может содержать по меньшей мере три остатка глицина. сСРР может содержать (i) 3, 4, 5 или 6 остатков глицина. сСРР может содержать (i) 3 остатка глицина. сСРР может содержать (i) 4 остатка глицина. сСРР может содержать (i) 5 остатков глицина. сСРР может содержать (i) 6 остатков глицина. сСРР может содержать (i) 3, 4 или 5 остатков глицина. сСРР может содержать (i) 3 или 4 остатка глицина.

Согласно вариантам реализации никакие из остатков глицина,  $\beta$ -аланина или 4-аминомасляной кислоты в сСРР не являются смежными. Два или три остатка глицина,  $\beta$ -аланина или 4-аминомасляной кислоты могут быть смежными. Два остатка глицина,  $\beta$ -аланина или 4-аминомасляной кислоты могут быть смежными.

Согласно вариантам реализации никакие из остатков глицина в сСРР не являются смежными. Каждый остаток глицина в сСРР может быть отделен аминокислотным остатком, который не может быть глицином. Два или три остатка глицина могут быть смежными. Два остатка глицина могут быть смежными.

*Боковые цепи аминокислот с ароматической или гетероароматической группой*

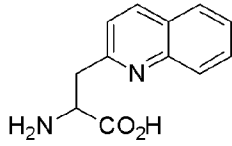
сСРР может содержать (ii) 2, 3, 4, 5 или 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 2 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 5 аминокислотных остатков, независимо имеющих боковую цепь, содержащую

ароматическую или гетероароматическую группу. сСРР может содержать (ii) 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 2, 3 или 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу. сСРР может содержать (ii) 2 или 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу.

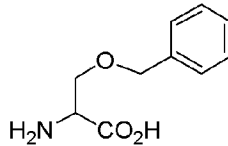
сСРР может содержать (ii) 2, 3, 4, 5 или 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 2 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 5 аминокислотных остатков, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 2, 3 или 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую группу. сСРР может содержать (ii) 2 или 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую ароматическую группу.

Ароматическая группа может представлять собой 6-14-членный арил. Арил может представлять собой фенил, нафтил или антраценил, каждый из которых необязательно замещен. Арил может представлять собой фенил или нафтил, каждый из которых необязательно замещен. Гетероароматическая группа может представлять собой 6-14-членный гетероарил, содержащий 1, 2 или 3 гетероатома, выбранных из N, O и S. Гетероарил может представлять собой пиридил, хинолил или изохинолил.

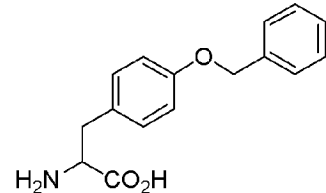
Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую или гетероароматическую группу, может независимо представлять собой бис(гомонафтилаланин), гомонафтилаланин, нафтилаланин, фенилглицин, бис(гомофенилаланин), гомофенилаланин, фенилаланин, триптофан, 3-(3-бензотиенил)-аланин, 3-(2-хинолил)-аланин, O-бензилсерин, 3-(4-(бензилокси)фенил)аланин, S-(4-метилбензил)цистеин, N-(нафталин-2-ил)глутамин, 3-(1,1'-бифенил-4-ил)-аланин, 3-(3-бензотиенил)-аланин или тирозин, каждый из которых необязательно замещен одним или более заместителями. Каждая аминокислота, имеющая боковую цепь, содержащую ароматическую или гетероароматическую группу, может быть независимо выбрана из:



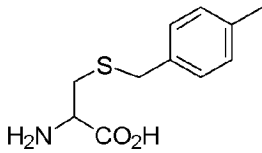
3-(2-хинолил)-аланин;



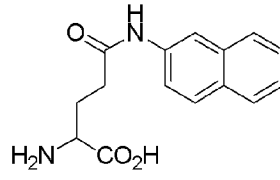
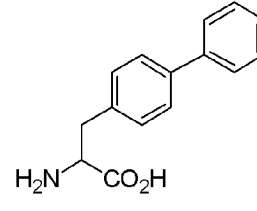
O-бензилсерин



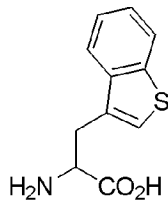
3-(4-(бензилокси)фенил)аланин,



S-(4-метилбензил)цистеин

N<sup>5</sup>-(нафталин-2-ил)глутамин

3-(1,1'-бифенил-4-ил)-аланин



3-(3-бензотиенил)-аланин,

где Н на N-конце и/или Н на С-конце заменены пептидной связью.

Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую или гетероароматическую группу, может независимо представлять собой остаток фенилаланина, нафтилаланина, фенилглицина, гомофенилаланина, гомонафтилаланина, бис(гомфенилаланина), бис-(гомонафтилаланина), триптофана или тирозина, каждый из которых необязательно замещен одним или более заместителями. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может независимо представлять собой остаток тирозина, фенилаланина, 1-нафтилаланина, 2-нафтилаланина, триптофана, 3-бензотиенилаланина, 4-фенилфенилаланина, 3,4-дифторфенилаланина, 4-трифторметилфенилаланина, 2,3,4,5,6-пентафторфенилаланина, гомофенилаланина,  $\beta$ -гомфенилаланина, 4-трет-бутилфенилаланина, 4-пиридинилаланина, 3-пиридинилаланина, 4-метилфенилаланина, 4-фторфенилаланина, 4-хлорфенилаланина, 3-(9-антрил)-аланина. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может независимо представлять собой остаток фенилаланина, нафтилаланина, фенилглицина, гомофенилаланина или гомонафтилаланина, каждый из которых необязательно замещен одним или более заместителями. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может независимо представлять собой остаток фенилаланина, нафтилаланина, гомофенилаланина, гомонафтилаланина, бис(гомонафтилаланина) или бис(гомонафтилаланина), каждый из которых необязательно замещен одним или более заместителями. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может независимо представлять



собой остаток фенилаланина или нафтилаланина, каждый из которых необязательно замещен одним или более заместителями. По меньшей мере один аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может представлять собой остаток фенилаланина. По меньшей мере два аминокислотных остатка, имеющих боковую цепь, содержащую ароматическую группу, могут представлять собой остатки фенилаланина. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу, может представлять собой остаток фенилаланина.

Согласно вариантам реализации никакие из аминокислот, имеющих боковую цепь, содержащую ароматическую или гетероароматическую группу, не являются смежными. Две аминокислоты, имеющие боковую цепь, содержащую ароматическую или гетероароматическую группу, могут быть смежными. Две смежные аминокислоты могут иметь противоположную стереохимию. Две смежные аминокислоты могут иметь одинаковую стереохимию. Три аминокислоты, имеющие боковую цепь, содержащую ароматическую или гетероароматическую группу, могут быть смежными. Три смежные аминокислоты могут иметь одинаковую стереохимию. Три смежные аминокислоты могут иметь чередующуюся стереохимию.

Аминокислотные остатки, содержащие ароматические или гетероароматические группы, могут представлять собой L-аминокислоты. Аминокислотные остатки, содержащие ароматические или гетероароматические группы, могут представлять собой D-аминокислоты. Аминокислотные остатки, содержащие ароматические или гетероароматические группы, могут представлять собой смесь D- и L-аминокислот.

Необязательный заместитель может представлять собой любой атом или группу, которая не снижает в значительной степени (например, более чем на 50%) эффективность цитозольной доставки сСРР, например, по сравнению с последовательностью, не содержащей заместителя, которая в остальном идентична. Необязательный заместитель может представлять собой гидрофобный заместитель или гидрофильный заместитель. Необязательный заместитель может представлять собой гидрофобный заместитель. Заместитель может увеличивать доступную для растворителя площадь поверхности (как определено в настоящем документе) гидрофобной аминокислоты. Заместитель может представлять собой галоген, алкил, алкенил, алкинил, циклоалкил, циклоалкенил, циклоалкинил, гетероциклил, арил, гетероарил, алкокси, арилокси, ацил, алкилкарбамоил, алкилкарбоксамидил, алкоксикарбонил, алкилтио или арилтио. Заместитель может представлять собой галоген.

Не ограничиваясь какой-либо теорией, полагают, что аминокислоты, имеющие ароматическую или гетероароматическую группу, имеющие более высокие показатели гидрофобности (т. е. аминокислоты, имеющие боковые цепи, содержащие ароматические или гетероароматические группы), могут улучшить эффективность цитозольной доставки сСРР по сравнению с аминокислотами, имеющими более низкий показатель гидрофобности. Каждая гидрофобная аминокислота может независимо иметь показатель

гидрофобности, превышающий показатель глицина. Каждая гидрофобная аминокислота может независимо представлять собой гидрофобную аминокислоту, имеющую показатель гидрофобности, превышающий показатель аланина. Каждая гидрофобная аминокислота может независимо иметь показатель гидрофобности, который превышает или равен таковому для фенилаланина. Гидрофобность может быть измерена с использованием шкал гидрофобности, известных в данной области техники. В **Таблице 5** перечислены показатели гидрофобности для различных аминокислот, сообщенные в работах Eisenberg and Weiss (Proc. Natl. Acad. Sci. U. S. A. 1984;81(1):140-144), Engleman, et al. (Ann. Rev. of Biophys. Biophys. Chem. 1986;1986(15):321-53), Kyte and Doolittle (J. Mol. Biol. 1982;157(1):105-132), Hoop and Woods (Proc. Natl. Acad. Sci. U. S. A. 1981;78(6):3824-3828), и Janin (Nature. 1979;277(5696):491-492), каждая из которых полностью включена в настоящий документ посредством ссылки. Гидрофобность можно измерить с использованием шкалы гидрофобности, сообщенной в Engleman, et al.

Таблица 5. Гидрофобность аминокислот

Аминокислота	Группа	Eisenberg and Weiss	Engleman et al.	Kyrie and Doolittle	Hoop and Woods	Janin
Ile	Неполярная	0,73	3,1	4,5	-1,8	0,7
Phe	Неполярная	0,61	3,7	2,8	-2,5	0,5
Val	Неполярная	0,54	2,6	4,2	-1,5	0,6
Leu	Неполярная	0,53	2,8	3,8	-1,8	0,5
Trp	Неполярная	0,37	1,9	-0,9	-3,4	0,3
Met	Неполярная	0,26	3,4	1,9	-1,3	0,4
Ala	Неполярная	0,25	1,6	1,8	-0,5	0,3
Gly	Неполярная	0,16	1,0	-0,4	0,0	0,3
Cys	Незаряженная/полярная	0,04	2,0	2,5	-1,0	0,9
Tyr	Незаряженная/полярная	0,02	-0,7	-1,3	-2,3	-0,4
Pro	Неполярная	-0,07	-0,2	-1,6	0,0	-0,3
Thr	Незаряженная/полярная	-0,18	1,2	-0,7	-0,4	-0,2
Ser	Незаряженная/полярная	-0,26	0,6	-0,8	0,3	-0,1
His	Заряженная	-0,40	-3,0	-3,2	-0,5	-0,1
Glu	Заряженная	-0,62	-8,2	-3,5	3,0	-0,7
Asn	Незаряженная/полярная	-0,64	-4,8	-3,5	0,2	-0,5
Gln	Незаряженная/полярная	-0,69	-4,1	-3,5	0,2	-0,7
Asp	Заряженная	-0,72	-9,2	-3,5	3,0	-0,6
Lys	Заряженная	-1,10	-8,8	-3,9	3,0	-1,8

Arg	Заряженная	-1,80	-12,3	-4,5	3,0	-1,4
-----	------------	-------	-------	------	-----	------

Размер ароматических или гетероароматических групп может быть выбран для улучшения эффективности цитозольной доставки сСРР. Не ограничиваясь какой-либо теорией, полагают, что более крупная ароматическая или гетероароматическая группа на боковой цепи аминокислоты может улучшить эффективность цитозольной доставки по сравнению с последовательностью, имеющей меньшую гидрофобную аминокислоту, но в остальном идентичной. Размер гидрофобной аминокислоты можно измерить с точки зрения молекулярной массы гидрофобной аминокислоты, стерических эффектов гидрофобной аминокислоты, доступной для растворителя площади поверхности (SASA) боковой цепи или их комбинаций. Размер гидрофобной аминокислоты можно измерить с точки зрения молекулярной массы гидрофобной аминокислоты, и гидрофобная аминокислота большего размера имеет боковую цепь с молекулярной массой по меньшей мере примерно 90 г/моль или по меньшей мере примерно 130 г/моль, или по меньшей мере примерно 141 г/моль. Размер аминокислоты может быть измерен с точки зрения SASA гидрофобной боковой цепи. Гидрофобная аминокислота может иметь боковую цепь с SASA, большей или равной таковой для аланина, или большей или равной таковой для глицина. Более крупные гидрофобные аминокислоты могут иметь боковую цепь с SASA, которая больше таковой для аланина или больше таковой для глицина. Гидрофобная аминокислота может иметь ароматическую или гетероароматическую группу с SASA, примерно большей или равной пиперидин-2-карбоновой кислоте, примерно большей или равной триптофану, примерно большей или равной фенилаланину или примерно большей или равной нафтилаланину. Первая гидрофобная аминокислота (AA<sub>H1</sub>) может иметь боковую цепь с SASA по меньшей мере примерно 200 Å<sup>2</sup>, по меньшей мере примерно 210 Å<sup>2</sup>, по меньшей мере примерно 220 Å<sup>2</sup>, по меньшей мере примерно 240 Å<sup>2</sup>, по меньшей мере примерно 250 Å<sup>2</sup>, по меньшей мере примерно 260 Å<sup>2</sup>, по меньшей мере примерно 270 Å<sup>2</sup>, по меньшей мере примерно 280 Å<sup>2</sup>, по меньшей мере примерно 290 Å<sup>2</sup>, по меньшей мере примерно 300 Å<sup>2</sup>, по меньшей мере примерно 310 Å<sup>2</sup>, по меньшей мере примерно 320 Å<sup>2</sup> или по меньшей мере примерно 330 Å<sup>2</sup>. Вторая гидрофобная аминокислота (AA<sub>H2</sub>) может иметь боковую цепь с SASA по меньшей мере примерно 200 Å<sup>2</sup>, по меньшей мере примерно 210 Å<sup>2</sup>, по меньшей мере примерно 220 Å<sup>2</sup>, по меньшей мере примерно 240 Å<sup>2</sup>, по меньшей мере примерно 250 Å<sup>2</sup>, по меньшей мере примерно 260 Å<sup>2</sup>, по меньшей мере примерно 270 Å<sup>2</sup>, по меньшей мере примерно 280 Å<sup>2</sup>, по меньшей мере примерно 290 Å<sup>2</sup>, по меньшей мере примерно 300 Å<sup>2</sup>, по меньшей мере примерно 310 Å<sup>2</sup>, по меньшей мере примерно 320 Å<sup>2</sup> или по меньшей мере примерно 330 Å<sup>2</sup>. Боковые цепи AA<sub>H1</sub> и AA<sub>H2</sub> могут иметь комбинированную SASA по меньшей мере примерно 350 Å<sup>2</sup>, по меньшей мере примерно 360 Å<sup>2</sup>, по меньшей мере примерно 370 Å<sup>2</sup>, по меньшей мере примерно 380 Å<sup>2</sup>, по меньшей мере примерно 390 Å<sup>2</sup>, по меньшей мере примерно 400 Å<sup>2</sup>, по меньшей мере примерно 410 Å<sup>2</sup>, по меньшей мере примерно 420 Å<sup>2</sup>, по меньшей мере примерно 430 Å<sup>2</sup>, по меньшей мере примерно 440 Å<sup>2</sup>, по меньшей мере примерно 450 Å<sup>2</sup>, по меньшей мере примерно 460 Å<sup>2</sup>, по меньшей мере примерно 470 Å<sup>2</sup>, по меньшей мере примерно 480 Å<sup>2</sup>.

$\text{\AA}^2$ , по меньшей мере примерно  $490 \text{\AA}^2$ , более примерно  $500 \text{\AA}^2$ , по меньшей мере примерно  $510 \text{\AA}^2$ , по меньшей мере примерно  $520 \text{\AA}^2$ , по меньшей мере примерно  $530 \text{\AA}^2$ , по меньшей мере примерно  $540 \text{\AA}^2$ , по меньшей мере примерно  $550 \text{\AA}^2$ , по меньшей мере примерно  $560 \text{\AA}^2$ , по меньшей мере примерно  $570 \text{\AA}^2$ , по меньшей мере примерно  $580 \text{\AA}^2$ , по меньшей мере примерно  $590 \text{\AA}^2$ , по меньшей мере примерно  $600 \text{\AA}^2$ , по меньшей мере примерно  $610 \text{\AA}^2$ , по меньшей мере примерно  $620 \text{\AA}^2$ , по меньшей мере примерно  $630 \text{\AA}^2$ , по меньшей мере примерно  $640 \text{\AA}^2$ , более примерно  $650 \text{\AA}^2$ , по меньшей мере примерно  $660 \text{\AA}^2$ , по меньшей мере примерно  $670 \text{\AA}^2$ , по меньшей мере примерно  $680 \text{\AA}^2$ , по меньшей мере примерно  $690 \text{\AA}^2$  или по меньшей мере примерно  $700 \text{\AA}^2$ . AA<sub>H2</sub> может представлять собой гидрофобный аминокислотный остаток с боковой цепью, имеющей SASA, которая меньше или равна SASA гидрофобной боковой цепи AA<sub>H1</sub>. В качестве примера и без ограничения, сCPP, имеющий мотив Nal-Arg, может характеризоваться улучшенной эффективностью цитозольной доставки по сравнению с идентичным в остальном сCPP, имеющими мотив Phe-Arg; сCPP, имеющий мотив Phe-Nal-Arg, может характеризоваться улучшенной эффективностью цитозольной доставки по сравнению с идентичным в остальном сCPP, имеющим мотив Nal-Phe-Arg; и мотив phe-Nal-Arg может характеризоваться улучшенной эффективностью цитозольной доставки по сравнению с идентичным в остальном сCPP, имеющим мотив nal-Phe-Arg.

В настоящем документе термин «площадь гидрофобной поверхности» или «SASA» относится к площади поверхности (сообщенной в виде квадрата значения в ангстремах;  $\text{\AA}^2$ ) аминокислотной боковой цепи, доступной растворителю, SASA может быть рассчитана с использованием алгоритма «катящегося мяча», разработанного Shrake & Rupley (*J Mol Biol.* **79** (2): 351-71), который полностью включен в настоящий документ посредством ссылки. В данном алгоритме используется «сфера» растворителя конкретного радиуса для зондирования поверхности молекулы. Типичное значение сферы составляет  $1,4 \text{\AA}$ , что приблизительно соответствует радиусу молекулы воды.

Значения SASA для некоторых боковых цепей показаны ниже в **Таблице 6**. Значения SASA, описанные в настоящем документе, основаны на теоретических значениях, перечисленных в **Таблице 6** ниже, сообщенных Tien, et al. (PLOS ONE 8(11): e80635, доступно по адресу [doi.org/10.1371/journal.pone.0080635](https://doi.org/10.1371/journal.pone.0080635)), которая полностью включена в настоящий документ посредством ссылки для всех целей.

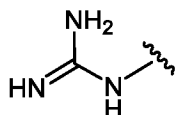
Таблица 6. Значения SASA аминокислот

Остаток	Теоретическое	Эмпирическое	Miller et al. (1987)	Rose et al. (1985)
Аланин	129,0	121,0	113,0	118,1
Аргинин	274,0	265,0	241,0	256,0
Аспарагин	195,0	187,0	158,0	165,5
Аспарат	193,0	187,0	151,0	158,7
Цистеин	167,0	148,0	140,0	146,1

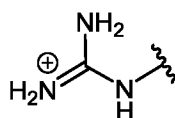
Остаток	Теоретическое	Эмпирическое	Miller <i>et al.</i> (1987)	Rose <i>et al.</i> (1985)
Глутамат	223,0	214,0	183,0	186,2
Глутамин	225,0	214,0	189,0	193,2
Глицин	104,0	97,0	85,0	88,1
Гистидин	224,0	216,0	194,0	202,5
Изолейцин	197,0	195,0	182,0	181,0
Лейцин	201,0	191,0	180,0	193,1
Лизин	236,0	230,0	211,0	225,8
Метионин	224,0	203,0	204,0	203,4
Фенилаланин	240,0	228,0	218,0	222,8
Пролин	159,0	154,0	143,0	146,8
Серин	155,0	143,0	122,0	129,8
Треонин	172,0	163,0	146,0	152,5
Триптофан	285,0	264,0	259,0	266,3
Тирозин	263,0	255,0	229,0	236,8
Валин	174,0	165,0	160,0	164,5

*Аминокислотные остатки, имеющие боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу, или ее протонированную форму*

В настоящем документе гуанидин относится к структуре:



В настоящем документе протонированная форма гуанидина относится к структуре:

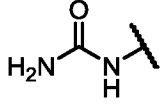
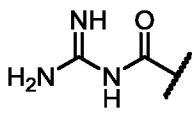
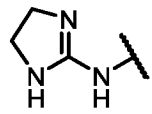
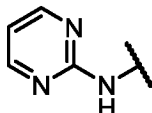
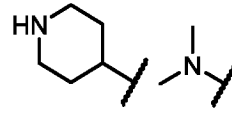


Заменяющие гуанидин группы относятся к функциональным группам на боковой цепи аминокислот, которые будут положительно заряжены при физиологическом pH или выше, или к группам, которые могут воспроизводить донорную или акцепторную активность в отношении водородной связи групп гуанидиния.

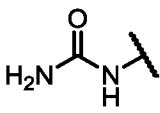
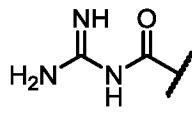
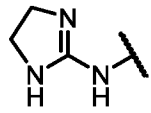
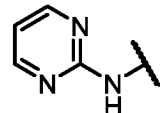
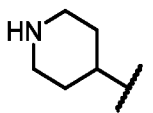

Заменяющие гуанидин группы облегчают проникновение в клетку и доставку терапевтических агентов при одновременном снижении токсичности, ассоциированной с гуанидиновыми группами или их протонированными формами. сСРР может содержать по меньшей мере одну аминокислоту, имеющую боковую цепь, содержащую гуанидин или заменяющую гуанидиний группу. сСРР может содержать по меньшей мере две аминокислоты, имеющие боковую цепь, содержащую гуанидин или заменяющую

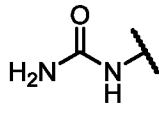
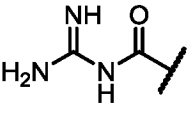
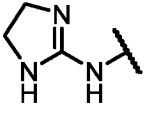
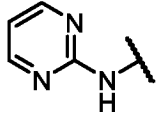
гуанидиний группу. сСРР может содержать по меньшей мере три аминокислоты, имеющие боковую цепь, содержащую гуанидин или заменяющую гуанидиний группу.

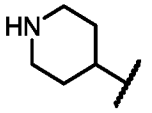

Гуанидин или группа гуанидиния может представлять собой изомер гуанидина или гуанидиния. Гуанидин или заменяющая гуанидиний группа может быть менее основной, чем гуанидин.

В настоящем документе заменяющая гуанидин группа относится к , , , , , или их протонированной форме.

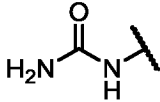
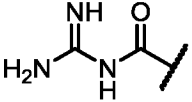
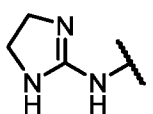
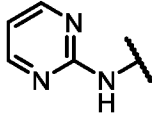
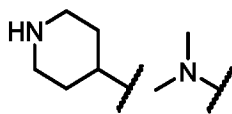
Настоящее изобретение относится к сСРР, содержащему от 4 до 20 аминокислотных остатков, причем: (i) по меньшей мере одна аминокислота имеет боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму; (ii) по меньшей мере один аминокислотный остаток не имеет боковой цепи или имеет

боковую цепь, содержащую , , , , , , или ее протонированную форму; и (iii) по меньшей мере два аминокислотных остатка независимо имеют боковую цепь, содержащую ароматическую или гетероароматическую группу.

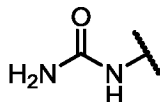
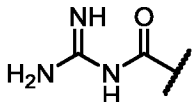
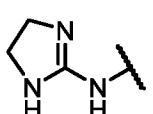
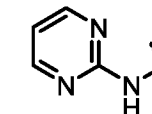
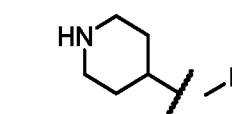
По меньшей мере два аминокислотных остатка могут не иметь боковой цепи или имеют боковую цепь, содержащую , , , ,

, , или ее протонированную форму. В настоящем документе, если боковая цепь отсутствует, аминокислотный остаток содержит два атома водорода на атоме(атомах) углерода (например, -CH<sub>2</sub>-), связывающих амин и карбоновую кислоту.

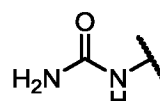
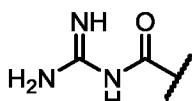
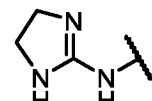
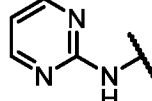
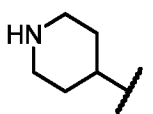
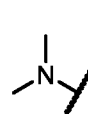
cSPР может содержать по меньшей мере одну аминокислоту, имеющую боковую

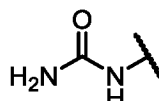
цепь, содержащую один из следующих фрагментов: , , , ,  или их протонированную форму.

cSPР может содержать по меньшей мере две аминокислоты, каждая из которых

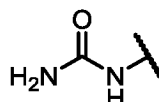
независимо содержит один из следующих фрагментов , , , , , или их протонированную форму. По

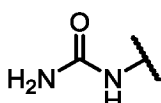
меньшей мере две аминокислоты могут иметь боковую цепь, содержащую один и тот же

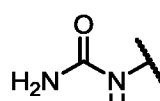
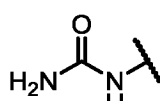
фрагмент, выбранный из: , , , , , , или их протонированной формы. По меньшей мере одна

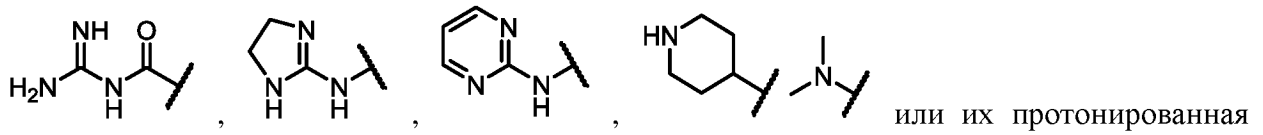
аминокислота может иметь боковую цепь, содержащую  или ее протонированную форму. По меньшей мере две аминокислоты могут иметь боковую цепь,

содержащую  или ее протонированную форму. Одна, две, три или четыре

аминокислоты могут иметь боковую цепь, содержащую  или ее протонированную форму. Одна аминокислота может иметь боковую цепь, содержащую

 или ее протонированную форму. Две аминокислоты могут иметь боковую

цепь, содержащую  или ее протонированную форму. ,



форма может быть присоединена к концу боковой цепи аминокислоты.

может быть присоединена к концу боковой цепи аминокислоты.

сСРР может содержать (iii) 2, 3, 4, 5 или 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 2 аминокислотных остатка, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 5 аминокислотных остатков, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 6 аминокислотных остатков, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 2, 3, 4 или 5 аминокислотных остатков, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 2, 3 или 4 аминокислотных остатка, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) 2 или 3 аминокислотных остатка, независимо имеющих боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму. сСРР может содержать (iii) по меньшей мере один аминокислотный остаток, имеющий боковую цепь, содержащую гуанидиновую группу или ее протонированную форму. сСРР может содержать (iii) два аминокислотных остатка, имеющих боковую цепь, содержащую гуанидиновую группу или ее протонированную форму. сСРР может содержать (iii) два аминокислотных остатка, имеющих боковую цепь, содержащую гуанидиновую группу или ее протонированную форму.

Аминокислотные остатки могут независимо иметь боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, которые не являются смежными. Два аминокислотных остатка могут независимо иметь боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, и могут быть смежными. Три аминокислотных остатка могут независимо иметь боковую цепь, содержащую гуанидиновую группу, заменяющую

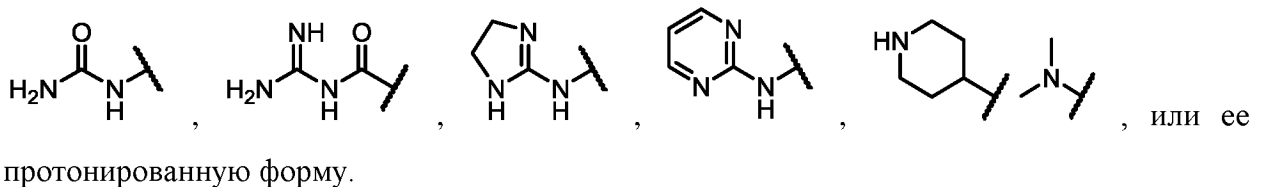


гуанидин группу или ее протонированную форму, и могут быть смежными. Четыре аминокислотных остатка могут независимо иметь боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, и могут быть смежными. Смежные аминокислотные остатки могут иметь одинаковую стереохимию. Смежные аминокислоты могут иметь чередующуюся стереохимию.

Аминокислотные остатки, независимо имеющие боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, могут представлять собой L-аминокислоты. Аминокислотные остатки, независимо имеющие боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, могут представлять собой D-аминокислоты. Аминокислотные остатки, независимо имеющие боковую цепь, содержащую гуанидиновую группу, заменяющую гуанидин группу или ее протонированную форму, могут представлять собой смесь L- или D-аминокислот.

Каждый аминокислотный остаток, имеющий боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму, может независимо представлять собой остаток аргинина, гомоаргинина, 2-амино-3-пропионовой кислоты, 2-амино-4-гуанидинмасляной кислоты или их протонированную форму. Каждый аминокислотный остаток, имеющий боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму, может независимо представлять собой остаток аргинина или его протонированную форму.

Каждая аминокислота, имеющая боковую цепь, содержащую заменяющую гуанидин группу или ее протонированную форму, может независимо представлять собой



Не ограничиваясь какой-либо теорией, предполагается, что заменяющие гуанидин группы имеют сниженную основность по сравнению с аргинином и в некоторых случаях не заряжены при физиологическом значении pH (например,  $-N(H)C(O)$ ) и способны поддерживать взаимодействия посредством бидентатных водородных связей с фосфолипидами на плазматической мембране, что, как полагают, способствует эффективной мембранной ассоциации и последующей интернализации. Также считается, что удаление положительного заряда снижает токсичность сСРР.

Специалистам в данной области техники будет понятно, что N- и/или C-концы упомянутых выше не природных ароматических гидрофобных аминокислот образуют амидные связи при включении в пептиды, раскрытые в настоящем документе.

сСРР может содержать первую аминокислоту, имеющую боковую цепь, содержащую ароматическую или гетероароматическую группу, и вторую аминокислоту, имеющую боковую цепь, содержащую ароматическую или гетероароматическую группу,

где N-конец первого глицина образует пептидную связь с первой аминокислотой, имеющей боковую цепь, содержащую ароматическую или гетероароматическую группу, и C-конец первого глицина образует пептидную связь со второй аминокислотой, имеющей боковую цепь, содержащую ароматическую или гетероароматическую группу. Хотя в соответствии с правилом термин «первая аминокислота» часто относится к N-концевой аминокислоте пептидной последовательности, в контексте настоящего документа «первая аминокислота» используется, чтобы отличить референтную аминокислоту от другой аминокислоты (например, «второй аминокислоты») в сСРР так, что термин «первая аминокислота» может или может относиться к аминокислоте, расположенной на N-конце пептидной последовательности.

сСРР может содержать N-конец второго глицина, образующего пептидную связь с аминокислотой, имеющей боковую цепь, содержащую ароматическую или гетероароматическую группу, и C-конец второго глицина образует пептидную связь с аминокислотой, имеющей боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму.

сСРР может содержать первую аминокислоту, имеющую боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму, и вторую аминокислоту, имеющую боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму, причем N-конец третьего глицина образует пептидную связь с первой аминокислотой, имеющей боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму, и C-конец третьего глицина образует пептидную связь со второй аминокислотой, имеющей боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму.

сСРР может содержать остаток аспарагина, аспарагиновой кислоты, глутамина, глутаминовой кислоты или гомоглутамина. сСРР может содержать остаток аспарагина. сСРР может содержать остаток глутамина.

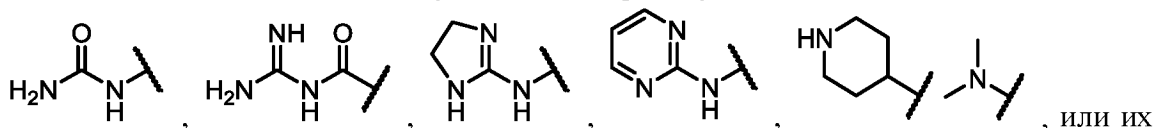
сСРР может содержать остаток тирозина, фенилаланина, 1-нафтилаланина, 2-нафтилаланина, триптофана, 3-бензотиенилаланина, 4-фенилфенилаланина, 3,4-дифторфенилаланина, 4-трифторметилфенилаланина, 2,3,4,5,6-пентафторфенилаланина, гомофенилаланина,  $\beta$ -гомофенилаланина, 4-трет-бутилфенилаланина, 4-пиридинилаланина, 3-пиридинилаланина, 4-метилфенилаланина, 4-фторфенилаланина, 4-хлорфенилаланина, 3-(9-антрил)аланина.

Не ограничиваясь какой-либо теорией, полагают, что хиральность аминокислот в сСРР может влиять на эффективность цитозольного поглощения. сСРР может содержать по меньшей мере одну D-аминокислоту. сСРР может содержать от одной до пятнадцати D-аминокислот. сСРР может содержать от одной до десяти D-аминокислот. сСРР может содержать 1, 2, 3 или 4 D-аминокислоты. сСРР может содержать 2, 3, 4, 5, 6, 7 или 8 смежных аминокислот, имеющих чередующуюся D и L-хиральность. сСРР может содержать три смежные аминокислоты, имеющие одинаковую хиральность. сСРР может содержать две смежные аминокислоты, имеющие одинаковую хиральность. По меньшей

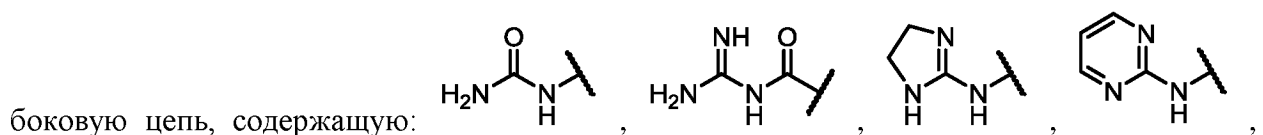
мере две аминокислоты могут иметь противоположную хиральность. По меньшей мере две аминокислоты, имеющие противоположную хиральность, могут быть смежными друг с другом. По меньшей мере три аминокислоты могут иметь чередующуюся стереохимию относительно друг друга. По меньшей мере три аминокислоты, имеющие чередующуюся хиральность относительно друг друга, могут быть смежными друг с другом. По меньшей мере четыре аминокислоты имеют чередующуюся стереохимию относительно друг друга. По меньшей мере четыре аминокислоты, имеющие чередующуюся хиральность относительно друг друга, могут быть смежными друг с другом. По меньшей мере две аминокислоты могут иметь одинаковую хиральность. По меньшей мере две аминокислоты, имеющие одинаковую хиральность, могут быть смежными друг с другом. По меньшей мере две аминокислоты имеют одинаковую хиральность, и по меньшей мере две аминокислоты имеют противоположную хиральность. По меньшей мере две аминокислоты, имеющие противоположную хиральность, могут быть смежными по меньшей мере с двумя аминокислотами, имеющими одинаковую хиральность. Соответственно, смежные аминокислоты в сСРР могут иметь любую из следующих последовательностей: 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; или L-D-D-L-D. Все аминокислотные остатки, которые образуют сСРР, могут представлять собой L-аминокислоты. Все аминокислотные остатки, которые образуют сСРР, могут представлять собой D-аминокислоты.

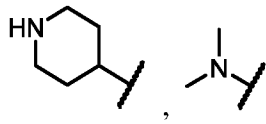
По меньшей мере две аминокислоты могут иметь различную хиральность. По меньшей мере две аминокислоты, имеющие разную хиральность, могут быть смежными друг с другом. По меньшей мере три аминокислоты могут иметь хиральность, отличную от хиральности смежной аминокислоты. По меньшей мере четыре аминокислоты могут иметь хиральность, отличную от хиральности смежной аминокислоты. По меньшей мере две аминокислоты имеют одинаковую хиральность, и по меньшей мере две аминокислоты имеют разную хиральность. Один или более аминокислотных остатков, образующих сСРР, могут быть ахиральными. сСРР может содержать мотив из 3, 4 или 5 аминокислот, в котором две аминокислоты, имеющие одинаковую хиральность, могут быть разделены ахиральной аминокислотой. сСРР могут содержать следующие последовательности: D-X-D; D-X-D-X; D-X-D-X-D; L-X-L; L-X-L-X; или L-X-L-X-L, где X представляет собой ахиральную аминокислоту. Ахиральная аминокислота может представлять собой глицин.

Аминокислота, имеющая боковую цепь, содержащую:



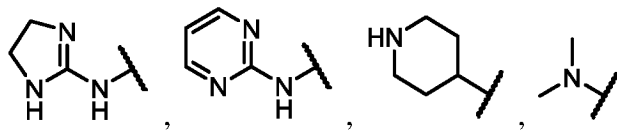
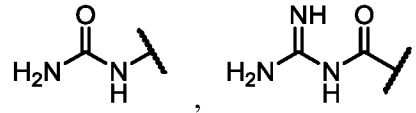
протонированную форму, может быть смежной с аминокислотой, имеющей боковую цепь, содержащую ароматическую или гетероароматическую группу. Аминокислота, имеющая





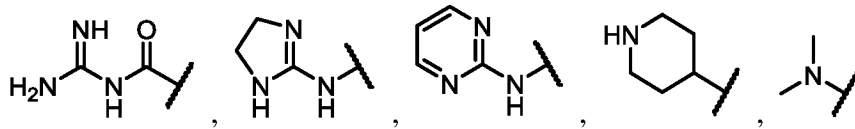
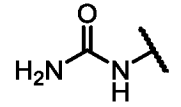
, или их протонированную форму, может быть смежной по меньшей мере с одной аминокислотой, имеющей боковую цепь, содержащую гуанидин или его протонированную форму. Аминокислота, имеющая боковую цепь, содержащую гуанидин или его протонированную форму, может быть смежной с аминокислотой, имеющей боковую цепь, содержащую ароматическую или гетероароматическую группу. Две

аминокислоты, имеющие боковую цепь, содержащую:

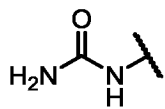


или их протонированную форму, могут быть смежными друг с другом. Две аминокислоты, имеющие боковую цепь, содержащую гуанидин или его протонированную форму, являются смежными друг с другом. сСРР могут содержать по меньшей мере две смежные аминокислоты, имеющие боковую цепь, которая может содержать ароматическую или гетероароматическую группу, и по меньшей

мере две несмежные аминокислоты, имеющие боковую цепь, содержащую:

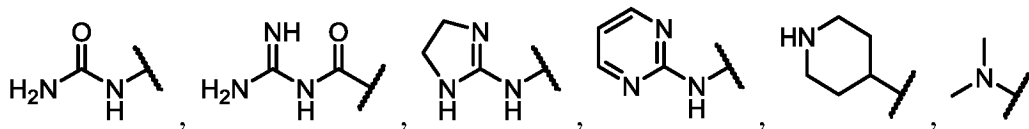


, или их протонированную форму. сСРР могут содержать по меньшей мере две смежные аминокислоты, имеющие боковую цепь, содержащую ароматическую или гетероароматическую группу, и по меньшей мере две несмежные аминокислоты, имеющие боковую цепь, содержащую



или ее протонированную форму. Смежные аминокислоты могут иметь одинаковую хиральность. Смежные аминокислоты могут иметь противоположную хиральность. Другие комбинации аминокислот могут иметь любую компоновку D и L-аминокислот, например, любую из последовательностей, описанных в предыдущем абзаце.

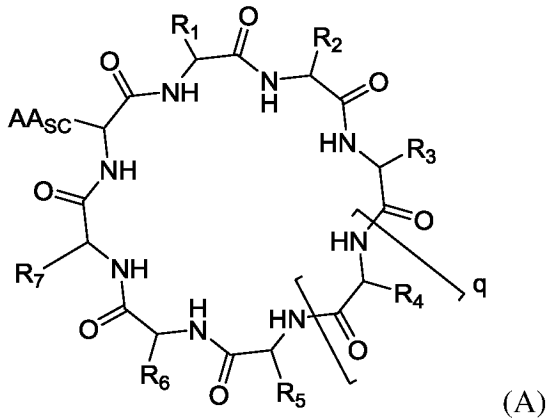
По меньшей мере две аминокислоты, имеющие боковую цепь, содержащую:



, или их протонированную форму, чередуются по меньшей мере с двумя аминокислотами,

имеющими боковую цепь, содержащую гуанидиновую группу или ее протонированную форму.

cCPP может содержать структуру Формулы (А):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или ароматическую или гетероароматическую боковую цепь аминокислоты;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  представляет собой боковую цепь 3-гуанидин-2-аминопропионовой кислоты, 4-гуанидин-2-аминобутановой кислоты, аргинина, гомоаргинина, N-метиларгинина, N, N-диметиларгинина, 2,3-диаминопропионовой кислоты, 2,4-диаминобутановой кислоты, лизина, N-метиллизина, N, N-диметиллизина, N-этиллизина, N, N, N-триметиллизина, 4-гуанидинфенилаланина, цитруллина, N, N-диметиллизина,  $\beta$ -гомоаргинина, 3-(1-пиперидинил)аланина;

$AA_{sc}$  представляет собой боковую цепь аминокислоты; и

$q$  представляет собой 1, 2, 3 или 4.

Согласно вариантам реализации по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляет собой незаряженную неароматическую боковую цепь аминокислоты. Согласно вариантам реализации по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляет собой H или боковую цепь цитруллина.

Согласно вариантам реализации предложены соединения, которые включают циклический пептид, содержащий от 6 до 12 аминокислот, причем по меньшей мере две аминокислоты циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты циклического пептида представляют собой ароматические гидрофобные аминокислоты и по меньшей мере две аминокислоты циклического пептида представляют собой незаряженные неароматические аминокислоты. Согласно вариантам реализации по меньшей мере две заряженные аминокислоты циклического пептида представляют собой аргинин. Согласно вариантам

реализации по меньшей мере две ароматические гидрофобные аминокислоты циклического пептида представляют собой фенилаланин или нафтилаланин. Согласно вариантам реализации по меньшей мере две незаряженные неароматические аминокислоты циклического пептида представляют собой цитруллин или глицин.

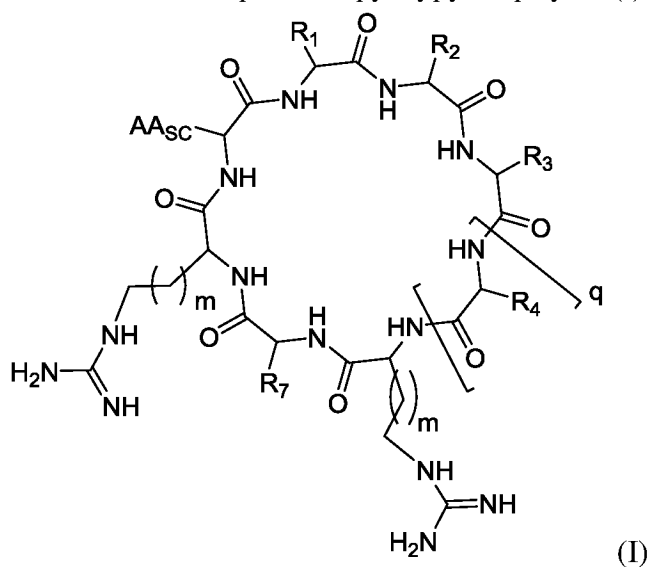
Согласно вариантам реализации циклический пептид Формулы (A) не выбран из циклического пептида, имеющего последовательность SEQ ID NO: 89-117.

Согласно вариантам реализации циклический пептид Формулы (A) выбран из циклического пептида, имеющего последовательность SEQ ID NO: 89-117.

Последовательности CPP и 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		

Φ=L-нафтилаланин; φ= D-нафтилаланин; Ω=L-норлейцин

cCPP может содержать структуру Формулы (I):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или

гетероароматическую боковую цепь аминокислоты;

$R_4$  и  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

$AA_{SC}$  представляет собой боковую цепь аминокислоты;

$q$  представляет собой 1, 2, 3 или 4; и

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3.

Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H, -алкилен-арил или -алкилен-гетероарил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H,  $-C_{1-3}$ алкилен-арил или  $-C_{1-3}$ алкилен-гетероарил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или -алкилен-арил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или  $-C_{1-3}$ алкилен-арил.  $C_{1-3}$ алкилен может представлять собой метилен. Арил может представлять собой 6-14-членный арил. Гетероарил может представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O и S. Арил может быть выбран из фенила, нафтила или антраценила. Арил может представлять собой фенил или нафтил. Арил может представлять собой фенил. Гетероарил может представлять собой пиридил, хинолил и изохинолил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H,  $-C_{1-3}$ алкилен-Ph или  $-C_{1-3}$ алкилен-нафтил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H,  $-CH_2Ph$  или  $-CH_2$ нафтил. Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или  $-CH_2Ph$ .

Каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой боковую цепь тирозина, фенилаланина, 1-нафтилаланина, 2-нафтилаланина, триптофана, 3-бензотиенилаланина, 4-фенилфенилаланина, 3,4-дифторфенилаланина, 4-трифторметилфенилаланина, 2,3,4,5,6-пентафторфенилаланина, гомофенилаланина,  $\beta$ -гомофенилаланина, 4-трет-бутилфенилаланина, 4-пиридинилаланина, 3-пиридинилаланина, 4-метилфенилаланина, 4-фторфенилаланина, 4-хлорфенилаланина, 3-(9-антрил)-аланина.

$R_1$  может представлять собой боковую цепь тирозина.  $R_1$  может представлять собой боковую цепь фенилаланина.  $R_1$  может представлять собой боковую цепь 1-нафтилаланина.  $R_1$  может представлять собой боковую цепь 2-нафтилаланина.  $R_1$  может представлять собой боковую цепь триптофана.  $R_1$  может представлять собой боковую цепь 3-бензотиенилаланина.  $R_1$  может представлять собой боковую цепь 4-фенилфенилаланина.  $R_1$  может представлять собой боковую цепь 3,4-дифторфенилаланина.  $R_1$  может представлять собой боковую цепь 4-трифторметилфенилаланина.  $R_1$  может представлять собой боковую цепь 2,3,4,5,6-пентафторфенилаланина.  $R_1$  может представлять собой боковую цепь гомофенилаланина.  $R_1$  может представлять собой боковую цепь  $\beta$ -гомофенилаланина.  $R_1$  может представлять собой боковую цепь 4-трет-бутилфенилаланина.  $R_1$  может представлять собой боковую цепь 4-пиридинилаланина.  $R_1$  может представлять собой боковую цепь 3-пиридинилаланина.  $R_1$  может представлять собой боковую цепь 4-метилфенилаланина.  $R_1$  может представлять собой боковую цепь 4-фторфенилаланина.  $R_1$  может представлять собой боковую цепь 4-хлорфенилаланина.  $R_1$  может представлять собой боковую цепь 3-(9-антрил)-аланина.

$R_2$  может представлять собой боковую цепь тирозина.  $R_2$  может представлять собой боковую цепь фенилаланина.  $R_2$  может представлять собой боковую цепь 1-нафтилаланина.  $R_1$  может представлять собой боковую цепь 2-нафтилаланина.  $R_2$  может представлять собой боковую цепь триптофана.  $R_2$  может представлять собой боковую цепь 3-бензотиенилаланина.  $R_2$  может представлять собой боковую цепь 4-фенилфенилаланина.  $R_2$  может представлять собой боковую цепь 3,4-дифторфенилаланина.  $R_2$  может представлять собой боковую цепь 4-трифторметилфенилаланина.  $R_2$  может представлять собой боковую цепь 2,3,4,5,6-пентафторфенилаланина.  $R_2$  может представлять собой боковую цепь гомофенилаланина.  $R_2$  может представлять собой боковую цепь  $\beta$ -гомофенилаланина.  $R_2$  может представлять собой боковую цепь 4-трет-бутилфенилаланина.  $R_2$  может представлять собой боковую цепь 4-пиридинилаланина.  $R_2$  может представлять собой боковую цепь 3-пиридинилаланина.  $R_2$  может представлять собой боковую цепь 4-метилфенилаланина.  $R_2$  может представлять собой боковую цепь 4-фторфенилаланина.  $R_2$  может представлять собой боковую цепь 4-хлорфенилаланина.  $R_2$  может представлять собой боковую цепь 3-(9-антрил)-аланина.

$R_3$  может представлять собой боковую цепь тирозина.  $R_3$  может представлять собой боковую цепь фенилаланина.  $R_3$  может представлять собой боковую цепь 1-нафтилаланина.  $R_3$  может представлять собой боковую цепь 2-нафтилаланина.  $R_3$  может представлять собой боковую цепь триптофана.  $R_3$  может представлять собой боковую цепь 3-бензотиенилаланина.  $R_3$  может представлять собой боковую цепь 4-фенилфенилаланина.  $R_3$  может представлять собой боковую цепь 3,4-дифторфенилаланина.  $R_3$  может представлять собой боковую цепь 4-трифторметилфенилаланина.  $R_3$  может представлять собой боковую цепь 2,3,4,5,6-пентафторфенилаланина.  $R_3$  может представлять собой боковую цепь гомофенилаланина.  $R_3$  может представлять собой боковую цепь  $\beta$ -гомофенилаланина.  $R_3$  может представлять собой боковую цепь 4-трет-бутилфенилаланина.  $R_3$  может представлять собой боковую цепь 4-пиридинилаланина.  $R_3$  может представлять собой боковую цепь 3-пиридинилаланина.  $R_3$  может представлять собой боковую цепь 4-метилфенилаланина.  $R_3$  может представлять собой боковую цепь 4-фторфенилаланина.  $R_3$  может представлять собой боковую цепь 4-хлорфенилаланина.  $R_3$  может представлять собой боковую цепь 3-(9-антрил)-аланина.

$R_4$  может представлять собой H, -алкилен-арил, -алкилен-гетероарил.  $R_4$  может представлять собой H,  $-C_{1-3}$ алкилен-арил или  $-C_{1-3}$ алкилен-гетероарил.  $R_4$  может представлять собой H или -алкилен-арил.  $R_4$  может представлять собой H или  $-C_{1-3}$ алкилен-арил.  $C_{1-3}$ алкилен может представлять собой метилен. Арил может представлять собой 6-14-членный арил. Гетероарил может представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O и S. Арил может быть выбран из фенила, нафтила или антраценила. Арил может представлять собой фенил или нафтил. Арил может представлять собой фенил. Гетероарил может



представлять собой пиридил, хинолил и изохинолил. R<sub>4</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-Ph или -C<sub>1-3</sub>алкилен-нафтил. R<sub>4</sub> может представлять собой H или боковую цепь аминокислоты в **Таблице 4** или **Таблице 6**. R<sub>4</sub> может представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу. R<sub>4</sub> может представлять собой H, -CH<sub>2</sub>Ph или -CH<sub>2</sub>нафтил. R<sub>4</sub> может представлять собой H или -CH<sub>2</sub>Ph.

R<sub>5</sub> может представлять собой H, -алкилен-арил, -алкилен-гетероарил. R<sub>5</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-арил или -C<sub>1-3</sub>алкилен-гетероарил. R<sub>5</sub> может представлять собой H или -алкилен-арил. R<sub>5</sub> может представлять собой H или -C<sub>1-3</sub>алкилен-арил. C<sub>1-3</sub>алкилен может представлять собой метилен. Арил может представлять собой 6-14-членный арил. Гетероарил может представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O и S. Арил может быть выбран из фенила, нафтила или антраценила. Арил может представлять собой фенил или нафтил. Арил может представлять собой фенил. Гетероарил может представлять собой пиридил, хинолил и изохинолил. R<sub>5</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-Ph или -C<sub>1-3</sub>алкилен-нафтил. R<sub>5</sub> может представлять собой H или боковую цепь аминокислоты в **Таблице 4** или **Таблице 6**. R<sub>4</sub> может представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу. R<sub>5</sub> может представлять собой H, -CH<sub>2</sub>Ph или -CH<sub>2</sub>нафтил. R<sub>4</sub> может представлять собой H или -CH<sub>2</sub>Ph.

R<sub>6</sub> может представлять собой H, -алкилен-арил, -алкилен-гетероарил. R<sub>6</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-арил или -C<sub>1-3</sub>алкилен-гетероарил. R<sub>6</sub> может представлять собой H или -алкилен-арил. R<sub>6</sub> может представлять собой H или -C<sub>1-3</sub>алкилен-арил. C<sub>1-3</sub>алкилен может представлять собой метилен. Арил может представлять собой 6-14-членный арил. Гетероарил может представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O и S. Арил может быть выбран из фенила, нафтила или антраценила. Арил может представлять собой фенил или нафтил. Арил может представлять собой фенил. Гетероарил может представлять собой пиридил, хинолил и изохинолил. R<sub>6</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-Ph или -C<sub>1-3</sub>алкилен-нафтил. R<sub>6</sub> может представлять собой H или боковую цепь аминокислоты в **Таблице 4** или **Таблице 6**. R<sub>6</sub> может представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу. R<sub>6</sub> может представлять собой H, -CH<sub>2</sub>Ph или -CH<sub>2</sub>нафтил. R<sub>6</sub> может представлять собой H или -CH<sub>2</sub>Ph.

R<sub>7</sub> может представлять собой H, -алкилен-арил, -алкилен-гетероарил. R<sub>7</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-арил или -C<sub>1-3</sub>алкилен-гетероарил. R<sub>7</sub> может представлять собой H или -алкилен-арил. R<sub>7</sub> может представлять собой H или -C<sub>1-3</sub>алкилен-арил. C<sub>1-3</sub>алкилен может представлять собой метилен. Арил может представлять собой 6-14-членный арил. Гетероарил может представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O и S. Арил

может быть выбран из фенила, нафтила или антраценила. Арил может представлять собой фенил или нафтил. Арил может представлять собой фенил. Гетероарил может представлять собой пиридил, хинолил и изохинолил. R<sub>7</sub> может представлять собой H, -C<sub>1-3</sub>алкилен-Ph или -C<sub>1-3</sub>алкилен-нафтил. R<sub>7</sub> может представлять собой H или боковую цепь аминокислоты в **Таблице 4** или **Таблице 6**. R<sub>7</sub> может представлять собой H или боковую цепь, содержащую ароматическую группу. R<sub>7</sub> может представлять собой H, -CH<sub>2</sub>Ph или -CH<sub>2</sub>нафтил. R<sub>7</sub> может представлять собой H или -CH<sub>2</sub>Ph.

Один, два или три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой -CH<sub>2</sub>Ph. Один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой -CH<sub>2</sub>Ph. Два из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой -CH<sub>2</sub>Ph. Три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой -CH<sub>2</sub>Ph. По меньшей мере один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой -CH<sub>2</sub>Ph. Не более четырех из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой -CH<sub>2</sub>Ph.

Один, два или три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой -CH<sub>2</sub>Ph. Один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляет собой -CH<sub>2</sub>Ph. Два из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой -CH<sub>2</sub>Ph. Три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой -CH<sub>2</sub>Ph. По меньшей мере один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляет собой -CH<sub>2</sub>Ph.

Один, два или три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой H. Один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой H. Два из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> представляют собой H. Три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой H. По меньшей мере один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой H. Не более трех из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> могут представлять собой -CH<sub>2</sub>Ph.

Один, два или три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой H. Один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляет собой H. Два из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой H. Три из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляют собой H. По меньшей мере один из R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> и R<sub>4</sub> представляет собой H.

По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь 3-гуанидин-2-аминопропионовой кислоты. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь 4-гуанидин-2-аминобутановой кислоты. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь аргинина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь гомоаргинина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь N-метиларгинина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь N,N-диметиларгинина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь 2,3-диаминопропионовой кислоты. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь 2,4-диаминобутановой кислоты, лизина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь N-метиллизина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь N,N-диметиллизина. По меньшей мере один из R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> и R<sub>7</sub> может представлять собой боковую цепь N-этиллизина. По меньшей



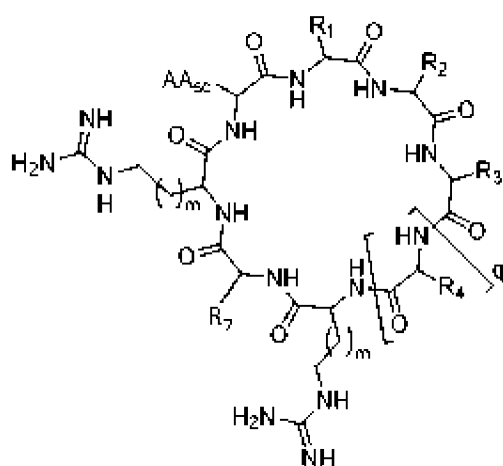
меньшей мере три из  $R_4$ ,  $R_5$ ,  $R_6$  и  $R_7$  могут представлять собой боковую цепь цитруллина. По меньшей мере три из  $R_4$ ,  $R_5$ ,  $R_6$  и  $R_7$  могут представлять собой боковую цепь N,N-диметиллизина,  $\beta$ -гомоаргинина. По меньшей мере три из  $R_4$ ,  $R_5$ ,  $R_6$  и  $R_7$  могут представлять собой боковую цепь 3-(1-пиперидинил)аланина.

$AA_{SC}$  может представлять собой боковую цепь остатка аспарагина, глутамина или гомоглутамина.  $AA_{SC}$  может представлять собой боковую цепь остатка глутамина.  $sCPR$  может дополнительно содержать линкер, конъюгированный с  $AA_{SC}$ , например, остатком аспарагина, глутамина или гомоглутамина. Следовательно,  $sCPR$  может дополнительно содержать линкер, конъюгированный с остатком аспарагина, глутамина или гомоглутамина.  $sCPR$  может дополнительно содержать линкер, конъюгированный с остатком глутамина.

$q$  может представлять собой 1, 2 или 3.  $q$  может представлять собой 1 или 2.  $q$  может представлять собой 1.  $q$  может представлять собой 2.  $q$  может представлять собой 3.  $q$  может представлять собой 4.

$m$  может представлять собой 1-3.  $m$  может представлять собой 1 или 2.  $m$  может представлять собой 0.  $m$  может представлять собой 1.  $m$  может представлять собой 2.  $m$  может представлять собой 3.

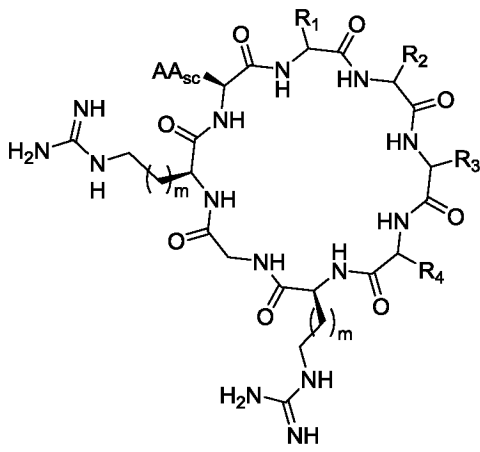
$sCPR$  Формулы (A) может содержать структуру Формулы (I)



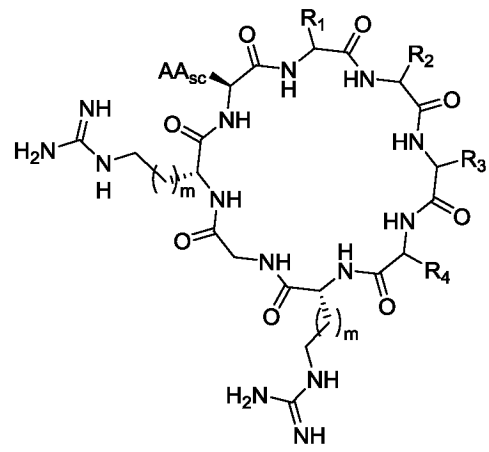
(I)

или ее протонированную форму, где  $AA_{SC}$ ,  $R_1$ ,  $R_2$ ,  $R_3$ ,  $R_4$ ,  $R_7$ ,  $m$  и  $q$  определены в настоящем документе.

$sCPR$  Формулы (A) может содержать структуру Формулы (I-a) или Формулы (I-b):



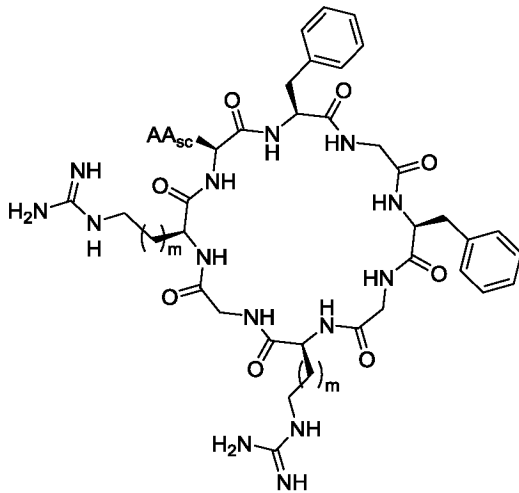
(I-a),



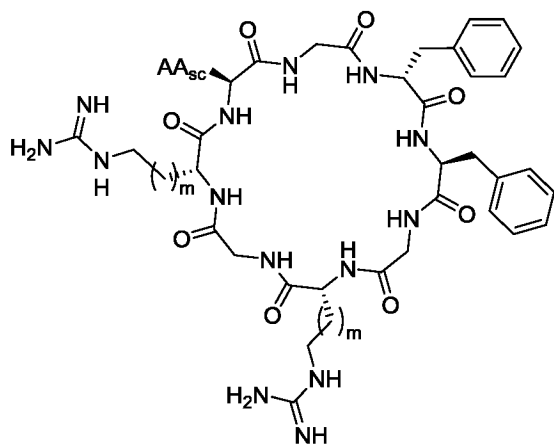
(I-b),

или ее протонированную форму, где AA<sub>sc</sub>, R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub> и m определены в настоящем документе.

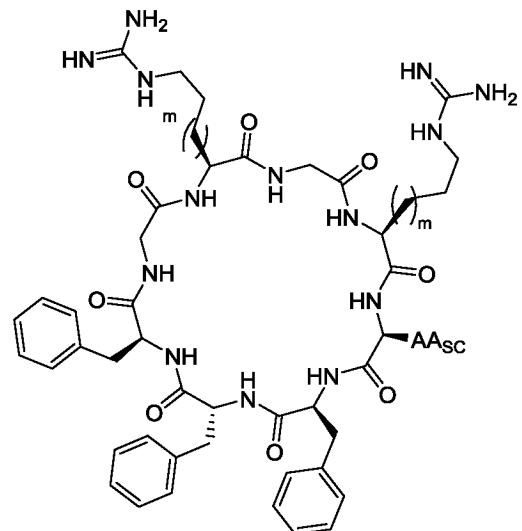
cSPR Формулы (A) может содержать структуру Формулы (I-1), (I-2), (I-3) или (I-4):



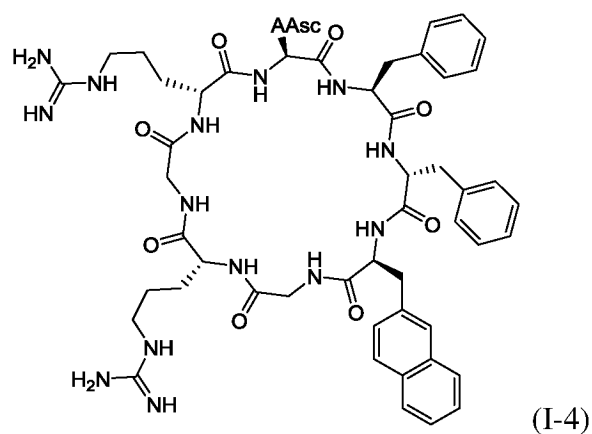
(I-1),



(I-2),

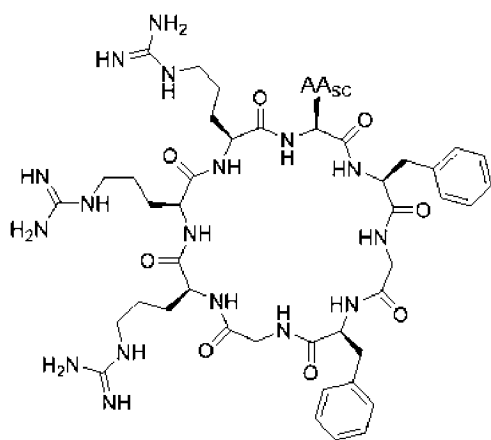


(I-3),



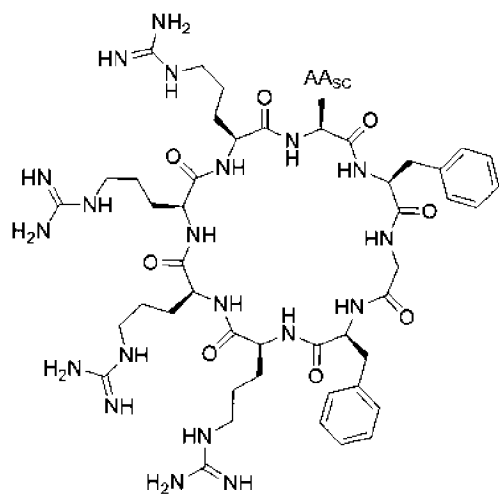
(I-4)

или ее протонированную форму, где AA<sub>SC</sub> и m определены в настоящем документе.  
сСРР Формулы (A) может содержать структуру Формулы (I-5) или (I-6):



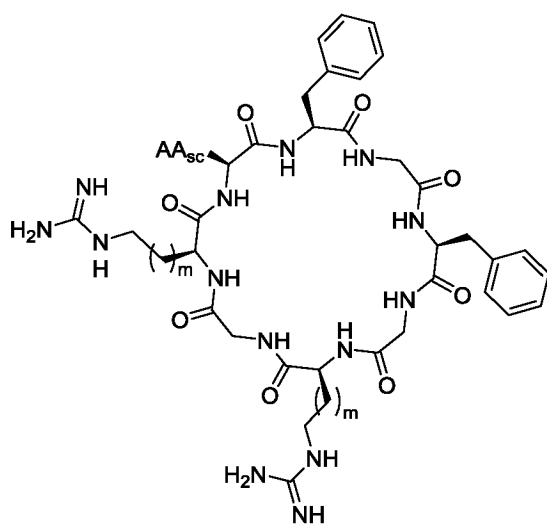
(I-5)

или



(I-6)

или ее протонированную форму, где AA<sub>SC</sub> определен в настоящем документе.  
сСРР Формулы (A) может содержать структуру Формулы (I-1):

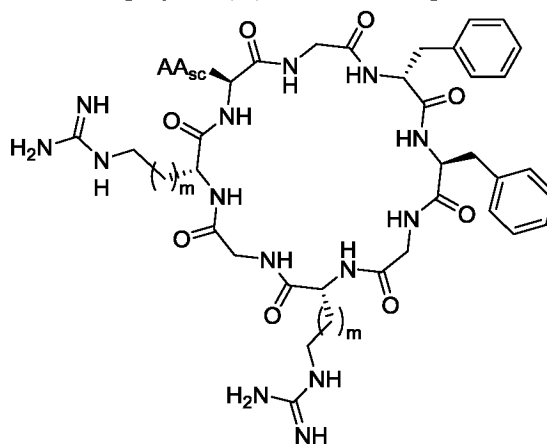


(I-1),

или ее протонированную форму,

где AA<sub>sc</sub> и m определены в настоящем документе.

cSPР Формулы (A) может содержать структуру Формулы (I-2):

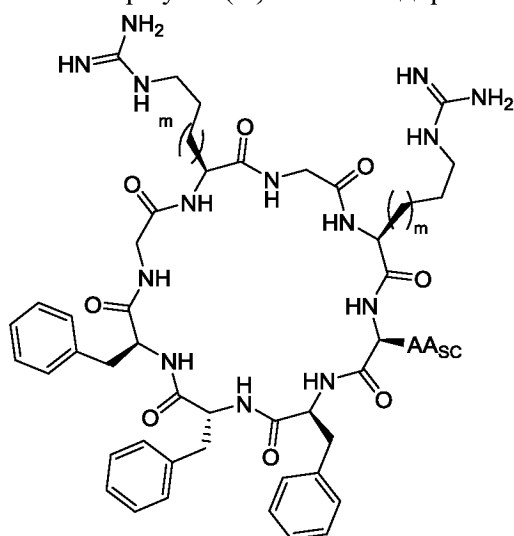


(I-2),

или ее протонированную форму,

где AA<sub>sc</sub> и m определены в настоящем документе.

cSPР Формулы (A) может содержать структуру Формулы (I-3):

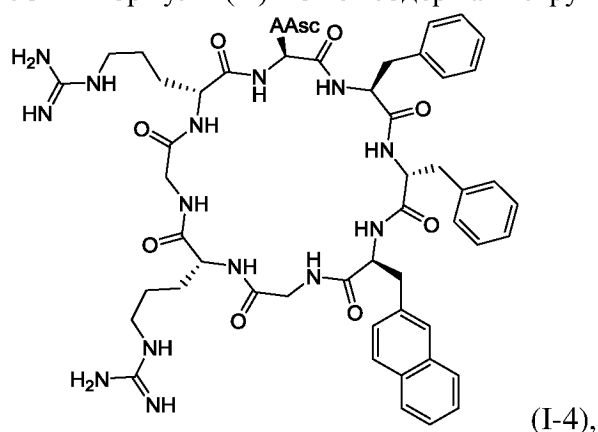


(I-3),

или ее протонированную форму,

где  $AA_{Sc}$  и  $m$  определены в настоящем документе.

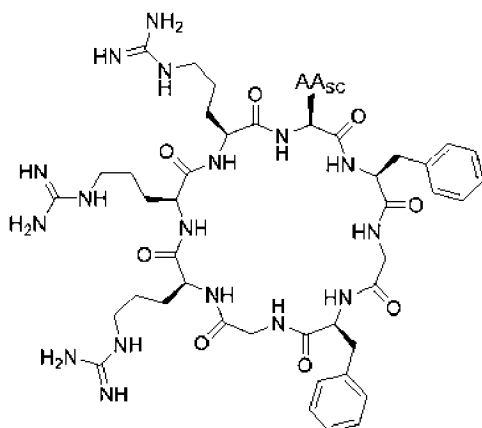
cCPP Формулы (A) может содержать структуру Формулы (I-4):



или ее протонированную форму,

где  $AA_{Sc}$  и  $m$  определены в настоящем документе.

cCPP Формулы (A) может содержать структуру Формулы (I-5):



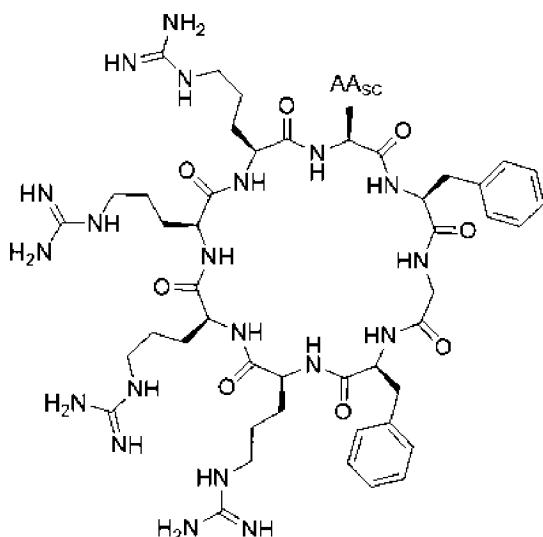
(I-5),

или ее протонированную форму,

где  $AA_{Sc}$  и  $m$  определены в настоящем документе.

cCPP Формулы (A) может содержать структуру Формулы (I-6):



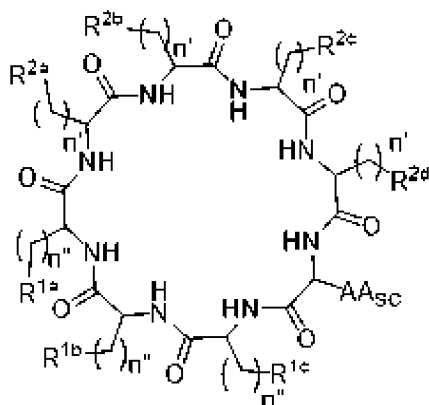


(I-6),

или ее протонированную форму, где AA<sub>SC</sub> и m определены в настоящем документе.

cCPP может содержать одну из следующих последовательностей: FGFGRR (SEQ ID NO:68); GfFGRR (SEQ ID NO:69), FfΦGRGR (SEQ ID NO:70); FfFGRR (SEQ ID NO:71); или FfΦGrGr (SEQ ID NO:72). cCPP может иметь одну из следующих последовательностей: FGFF (SEQ ID NO:73); GfFGRRQ (SEQ ID NO:74), FfΦGRGRQ (SEQ ID NO:75); FfFGRRQ (SEQ ID NO:76); или FfΦGrGrQ (SEQ ID NO:77).

Настоящее изобретение также относится к cCPP, имеющему структуру Формулы (II):



(II)

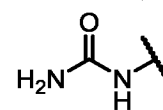
где:

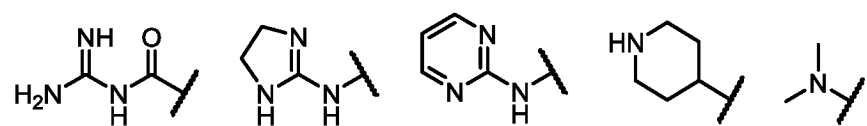
AA<sub>SC</sub> представляет собой боковую цепь аминокислоты;

каждый R<sup>1a</sup>, R<sup>1b</sup> и R<sup>1c</sup> независимо представляет собой 6-14-членный арил или 6-14-членный гетероарил;

R<sup>2a</sup>, R<sup>2b</sup>, R<sup>2c</sup> и R<sup>2d</sup> независимо представляют собой боковую цепь аминокислоты;

по меньшей мере один из R<sup>2a</sup>, R<sup>2b</sup>, R<sup>2c</sup> и R<sup>2d</sup> представляет собой




 , или их протонированную форму;

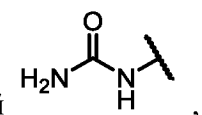
по меньшей мере один из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  представляет собой гуанидин или его протонированную форму;

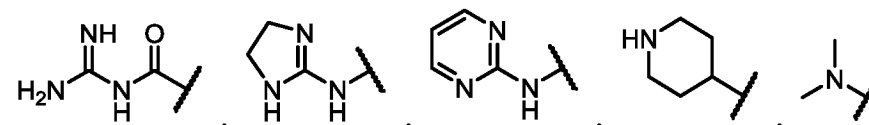
каждый  $n''$  независимо представляет собой целое число 0, 1, 2, 3, 4 или 5;

каждый  $n'$  независимо представляет собой целое число 0, 1, 2 или 3; и

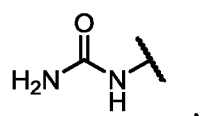
если  $n'$  представляет собой 0, то  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  или  $R^{2d}$  отсутствует.

По меньшей мере два из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой

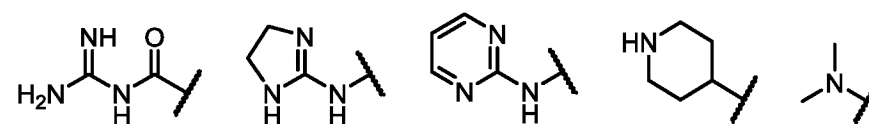



 , или их протонированную форму.

Два или три из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой



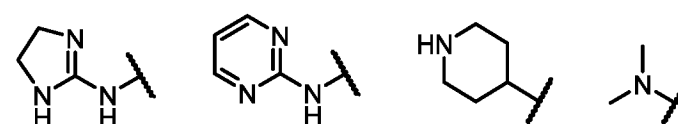
форму.


 , или их протонированную форму.

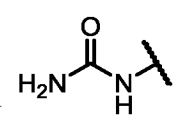
Один из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  может представлять собой



форму.

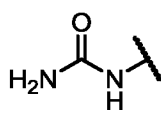

 , или их протонированную форму. По

меньшей мере один из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  может представлять собой



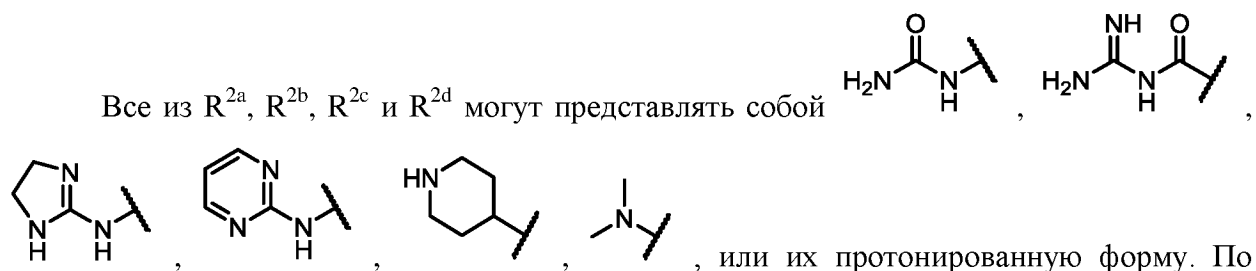
форму, и остальные  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой гуанидин или его протонированную форму. По меньшей мере два из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$

могут представлять собой

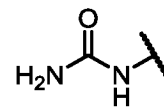


или его протонированную форму, и остальные  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой гуанидин или его протонированную форму.

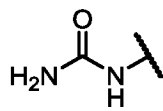
Все из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой



или их протонированную форму. По меньшей мере  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой



и остальные  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой гуанидин или его протонированную форму. По меньшей мере две группы  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  могут представлять собой



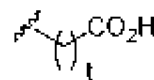
и остальные  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  представляют собой гуанидин, или его протонированную форму.

Каждый из  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$  и  $R^{2d}$  может независимо представлять собой 2,3-диаминопропионовую кислоту, 2,4-диаминомасляную кислоту, боковые цепи орнитина, лизина, метиллизина, диметиллизина, триметиллизина, гомолизина, серина, гомосерина, треонина, аллотреонина, гистидина, 1-метилгистидина, 2-аминобутандиовой кислоты, аспарагиновой кислоты, глутаминовой кислоты или гомоглутаминовой кислоты.

$AA_{SC}$  может представлять собой

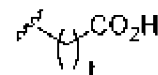


или



, где  $t$  может

представлять собой целое число от 0 до 5.  $AA_{SC}$  может представлять собой



где  $t$  может представлять собой целое число от 0 до 5.  $t$  может представлять собой целое число от 1 до 5.  $t$  представляет собой 2 или 3.  $t$  может представлять собой 2.  $t$  может представлять собой 3.

Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может независимо представлять собой 6-14-членный арил. Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может независимо представлять собой 6-14-членный гетероарил, содержащий один или более гетероатомов, выбранных из N, O или S. Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может быть независимо выбран из фенила, нафтила, антраценила, пиридила, хинолила или изохинолила. Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может быть независимо выбран из фенила, нафтила или антраценила. Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может независимо представлять собой фенил или нафтил. Каждый  $R^{1a}$ ,  $R^{1b}$  и  $R^{1c}$  может быть независимо выбран из пиридила, хинолила или изохинолила.

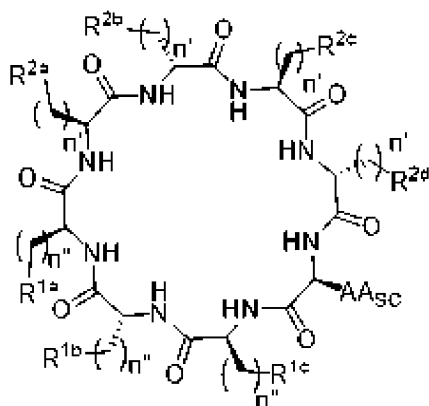
Каждый  $n'$  может независимо представлять собой 1 или 2. Каждый  $n'$  может представлять собой 1. Каждый  $n'$  может представлять собой 2. По меньшей мере один  $n'$  может представлять собой 0. По меньшей мере один  $n'$  может представлять собой 1. По меньшей мере один  $n'$  может представлять собой 2. По меньшей мере один  $n'$  может

представлять собой 3. По меньшей мере один  $n'$  может представлять собой 4. По меньшей мере один  $n'$  может представлять собой 5.

Каждый  $n''$  может независимо представлять собой целое число от 1 до 3. Каждый  $n''$  может независимо представлять собой 2 или 3. Каждый  $n''$  может представлять собой 2. Каждый  $n''$  может представлять собой 3. По меньшей мере один  $n''$  может представлять собой 0. По меньшей мере один  $n''$  может представлять собой 1. По меньшей мере один  $n''$  может представлять собой 2. По меньшей мере один  $n''$  может представлять собой 3.

Каждый  $n''$  может независимо представлять собой 1 или 2, и каждый  $n'$  может независимо представлять собой 2 или 3. Каждый  $n''$  может представлять собой 1, и каждый  $n'$  может независимо представлять собой 2 или 3. Каждый  $n''$  может представлять собой 1, и каждый  $n'$  может представлять собой 2. Каждый  $n''$  представляет собой 1, и каждый  $n'$  представляет собой 3.

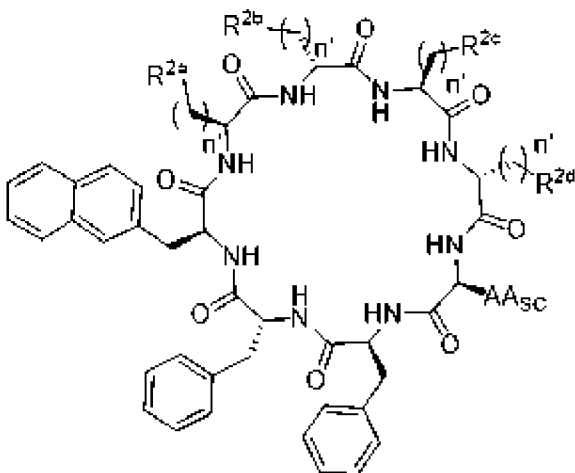
cCPP Формулы (II) может иметь структуру Формулы (II-1):



(II-1),

где  $R^{1a}$ ,  $R^{1b}$ ,  $R^{1c}$ ,  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$ ,  $R^{2d}$ ,  $AA_{5c}$ ,  $n'$  и  $n''$  определены в настоящем документе.

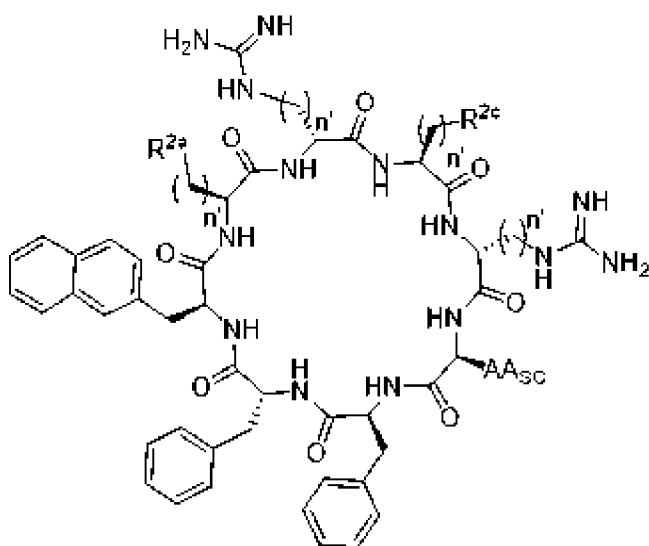
cCPP Формулы (II) может иметь структуру Формулы (IIa):



(IIa),

где  $R^{1a}$ ,  $R^{1b}$ ,  $R^{1c}$ ,  $R^{2a}$ ,  $R^{2b}$ ,  $R^{2c}$ ,  $R^{2d}$ ,  $AA_{5c}$  и  $n'$  определены в настоящем документе.

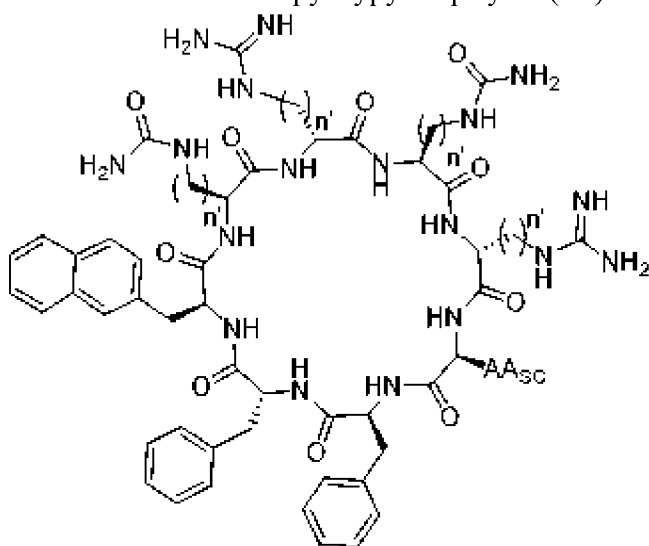
cCPP Формулы (II) может иметь структуру Формулы (IIb):



(Ib),

где  $R^{2a}$ ,  $R^{2b}$ ,  $AA_{5C}$  и  $n'$  определены в настоящем документе.

cCPR может иметь структуру Формулы (Ic):



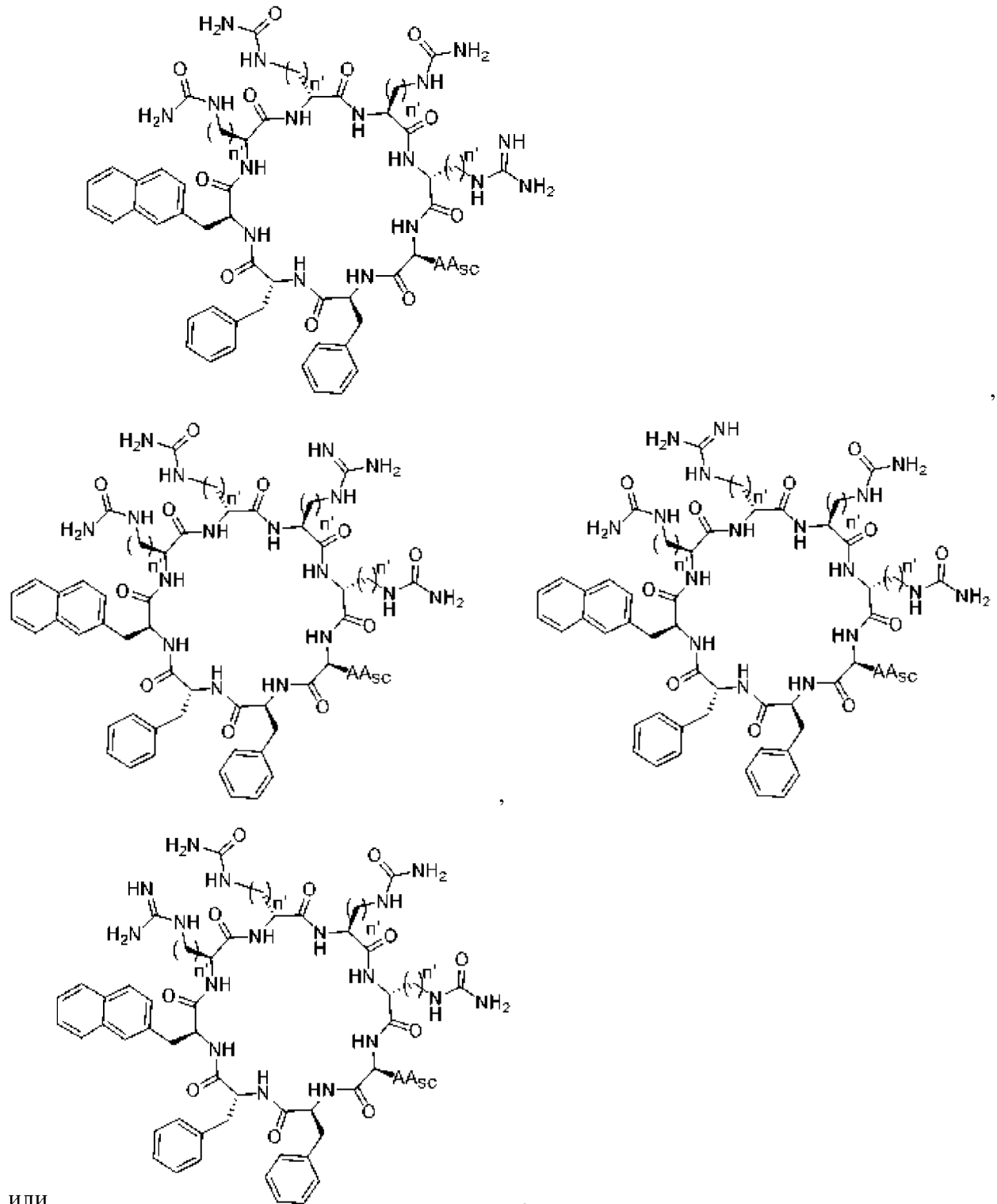
(Ic),

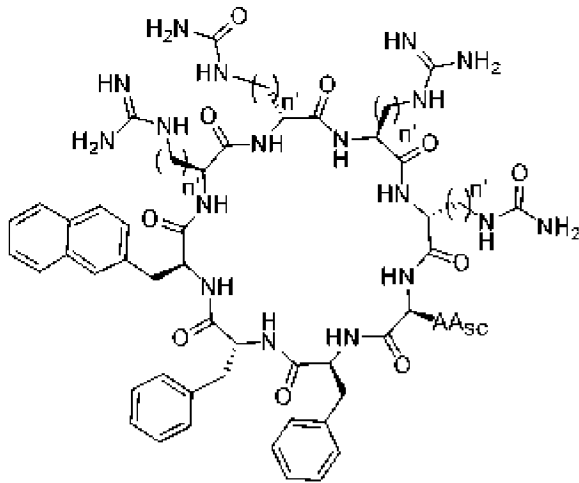
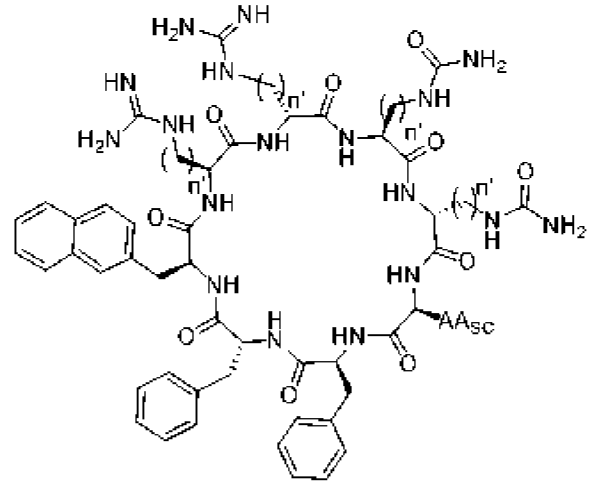
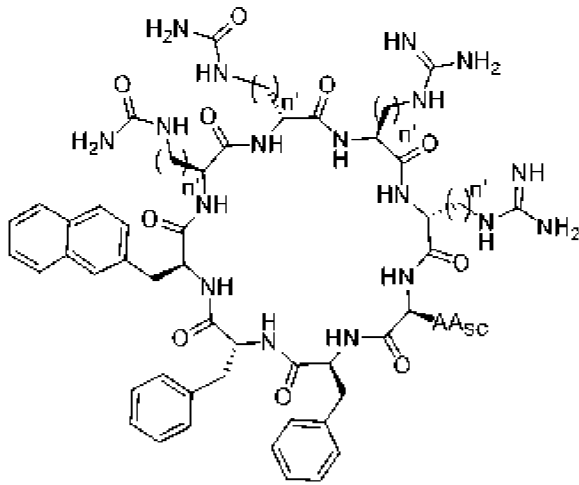
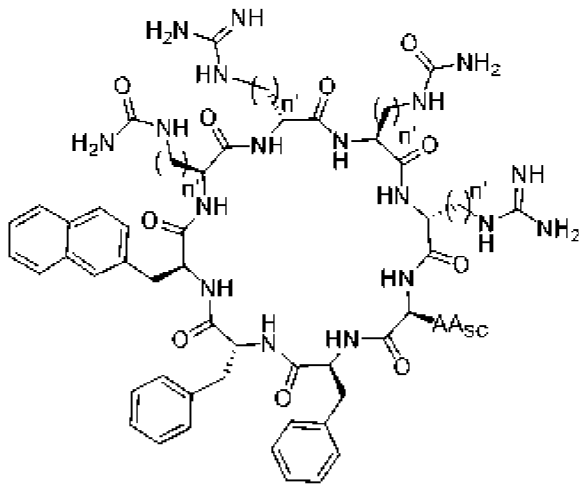
или ее протонированную форму,

где:

$AA_{5C}$  и  $n'$  определены в настоящем документе.

cCPR Формулы (Ia) имеет одну из следующих структур:

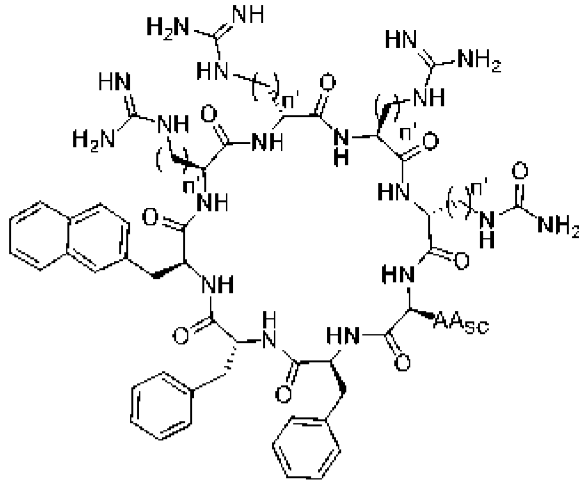
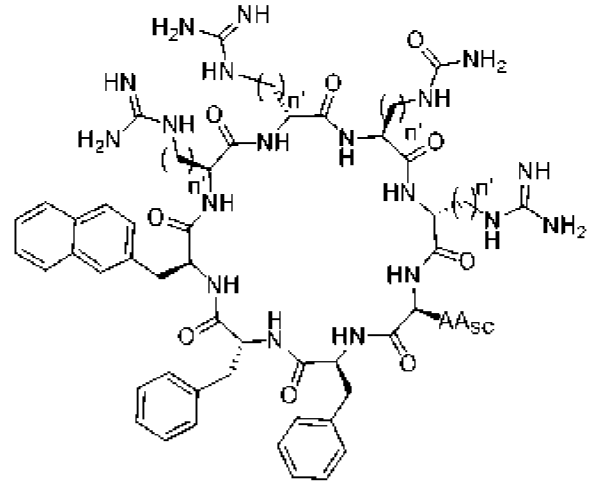
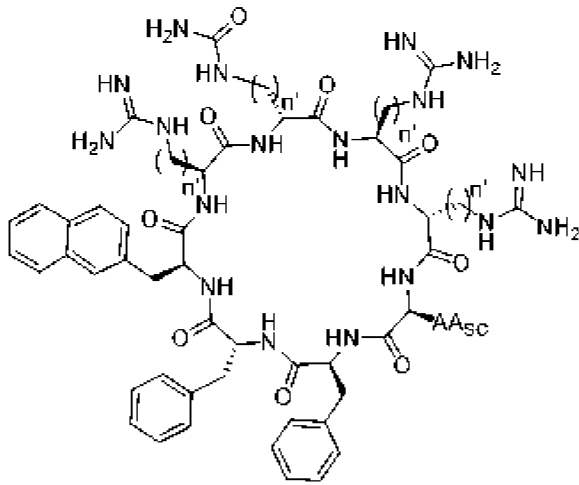
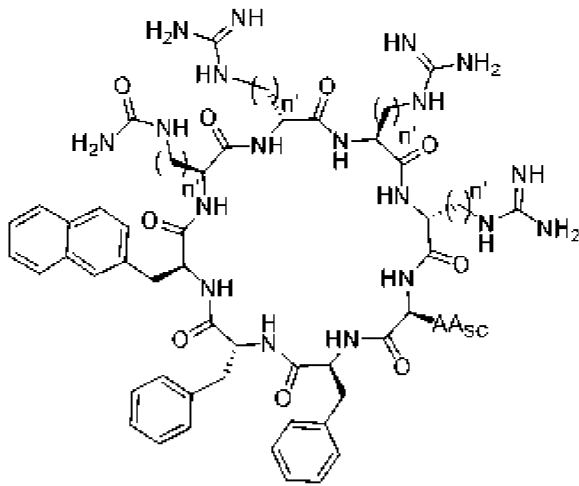




или

где AA<sub>Sc</sub> и n определены в настоящем документе

cCPR Формулы (IIa) имеет одну из следующих структур:

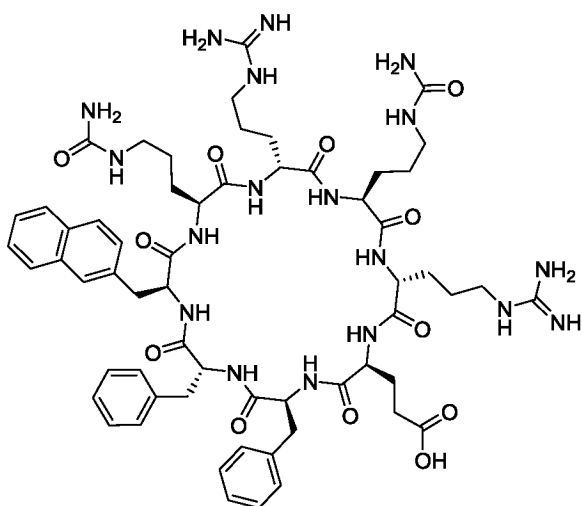


или  
документе.

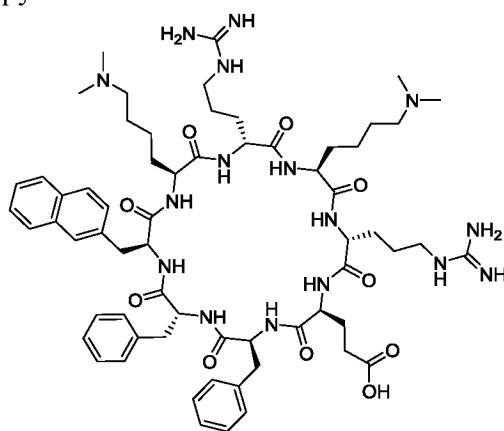
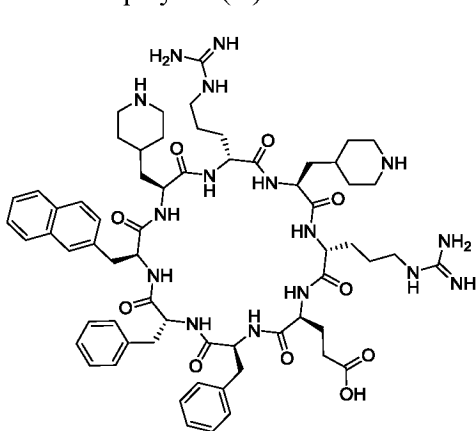
, где AA<sub>Sc</sub> и  $n$  определены в настоящем

cCPP Формулы (II) может иметь структуру:



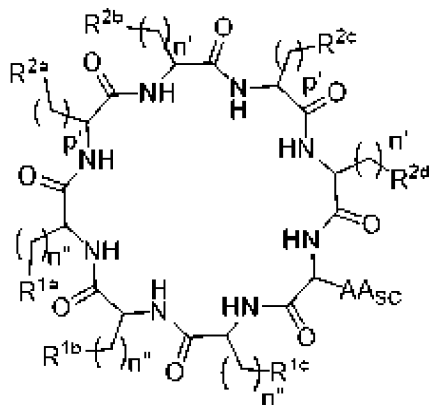


сСРР Формулы (II) может иметь структуру:



, или

сСРР может иметь структуру Формулы (III):

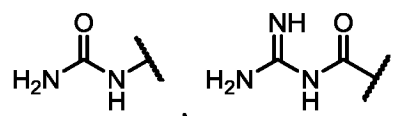


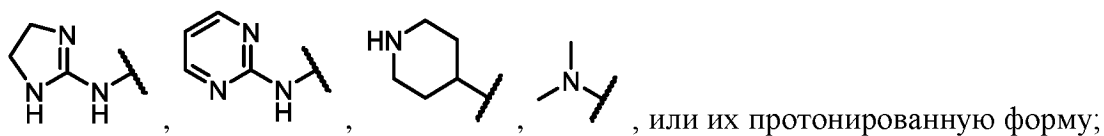
(III),

где:

AA<sub>sc</sub> представляет собой боковую цепь аминокислоты;  
каждый R<sup>1a</sup>, R<sup>1b</sup> и R<sup>1c</sup> независимо представляет собой 6-14-членный арил или 6-14-членный гетероарил;

каждый R<sup>2a</sup> и R<sup>2c</sup> независимо представляет собой H,





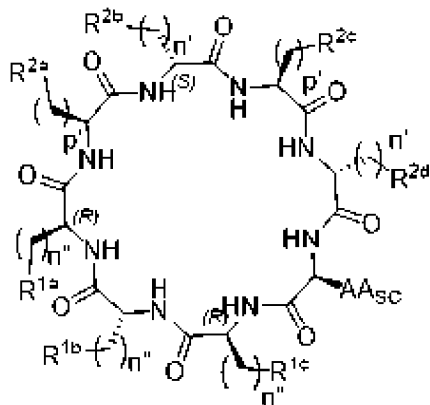
каждый  $R^{2b}$  и  $R^{2d}$  независимо представляет собой гуанидин или его протонированную форму;

каждый  $n^{cc}$  независимо представляет собой целое число от 1 до 3;

каждый  $n'$  независимо представляет собой целое число от 1 до 5; и

каждый  $p'$  независимо представляет собой целое число от 0 до 5.

cCPP Формулы (III) может иметь структуру Формулы (III-1):

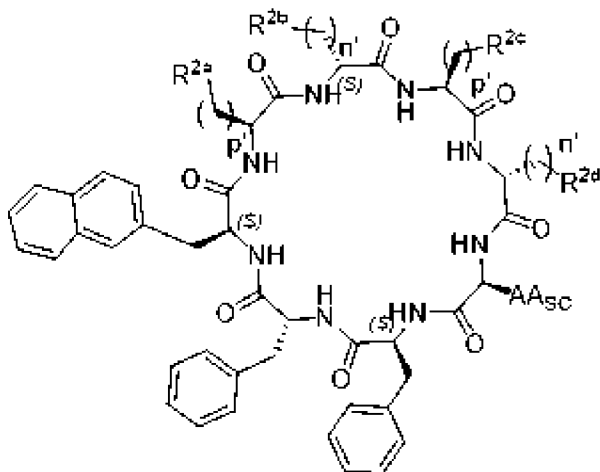


(III-1),

где:

$AA_{Sc}$ ,  $R^{1a}$ ,  $R^{1b}$ ,  $R^{1c}$ ,  $R^{2a}$ ,  $R^{2c}$ ,  $R^{2b}$ ,  $R^{2d}$ ,  $n'$ ,  $n^{cc}$  и  $p'$  определены в настоящем документе.

cCPP Формулы (III) может иметь структуру Формулы (IIIa):



(IIIa),

где:

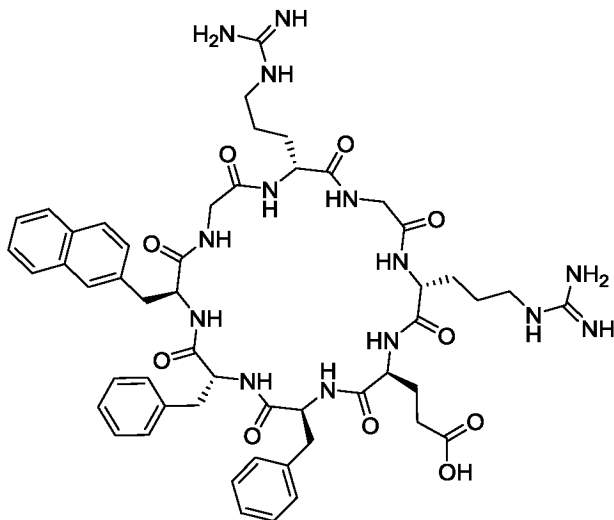
$AA_{Sc}$ ,  $R^{2a}$ ,  $R^{2c}$ ,  $R^{2b}$ ,  $R^{2d}$ ,  $n'$ ,  $n^{cc}$  и  $p'$  определены в настоящем документе.

В Формулах (III), (III-1) и (IIIa)  $R^a$  и  $R^c$  могут представлять собой H.  $R^a$  и  $R^c$  могут представлять собой H, и каждый  $R^b$  и  $R^d$  может независимо представлять собой гуанидин или его протонированную форму.  $R^a$  может представлять собой H.  $R^b$  может представлять собой H.  $p'$  может представлять собой 0.  $R^a$  и  $R^c$  могут представлять собой H, и каждый  $p'$  может представлять собой 0.

В Формулах (III), (III-1) и (IIIa)  $R^a$  и  $R^c$  могут представлять собой H, каждый  $R^b$  и  $R^d$  может независимо представлять собой гуанидин или его протонированную форму,  $n^c$  может представлять собой 2 или 3, и каждый  $p'$  может представлять собой 0.

$p'$  может представлять собой 0,  $p'$  может представлять собой 1,  $p'$  может представлять собой 2,  $p'$  может представлять собой 3,  $p'$  может представлять собой 4,  $p'$  может представлять собой 5.

cCPP может иметь структуру:



cCPP Формулы (A) может быть выбран из:

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

cCPP Формулы (A) может быть выбран из:

Последовательность CPP	SEQ ID NO:
FΦRRRRQ	86
fΦRrRrQ	87
FfΦRrRrQ	78
FfΦCit-r-Cit-rQ	79
FfΦGrGrQ	80
FfΦRGRGQ	88

FfFGRGRQ	81
FGFGRGRQ	82
GfFGrGrQ	83
FGFGRRRQ	84
FGFRRRRQ	85

Согласно вариантам реализации сCPP выбран из:

Последовательности CPP и 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		

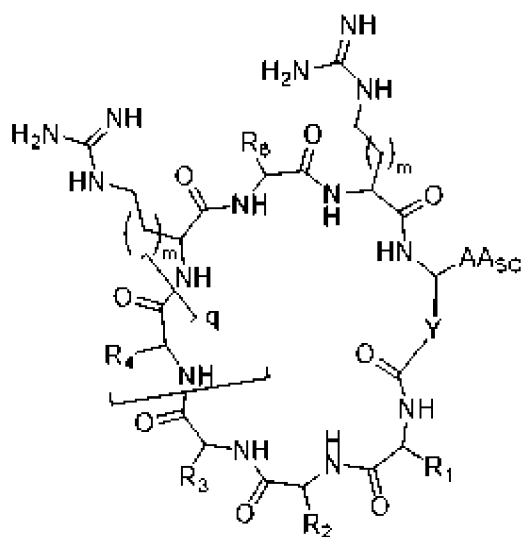
где Φ= L-нафтилаланин; φ= D-нафтилаланин; Ω=L-норлейцин

Согласно вариантам реализации сCPP не выбран из:

Последовательности CPP и 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		

где Φ= L-нафтилаланин; φ= D-нафтилаланин; Ω=L-норлейцин

сСРР может содержать структуру Формулы (D)



(D)

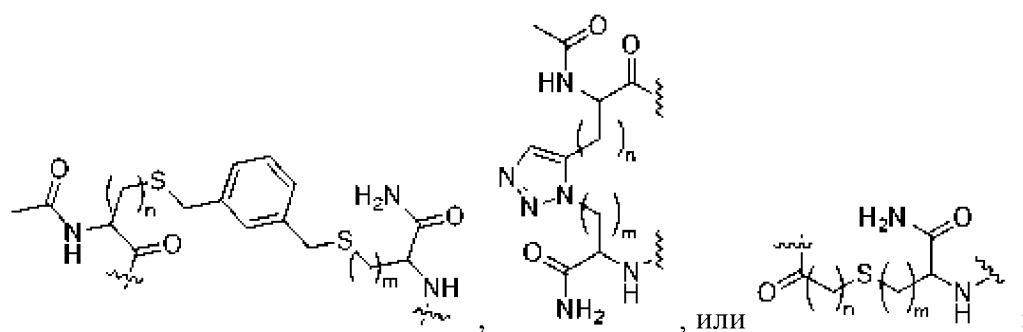
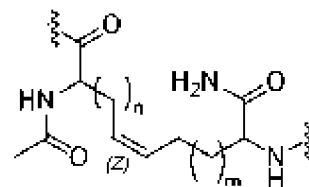
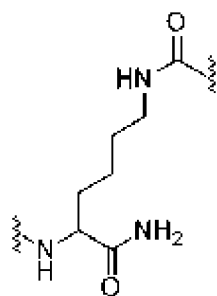
или ее протонированную форму, где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

Y представляет собой



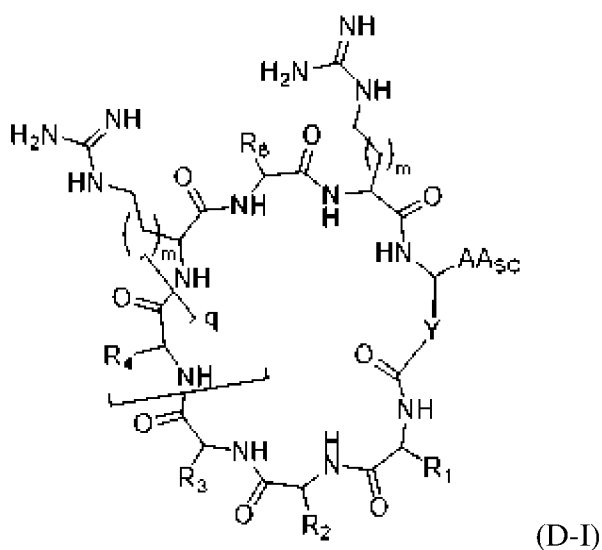
$AA_{3C}$  представляет собой боковую цепь аминокислоты;

q представляет собой 1, 2, 3 или 4;

каждый m независимо представляет собой целое число 0, 1, 2 или 3, и

каждый n независимо представляет собой целое число 0, 1, 2 или 3.

сСРР Формулы (D) может иметь структуру Формулы (D-I):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

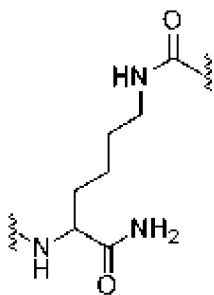
$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

$AA_{5C}$  представляет собой боковую цепь аминокислоты;

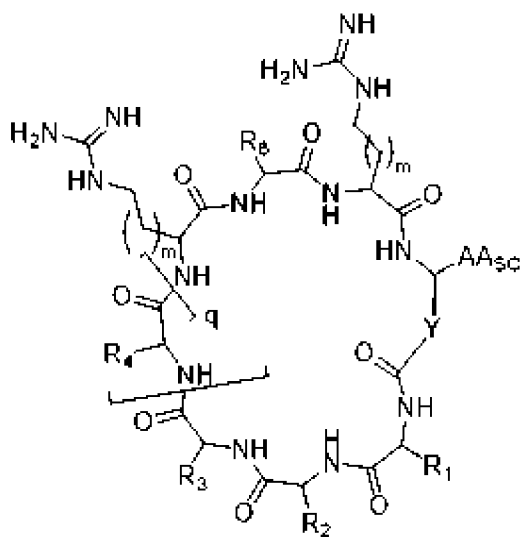
$q$  представляет собой 1, 2, 3 или 4;

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3, и

$Y$  представляет собой



cCPP Формулы (D) может иметь структуру Формулы (D-II):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

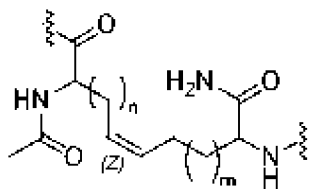
$AA_{5C}$  представляет собой боковую цепь аминокислоты;

$q$  представляет собой 1, 2, 3 или 4;

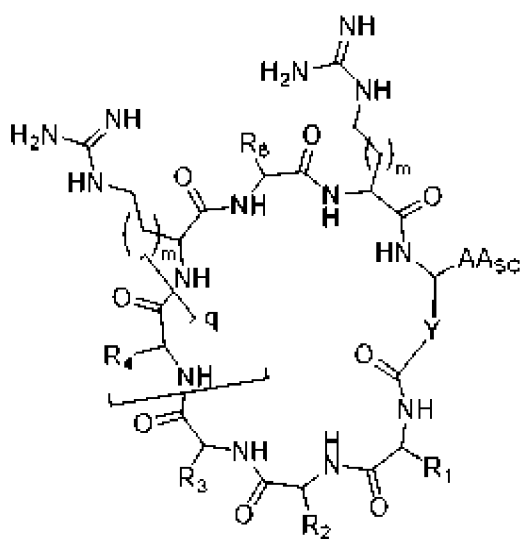
каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3,

каждый  $n$  независимо представляет собой целое число 0, 1, 2 или 3, и

$Y$  представляет собой



cSPR Формулы (D) может иметь структуру Формулы (D-III):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

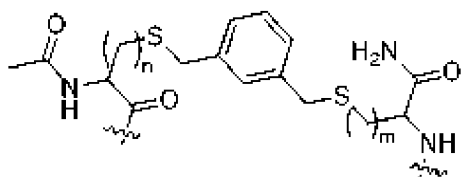
$AA_{3C}$  представляет собой боковую цепь аминокислоты;

$q$  представляет собой 1, 2, 3 или 4;

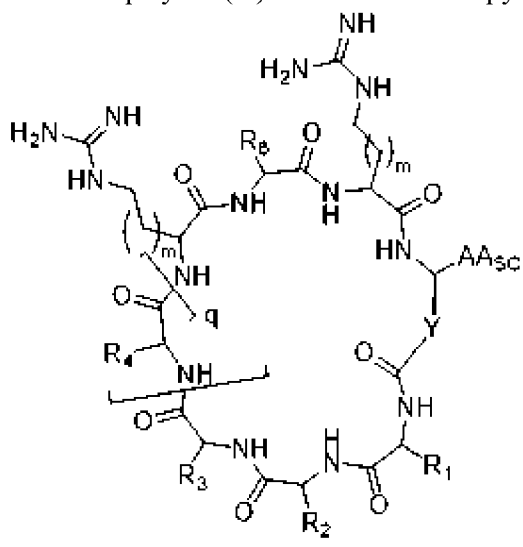
каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3,

каждый  $n$  независимо представляет собой целое число 0, 1, 2 или 3, и

$Y$  представляет собой



cCPP Формулы (D) может иметь структуру Формулы (D-IV):





или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

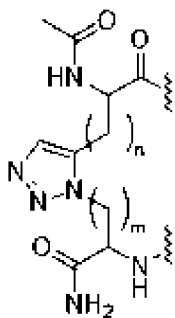
$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

$AA_{SC}$  представляет собой боковую цепь аминокислоты;

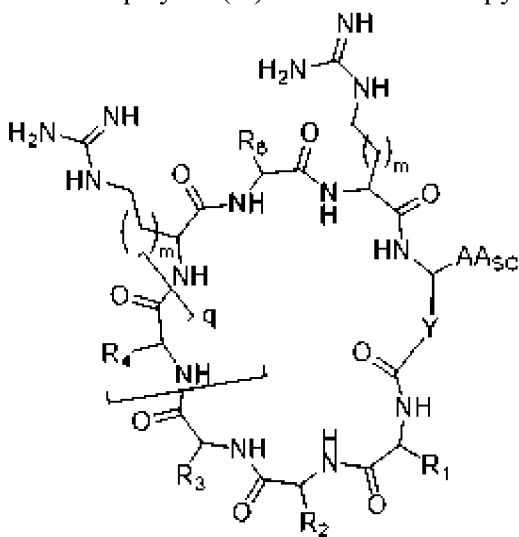
$q$  представляет собой 1, 2, 3 или 4;

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3, и

$Y$  представляет собой



cCPP Формулы (D) может иметь структуру Формулы (D-V):



(D-V)

или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

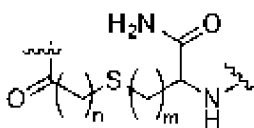
по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  и  $R_6$  независимо представляют собой H или боковую цепь аминокислоты;

$AA_{SC}$  представляет собой боковую цепь аминокислоты;

$q$  представляет собой 1, 2, 3 или 4;

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3, и  $Y$  представляет собой



$AA_{SC}$  может быть конъюгирован с линкером.

### *Линкер*

cCPR согласно настоящему изобретению может быть конъюгирован с линкером. Линкер может связывать груз с cCPR. Линкер может быть присоединен к боковой цепи аминокислоты cCPR, и груз может быть присоединен в подходящем положении на линкере.

Линкер может представлять собой любой подходящий фрагмент, с помощью которого можно конъюгировать cCPR с одним или более дополнительными фрагментами, например, экзоциклическим пептидом (EP) и/или грузом. Перед конъюгацией с cCPR и одним или более дополнительными фрагментами линкер содержит две или более функциональных групп, каждая из которых независимо способна образовывать ковалентную связь с cCPR и одним или более дополнительными фрагментами. Если груз представляет собой олигонуклеотид, линкер может быть ковалентно связан с 5'-концом груза или 3'-концом груза. Линкер может быть ковалентно связан с 5'-концом груза. Линкер может быть ковалентно связан с 3'-концом груза. Если груз представляет собой пептид, линкер может быть ковалентно связан с N-концом или C-концом груза. Линкер может быть ковалентно связан с остовом олигонуклеотида или пептидного груза. Линкер может представлять собой любой подходящий фрагмент, с помощью которого конъюгируют cCPR, описанный в настоящем документе, с грузом, таким как олигонуклеотид, пептид или малая молекула.

Линкер может содержать углеводородный линкер.

Линкер может содержать сайт расщепления. Сайт расщепления может представлять собой дисульфид или сайт расщепления каспазой (например, Val-Cit-PAVC).

Линкер может содержать: (i) одну или более D- или L-аминокислот, каждая из которых необязательно замещена; (ii) необязательно замещенный алкилен; (iii) необязательно замещенный алкенилен; (iv) необязательно замещенный алкинилен; (v) необязательно замещенный карбоциклл; (vi) необязательно замещенный гетероциклл; (vii) одну или более субъединиц  $-(R^1-J-R^2)z^{c'}$ , где каждый из  $R^1$  и  $R^2$ , в каждом случае, независимо выбран из алкилена, алкенилена, алкинилена, карбоциклила и гетероциклила, каждый J независимо представляет собой C,  $NR^3$ ,  $-NR^3C(O)-$ , S и O, где  $R^3$  независимо выбран из H, алкила, алкенила, алкинила, карбоциклила и гетероциклила, каждый из которых необязательно замещен, и  $z^{c'}$  представляет собой целое число от 1 до 50; (viii)  $(R^1-J)z^{c'}$  или  $-(J-R^1)z^{c'}$ , где каждый из  $R^1$ , в каждом случае, независимо представляет собой алкилен, алкенилен, алкинилен, карбоциклл или гетероциклл, каждый J независимо представляет собой C,  $NR^3$ ,  $-NR^3C(O)-$ , S или O, где  $R^3$  представляет собой H,

алкил, алкенил, алкинил, карбоцикллил или гетероцикллил, каждый из которых необязательно замещен, и  $z^{cc}$  представляет собой целое число от 1 до 50; или (ix) линкер может содержать одно или более из (i) - (x).

Линкер может содержать одну или более D или L-аминокислот и/или  $-(R^1-J-R^2)z^{cc}$ -, где каждый из  $R^1$  и  $R^2$ , в каждом случае, независимо представляет собой алкилен, каждый J независимо представляет собой C,  $NR^3$ ,  $-NR^3C(O)-$ , S и O, где  $R^4$  независимо выбран из H и алкила, и  $z^{cc}$  представляет собой целое число от 1 до 50; или их комбинации.

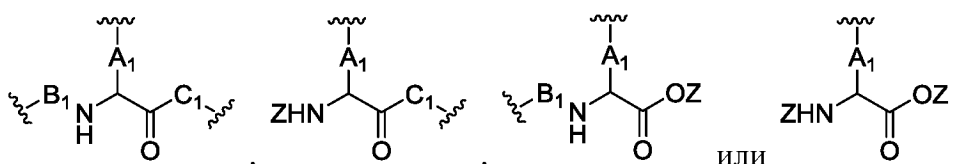
Линкер может содержать  $-(OCH_2CH_2)_{z'}$ - (например, в виде спейсера), где  $z'$  представляет собой целое число от 1 до 23, например, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 или 23.  $-(OCH_2CH_2)_{z'}$  может также называться полиэтиленгликолем (PEG).

Линкер может содержать одну или более аминокислот. Линкер может содержать пептид. Линкер может содержать  $-(OCH_2CH_2)_{z'}$ -, где  $z'$  представляет собой целое число от 1 до 23, и пептид. Пептид может содержать от 2 до 10 аминокислот. Линкер может дополнительно содержать функциональную группу (FG), способную взаимодействовать посредством клик-химии. FG может представлять собой азид или алкин, и триазол образуется, когда груз конъюгирован с линкером.

Линкер может содержать (i) остаток  $\beta$ -аланина и остаток лизина; (ii)  $-(J-R^1)z^{cc}$ ; или (iii) их комбинацию. Каждый  $R^1$  может независимо представлять собой алкилен, алкенилен, алкинилен, карбоцикллил или гетероцикллил, каждый J независимо представляет собой C,  $NR^3$ ,  $-NR^3C(O)-$ , S или O, где  $R^3$  представляет собой H, алкил, алкенил, алкинил, карбоцикллил или гетероцикллил, каждый из которых необязательно замещен, и  $z^{cc}$  может представлять собой целое число от 1 до 50. Каждый  $R^1$  может представлять собой алкилен, и каждый J может представлять собой O.

Линкер может содержать (i) остатки  $\beta$ -аланина, глицина, лизина, 4-аминомасляной кислоты, 5-аминопентановой кислоты, 6-аминогексановой кислоты или их комбинации; и (ii)  $-(R^1-J)z^{cc}$ - или  $-(J-R^1)z^{cc}$ . Каждый  $R^1$  может независимо представлять собой алкилен, алкенилен, алкинилен, карбоцикллил или гетероцикллил, каждый J независимо представляет собой C,  $NR^3$ ,  $-NR^3C(O)-$ , S или O, где  $R^3$  представляет собой H, алкил, алкенил, алкинил, карбоцикллил или гетероцикллил, каждый из которых необязательно замещен, и  $z^{cc}$  может представлять собой целое число от 1 до 50. Каждый  $R^1$  может представлять собой алкилен, и каждый J может представлять собой O. Линкер может содержать глицин, бета-аланин, 4-аминомасляную кислоту, 5-аминопентановую кислоту, 6-аминогексановую кислоту или их комбинацию.

Линкер может представлять собой трехвалентный линкер. Линкер может иметь структуру:



где  $A_1$ ,  $B_1$  и  $C_1$  могут независимо представлять собой углеводородный линкер (например,  $\text{NRH}-(\text{CH}_2)_n-\text{COOH}$ ), PEG-линкер (например,  $\text{NRH}-(\text{CH}_2\text{O})_n-\text{COOH}$ , где R представляет собой H, метил или этил) или один или более аминокислотных остатков, и Z независимо представляет собой защитную группу. Линкер также может включать сайт расщепления, включая дисульфид  $[\text{NH}_2-(\text{CH}_2\text{O})_n-\text{S}-\text{S}-(\text{CH}_2\text{O})_n-\text{COOH}]$  или сайт расщепления каспазой (Val-Cit-PABC).

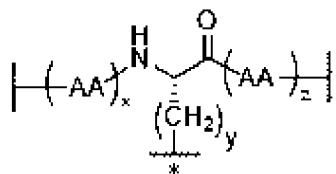
Углеводород может представлять собой остаток глицина или бета-аланина.

Линкер может быть двухвалентным и связывать сСРР с грузом. Линкер может быть двухвалентным и связывать сСРР с экзоциклическим пептидом (EP).

Линкер может быть трехвалентным и связывать сСРР с грузом и с EP.

Линкер может представлять собой двухвалентный или трехвалентный  $C_1$ - $C_{50}$  алкилен, в котором 1-25 метиленовых групп необязательно и независимо заменены -N(H)-, -N( $C_1$ - $C_4$  алкил)-, -N(циклоалкил)-, -O-, -C(O)-, -C(O)O-, -S-, -S(O)-, -S(O)<sub>2</sub>-, -S(O)<sub>2</sub>N( $C_1$ - $C_4$  алкил)-, -S(O)<sub>2</sub>N(циклоалкил)-, -N(H)C(O)-, -N( $C_1$ - $C_4$  алкил)C(O)-, -N(циклоалкил)C(O)-, -C(O)N(H)-, -C(O)N( $C_1$ - $C_4$  алкил), -C(O)N(циклоалкил), арилом, гетероциклилом, гетероарилом, циклоалкилом или циклоалкенилом. Линкер может представлять собой двухвалентный или трехвалентный  $C_1$ - $C_{50}$  алкилен, в котором 1-25 метиленовых групп необязательно и независимо заменены -N(H)-, -O-, -C(O)N(H)- или их комбинацией.

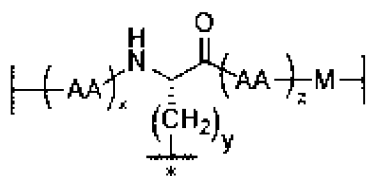
Линкер может иметь структуру:



где: каждый AA независимо представляет собой аминокислотный остаток; \* представляет собой положение присоединения к  $\text{AA}_{\text{SC}}$ , и  $\text{AA}_{\text{SC}}$  представляет собой боковую цепь аминокислотного остатка сСРР; x представляет собой целое число от 1 до 10; y представляет собой целое число от 1 до 5; и z представляет собой целое число от 1 до 10. x может представлять собой целое число от 1 до 5. x может представлять собой целое число от 1 до 3. x может представлять собой 1. y может представлять собой целое число от 2 до 4. y может представлять собой 4. z может представлять собой целое число от 1 до 5. z может представлять собой целое число от 1 до 3. z может представлять собой 1. Каждый AA может быть независимо выбран из глицина,  $\beta$ -аланина, 4-аминомасляной кислоты, 5-аминопентановой кислоты и 6-аминогексановой кислоты.

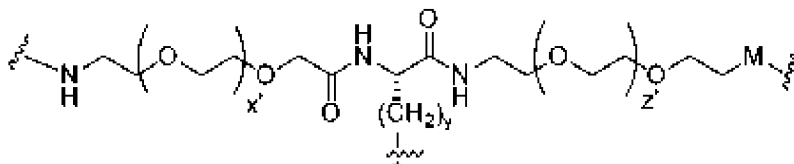
сСРР может быть присоединен к грузу посредством линкера («L»). Линкер может быть конъюгирован с грузом посредством связывающей группы («M»).

Линкер может иметь структуру:



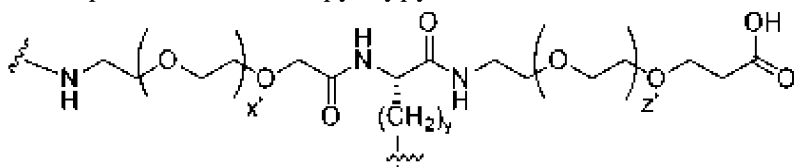
где  $x$  представляет собой целое число от 1 до 10;  $y$  представляет собой целое число от 1 до 5;  $z$  представляет собой целое число от 1 до 10; каждый AA независимо представляет собой аминокислотный остаток; \* представляет собой положение присоединения к AA<sub>SC</sub>, и AA<sub>SC</sub> представляет собой боковую цепь аминокислотного остатка сCPP; и M представляет собой связывающую группу, определенную в настоящем документе.

Линкер может иметь структуру:



где:  $x'$  представляет собой целое число от 1 до 23;  $y$  представляет собой целое число от 1 до 5;  $z'$  представляет собой целое число от 1 до 23; \* представляет собой положение присоединения к AA<sub>SC</sub>, и AA<sub>SC</sub> представляет собой боковую цепь аминокислотного остатка сCPP; и M представляет собой связывающую группу, определенную в настоящем документе.

Линкер может иметь структуру:



где:  $x'$  представляет собой целое число от 1 до 23;  $y$  представляет собой целое число от 1 до 5; и  $z'$  представляет собой целое число от 1 до 23; \* представляет собой положение присоединения к AA<sub>SC</sub>, и AA<sub>SC</sub> представляет собой боковую цепь аминокислотного остатка сCPP.

$x$  может представлять собой целое число от 1 до 10, например, 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10, включая все промежуточные диапазоны и поддиапазоны.

$x'$  может представлять собой целое число от 1 до 23, например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 или 23, включая все диапазоны и поддиапазоны между ними.  $x'$  может представлять собой целое число от 5 до 15.  $x'$  может представлять собой целое число от 9 до 13.  $x'$  может представлять собой целое число от 1 до 5.  $x'$  может представлять собой 1.

у может представлять собой целое число от 1 до 5, например, 1, 2, 3, 4 или 5, включая все диапазоны и поддиапазоны между ними. у может представлять собой целое число от 2 до 5. у может представлять собой целое число от 3 до 5. у может представлять собой 3 или 4. у может представлять собой 4 или 5. у может представлять собой 3. у может представлять собой 4. у может представлять собой 5.

z может представлять собой целое число от 1 до 10, например, 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10, включая все диапазоны и поддиапазоны между ними.

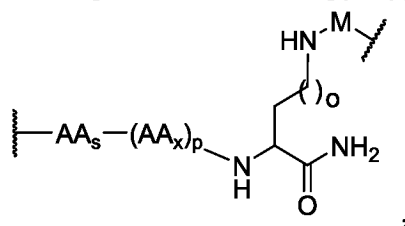
z' может представлять собой целое число от 1 до 23, например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 или 23, включая все диапазоны и поддиапазоны между ними. z' может представлять собой целое число от 5 до 15. z' может представлять собой целое число от 9 до 13. z' может представлять собой 11.

Как обсуждалось выше, линкер или М (где М представляет собой часть линкера) может быть ковалентно связан с грузом в любом подходящем месте на грузе. Линкер или М (где М представляет собой часть линкера) может быть ковалентно связан с 3'-концом олигонуклеотидного груза или 5'-концом олигонуклеотидного груза. Линкер или М (где М представляет собой часть линкера) может быть ковалентно связан с N-концом или С-концом пептидного груза. Линкер или М (где М представляет собой часть линкера) может быть ковалентно связан с остовом олигонуклеотида или пептидного груза.

Линкер может быть связан с боковой цепью аспарагиновой кислоты, глутаминовой кислоты, глутамина, аспарагина или лизина или модифицированной боковой цепью глутамина или аспарагина (например, с восстановленной боковой цепью, имеющей аминогруппу) на сСРР. Линкер может быть связан с боковой цепью лизина на сСРР.

Линкер может быть связан с боковой цепью аспарагиновой кислоты, глутаминовой кислоты, глутамина, аспарагина или лизина или модифицированной боковой цепью глутамина или аспарагина (например, с восстановленной боковой цепью, имеющей аминогруппу) на пептидном грузе. Линкер может быть связан с боковой цепью лизина на пептидном грузе.

Линкер может иметь структуру:



где

М представляет собой группу, с помощью которой конъюгируют L с грузом, например, олигонуклеотидом;

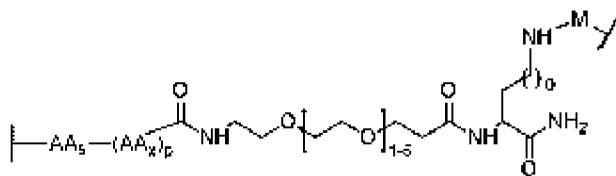
AA<sub>s</sub> представляет собой боковую цепь или конец аминокислоты на сСРР;

каждый AA<sub>x</sub> независимо представляет собой аминокислотный остаток;

o представляет собой целое число от 0 до 10; и

p представляет собой целое число от 0 до 5.

Линкер может иметь структуру:



где

M представляет собой группу, с помощью которой конъюгируют L с грузом, например, олигонуклеотидом;

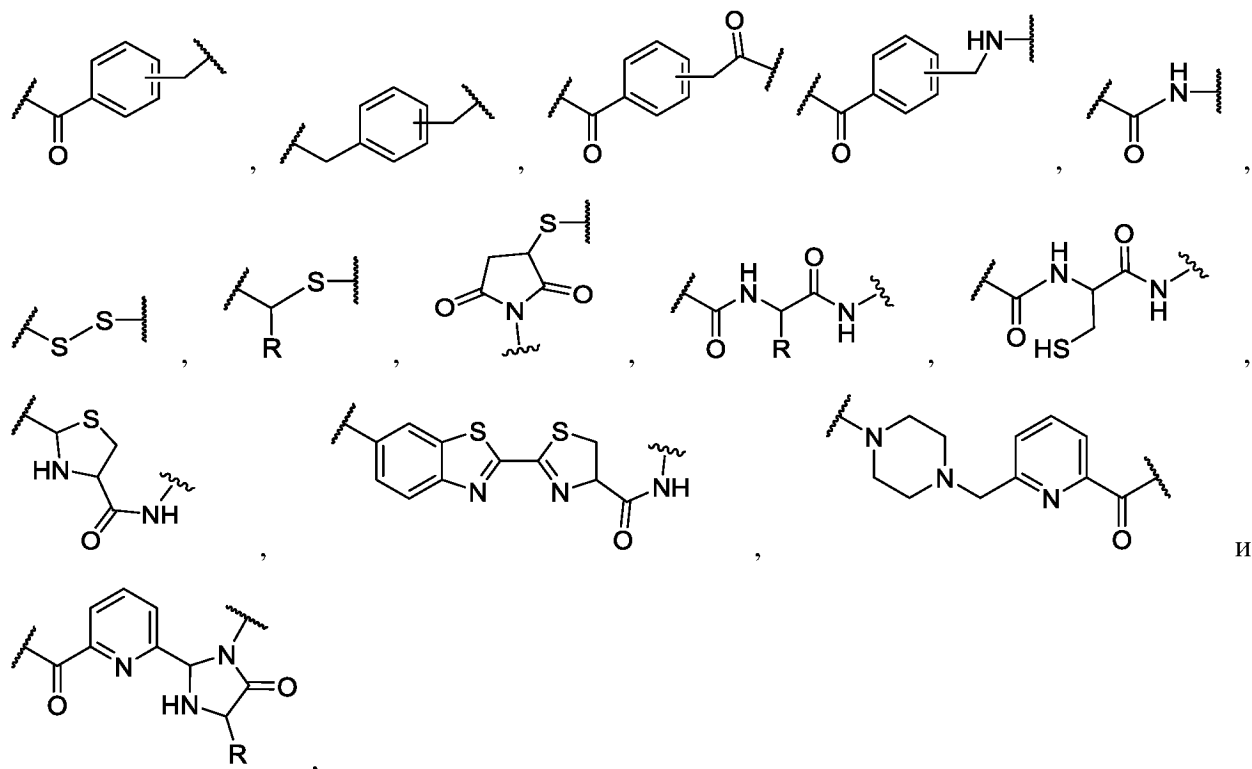
AA<sub>5</sub> представляет собой боковую цепь или конец аминокислоты на сСРР;

каждый AA<sub>x</sub> независимо представляет собой аминокислотный остаток;

o представляет собой целое число от 0 до 10; и

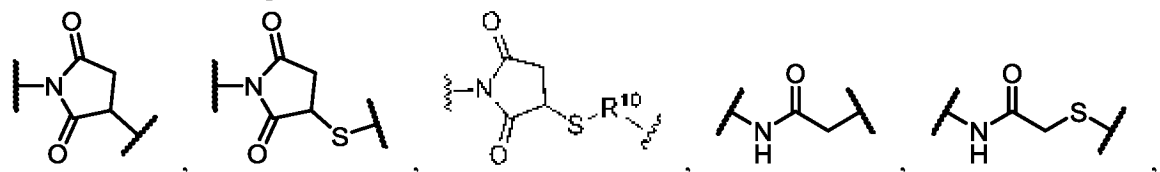
r представляет собой целое число от 0 до 5.

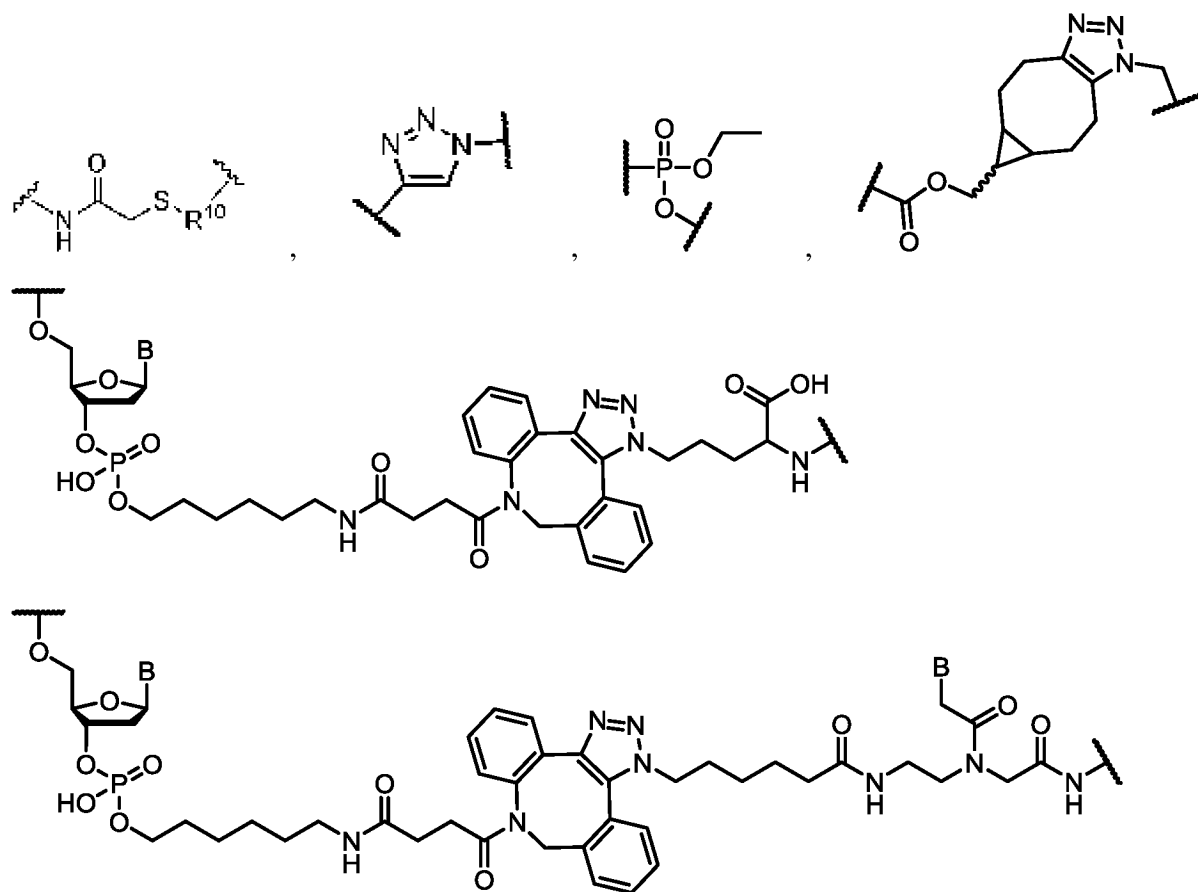
M может содержать алкилен, алкенилен, алкинилен, карбоциклил или гетероциклил, каждый из которых необязательно замещен. M может быть выбран из:



где R представляет собой алкил, алкенил, алкинил, карбоциклил или гетероциклил.

M может быть выбран из:

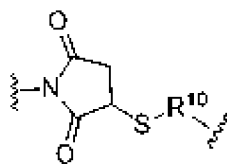




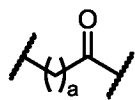
и

где:  $R^{10}$  представляет собой алкилен, циклоалкил или , где  $a$  представляет собой от 0 до 10.

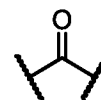
$M$  может представлять собой



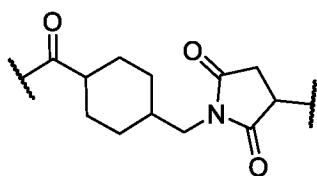
,  $R^{10}$  может представлять собой



, и  $a$  представляет собой от 0 до 10.  $M$  может представлять собой



$M$  может представлять собой гетеробифункциональный сшивающий агент,



например, , который раскрыт в Williams et al. *Curr. Protoc Nucleic Acid Chem.* **2010**, 42, 4.41.1-4.41.20, полностью включенной в настоящий документ посредством ссылки.

$M$  может представлять собой  $-C(O)-$ .



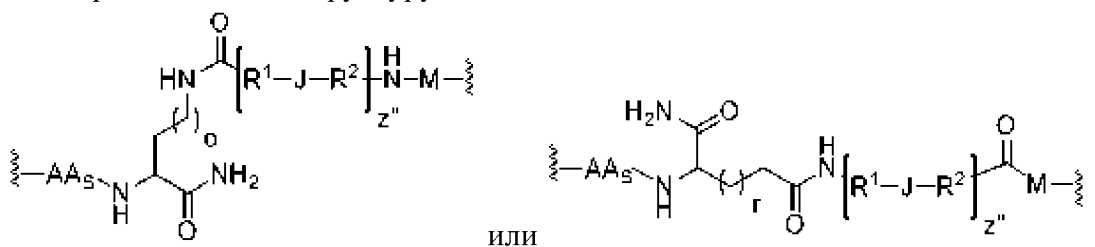
AA<sub>s</sub> может представлять собой боковую цепь или конец аминокислоты на сСРР. Неограничивающие примеры AA<sub>s</sub> включают аспарагиновую кислоту, глутаминовую кислоту, глутамин, аспарагин или лизин или модифицированную боковую цепь глутамина или аспарагина (например, восстановленную боковую цепь, содержащую аминогруппу). AA<sub>s</sub> может представлять собой AA<sub>SC</sub>, определенный в настоящем документе.

Каждый AA<sub>x</sub> независимо представляет собой природную или неприродную аминокислоту. Один или более AA<sub>x</sub> могут представлять собой природную аминокислоту. Один или более AA<sub>x</sub> могут представлять собой неприродную аминокислоту. Один или более AA<sub>x</sub> могут представлять собой β-аминокислоту. β-аминокислота может представлять собой β-аланин.

o может представлять собой целое число от 0 до 10, например, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 и 10. o может представлять собой 0, 1, 2 или 3. o может представлять собой 0. o может представлять собой 1. o может представлять собой 2. o может представлять собой 3.

r может представлять собой 0-5, например, 0, 1, 2, 3, 4 или 5. r может представлять собой 0. r может представлять собой 1. r может представлять собой 2. r может представлять собой 3. r может представлять собой 4. r может представлять собой 5.

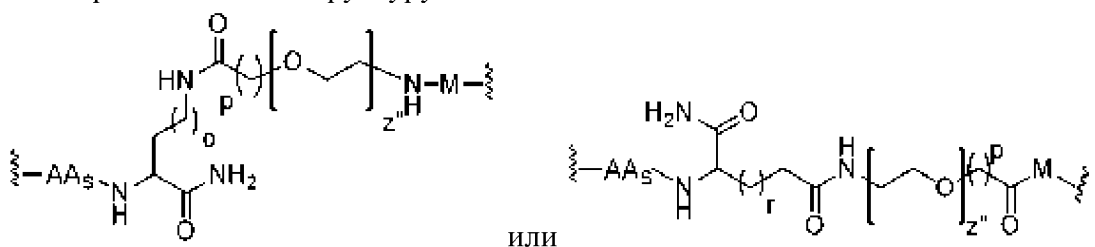
Линкер может иметь структуру:



где M, AA<sub>s</sub>, каждый  $-(R^1-J-R^2)z''$ , o и z'' определены в настоящем документе; r может представлять собой 0 или 1.

г может представлять собой 0. г может представлять собой 1.

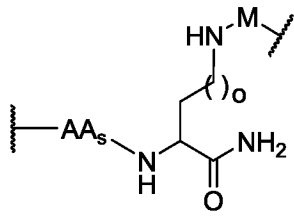
Линкер может иметь структуру:



где каждый M, AA<sub>s</sub>, o, p, q, г и z'' может быть таким, как определено в настоящем документе.

z'' может представлять собой целое число от 1 до 50, например, 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 и 50, включая все диапазоны и значения между ними. z'' может представлять собой целое число от 5 до 20. z'' может представлять собой целое число от 10 до 15.

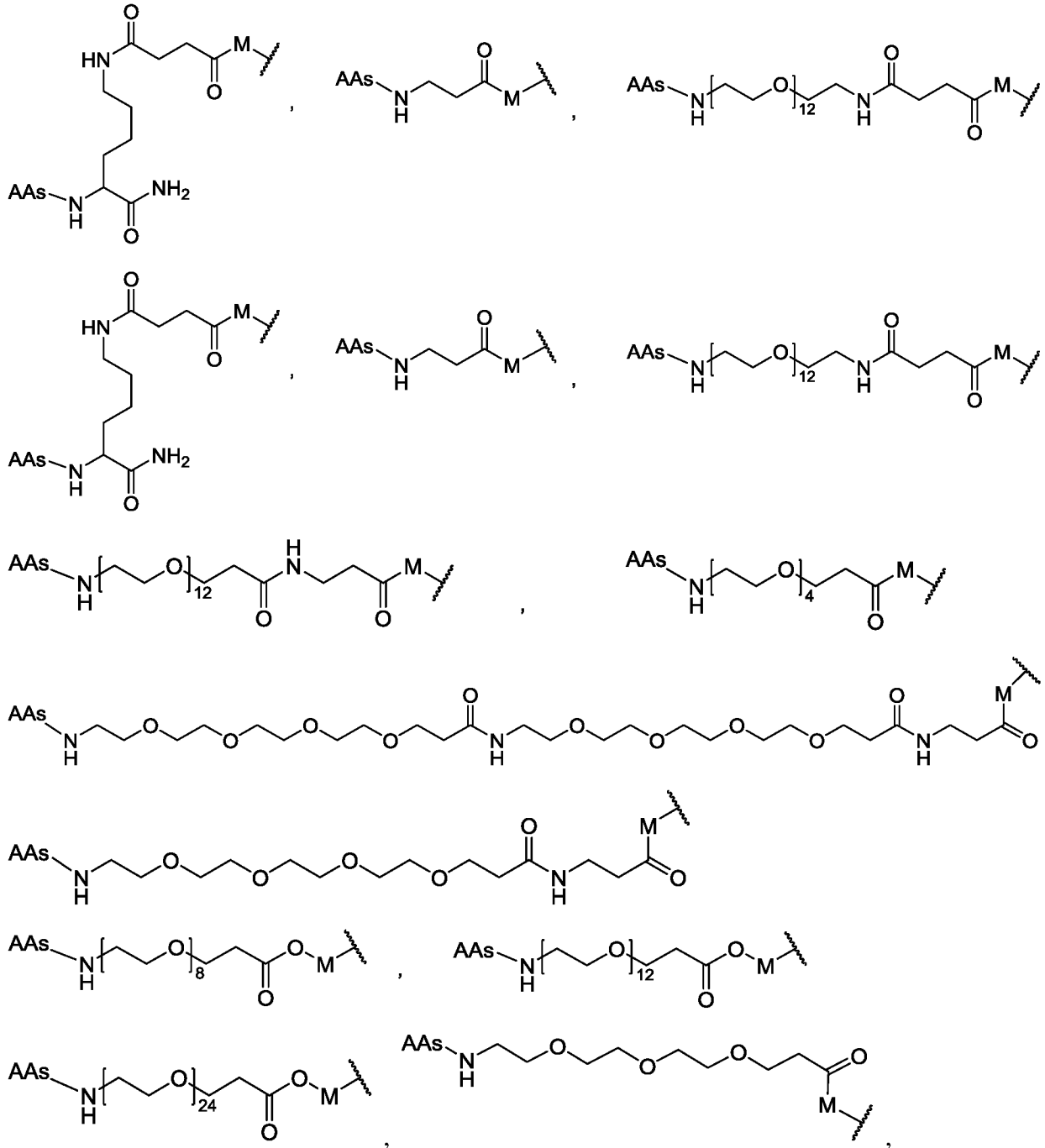
Линкер может иметь структуру:

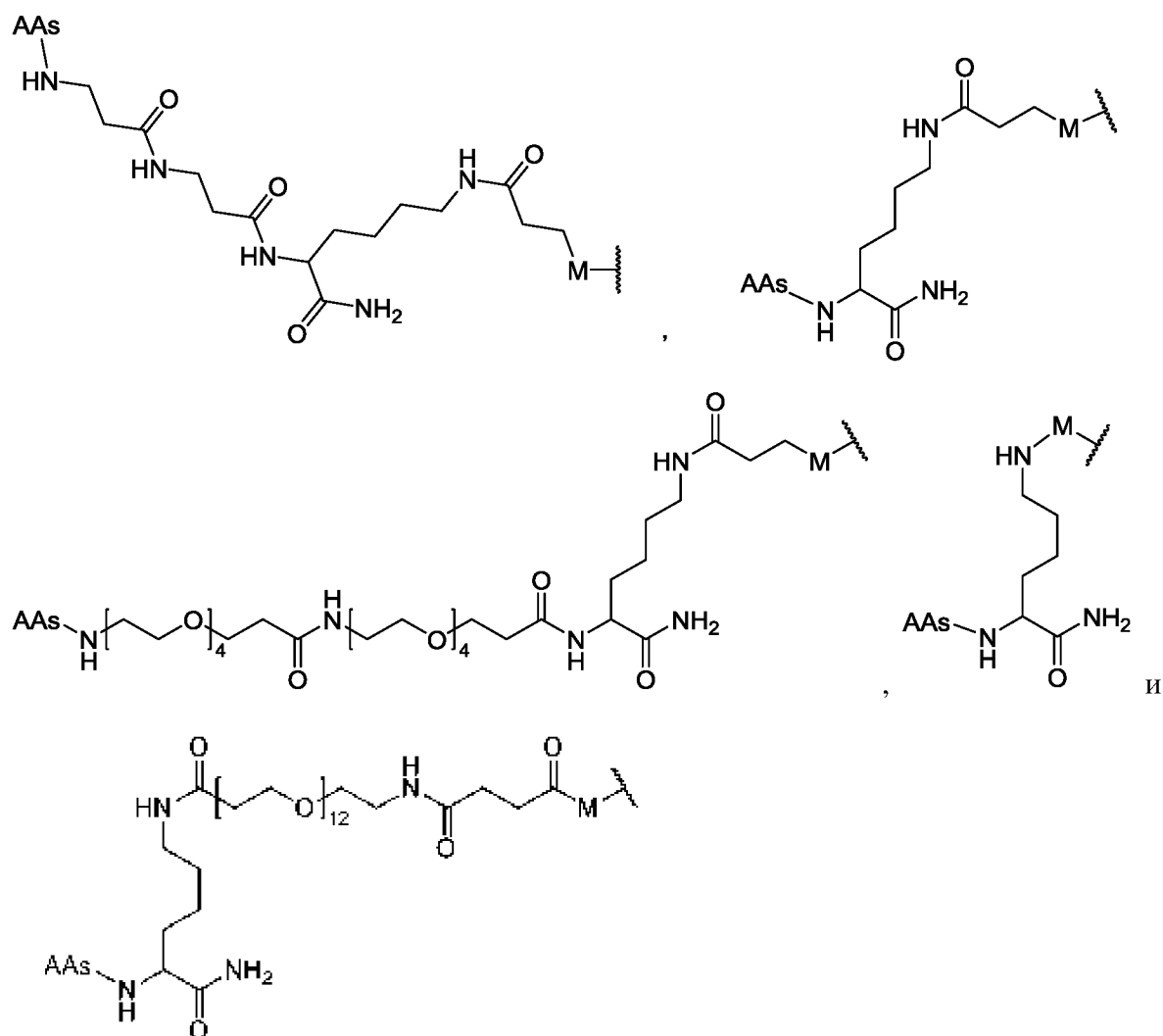


где:

$M$ ,  $\text{AA}_s$  и  $o$  определены в настоящем документе.

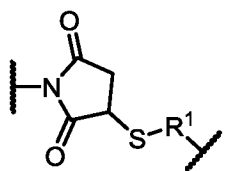
Другие неограничивающие примеры подходящих линкеров включают:



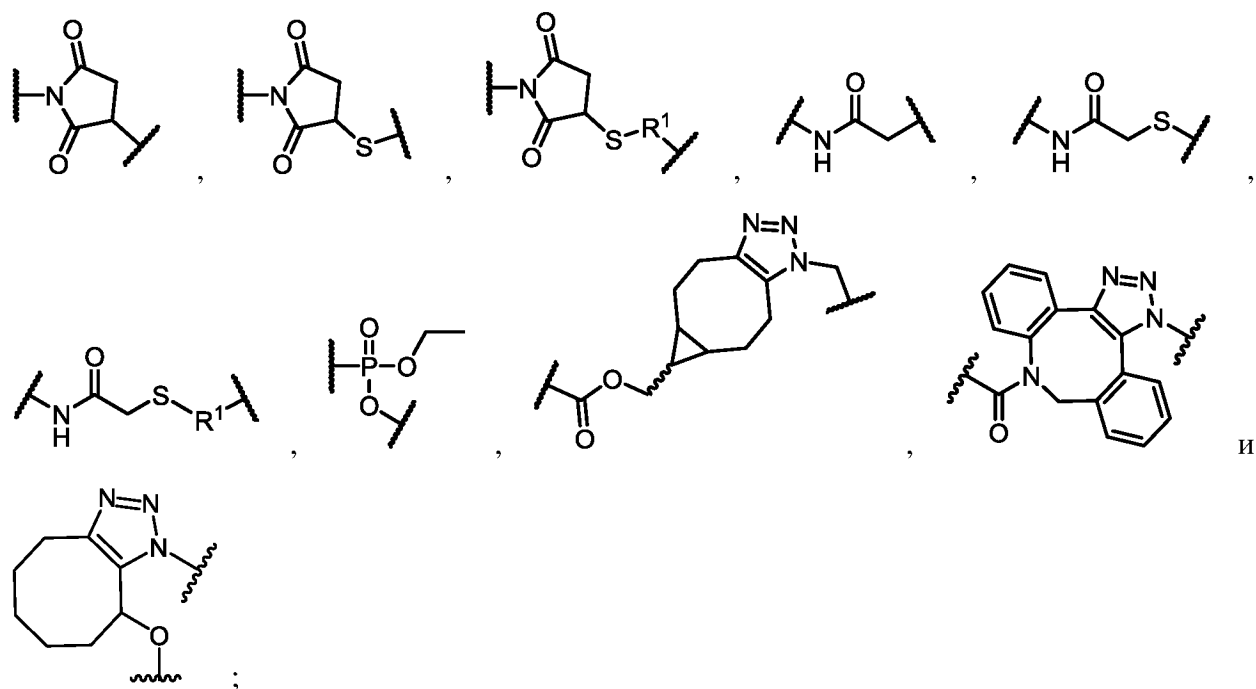


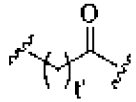
где M и AA<sub>s</sub> определены в настоящем документе.

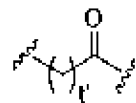
Согласно настоящему изобретению предложено соединение, содержащее сСРР и АС, который комплементарен мишени в последовательности пре-мРНК, дополнительно содержащее L, в котором линкер конъюгирован с АС посредством связывающей группы (M), где M представляет собой



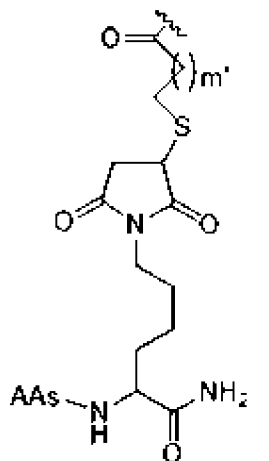
Согласно настоящему изобретению предложено соединение, содержащее сСРР и груз, который содержит антисмысловое соединение (АС), например, антисмысловый олигонуклеотид, который комплементарен мишени в последовательности пре-мРНК, причем указанное соединение дополнительно содержит L, при этом указанный линкер конъюгирован с АС посредством связывающей группы (M), где M выбран из:



где:  $R^1$  представляет собой алкилен, циклоалкил или , где  $t'$  представляет собой 0-10, где каждый  $R$  независимо представляет собой алкил, алкенил, алкинил,

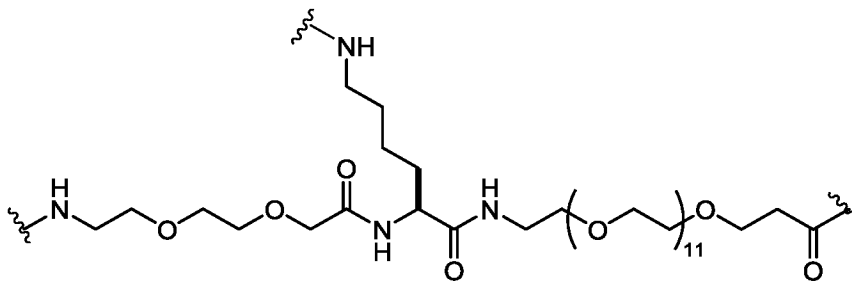
карбоциклил или гетероциклил, где  $R^1$  представляет собой , и  $t'$  представляет собой 2.

Линкер может иметь структуру:

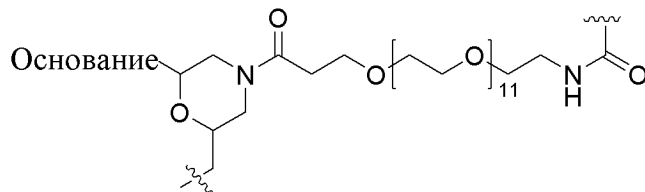


где  $AA_s$  определен в настоящем документе, и  $m'$  представляет собой 0-10.

Линкер может иметь формулу:

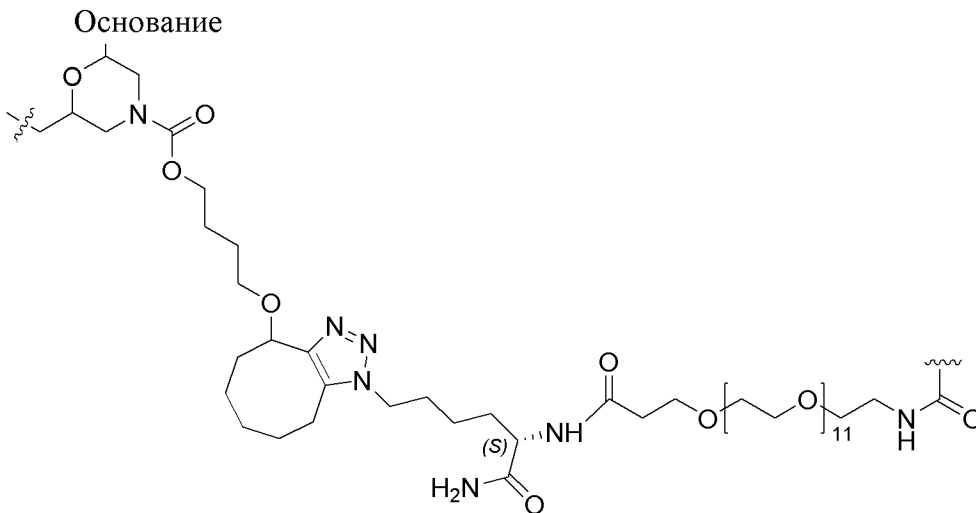


Линкер может иметь формулу:



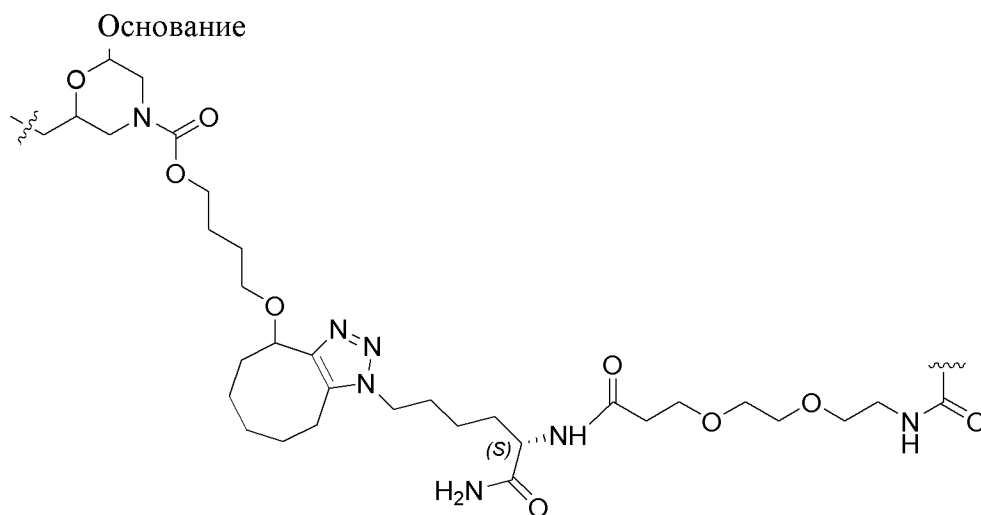
где «основание» представляет собой нуклеиновое основание на 3'-конце груза фосфордиамидат-морфолинового олигомера.

Линкер может иметь формулу:



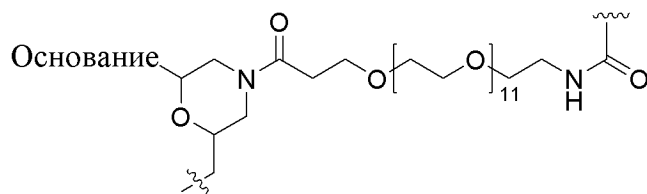
где «основание» соответствует нуклеиновому основанию на 3'-конце груза фосфордиамидат-морфолинового олигомера.

Линкер может иметь формулу:



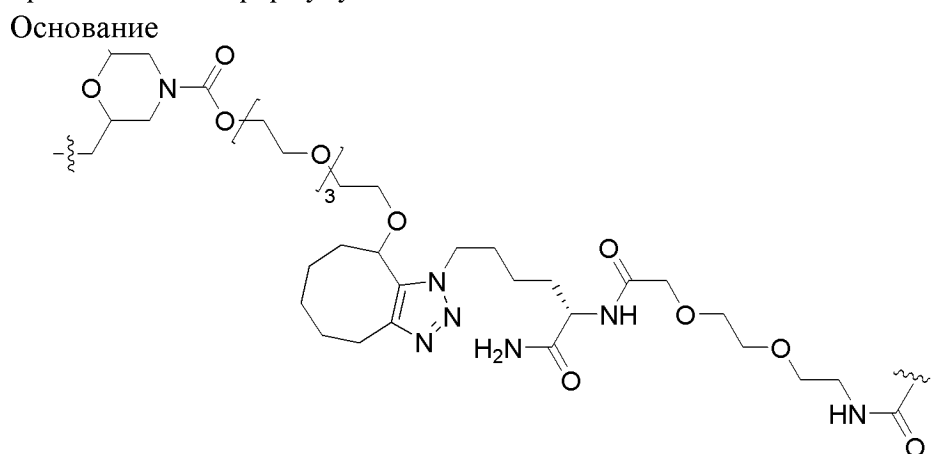
где «основание» представляет собой нуклеиновое основание на 3'-конце груза фосфордиамидат-морфолинового олигомера.

Линкер может иметь формулу:



где «основание» представляет собой нуклеиновое основание на 3'-конце груза фосфордиамидат-морфолинового олигомера.

Линкер может иметь формулу:



Линкер может быть ковалентно связан с грузом в любом подходящем месте на грузе. Линкер ковалентно связан с 3'-концом груза или 5'-концом олигонуклеотидного груза. Линкер может быть ковалентно связан с остовом груза.

Линкер может быть связан с боковой цепью аспарагиновой кислоты, глутаминовой кислоты, глутамина, аспарагина или лизина или модифицированной боковой цепью

глутамин или аспарагин (например, с восстановленной боковой цепью, имеющей аминогруппу) на сСРР. Линкер может быть связан с боковой цепью лизина на сСРР.

#### *Конъюгаты сСРР-линкер*

сСРР может быть конъюгирован с линкером, определенным в настоящем документе. Линкер может быть конъюгирован с AA<sub>5C</sub> сСРР, как определено в настоящем документе.

Линкер может содержать субъединицу  $-(OCH_2CH_2)_{z'}$  (например, в виде спейсера), где  $z'$  представляет собой целое число от 1 до 23, например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 или 23.  $-(OCH_2CH_2)_{z'}$  также называется PEG. Конъюгат сСРР-линкер может иметь структуру, выбранную из **Таблицы 7**:

Таблица 7: Конъюгаты сСРР-линкер и SEQ ID NO

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

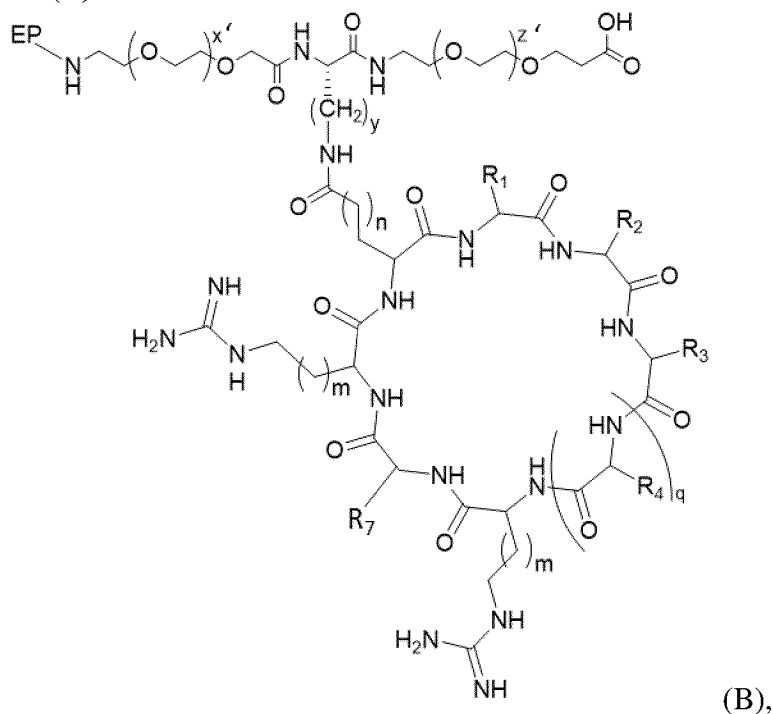
Линкер может содержать субъединицу  $-(OCH_2CH_2)_{z'}$ , где  $z'$  представляет собой целое число от 1 до 23, и пептидную субъединицу. Пептидная субъединица может содержать от 2 до 10 аминокислот. Конъюгат сСРР-линкер может иметь структуру, выбранную из **Таблицы 8**:

Таблица 8: Конъюгат сСРР-линкер и SEQ ID NO

Ac-РКККРКV-Lys(цикло[FfΦ-R-r-Cit-rQ])-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-Lys(цикло[SEQ ID NO:127])-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-Lys(цикло[FfΦ-Cit-r-R-rQ])-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-Lys(цикло[SEQ ID NO:128])-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-K(цикло(FfΦR-cit-R-cit-Q))-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-K(цикло(SEQ ID NO:129))-PEG <sub>12</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-B-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-PEG2-Lys(цикло[SEQ ID NO:130])-B-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-PEG2-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-PEG2-Lys(цикло[SEQ ID NO:130])-PEG2-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-PEG4-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-PEG2-Lys(цикло[SEQ ID NO:130])-PEG4-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-РКККРКV-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac-SEQ ID NO:42-Lys(цикло[SEQ ID NO:130])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>

Ac-pkkkrkv-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:131-PEG2-Lys(цикло[SEQ ID NO:130])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-rrv-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-rQ])-PEG12-OH	Ac-rrv-PEG2-Lys(цикло[SEQ ID NO:130])-PEG12-OH
Ac-PKKKRKV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-r-Q])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:42-PEG2-Lys(цикло[SEQ ID NO:130])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-PKKK-Cit-KV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-r-Q])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:126-PEG2-Lys(цикло[SEQ ID NO:130])-PEG12-k(N <sub>3</sub> )-NH <sub>2</sub>
Ac-PKKKRKV-PEG2-Lys(цикло[FfΦ-Cit-r-Cit-r-Q])-PEG12-K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:42-PEG2-Lys(цикло[SEQ ID NO:130])-PEG12-K(N <sub>3</sub> )-NH <sub>2</sub>

Предложены EEV, содержащие проникающий в клетку циклический пептид (сСРР), линкер и экзоциклический пептид (EP). EEV может содержать структуру Формулы (B):



или ее протонированную форму,

где:

каждый R<sub>1</sub>, R<sub>2</sub> и R<sub>3</sub> независимо представляет собой H или ароматическую или гетероароматическую боковую цепь аминокислоты;

R<sub>4</sub> и R<sub>7</sub> независимо представляют собой H или боковую цепь аминокислоты;

EP представляет собой экзоциклический пептид, определенный в настоящем документе;

каждый m независимо представляет собой целое число от 0 до 3;

n представляет собой целое число от 0 до 2;

x' представляет собой целое число от 1 до 20;



у представляет собой целое число от 1 до 5;

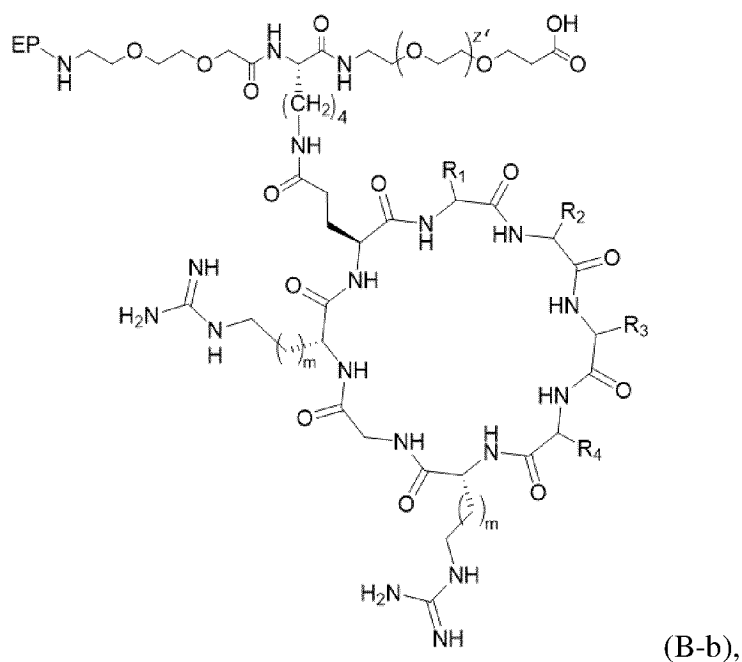
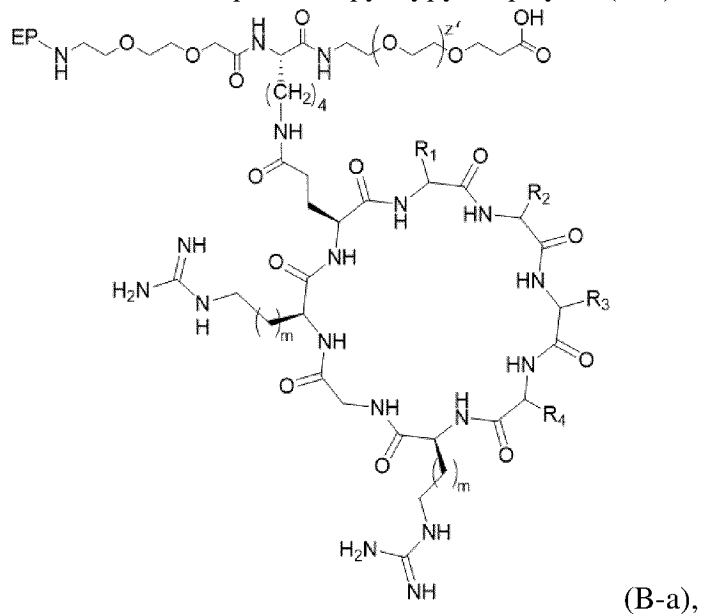
q представляет собой 1-4; и

z' представляет собой целое число от 1 до 23.

R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>7</sub>, EP, m, q, y, x', z' описаны в настоящем документе.

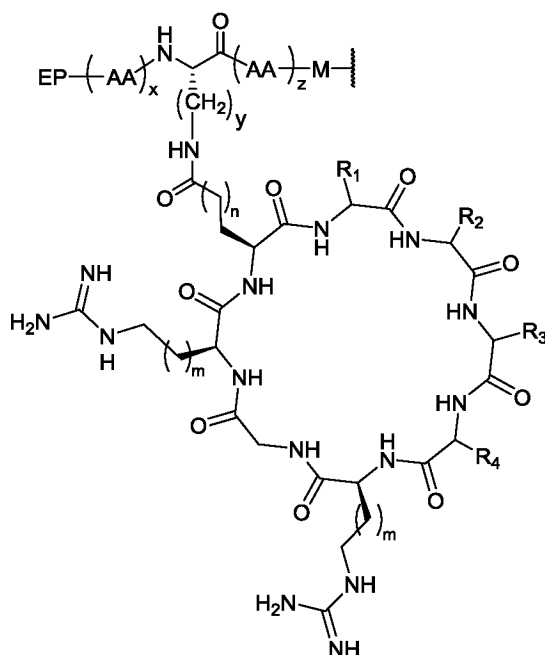
n может представлять собой 0. n может представлять собой 1. n может представлять собой 2.

EEV может содержать структуру Формулы (B-a) или (B-b):



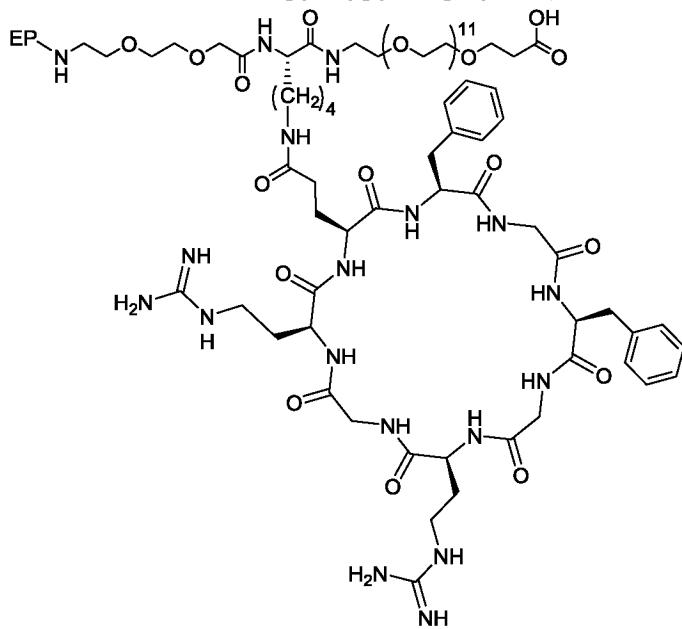
или ее протонированную форму, где EP, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup>, m и z' определены выше в Формуле (B).

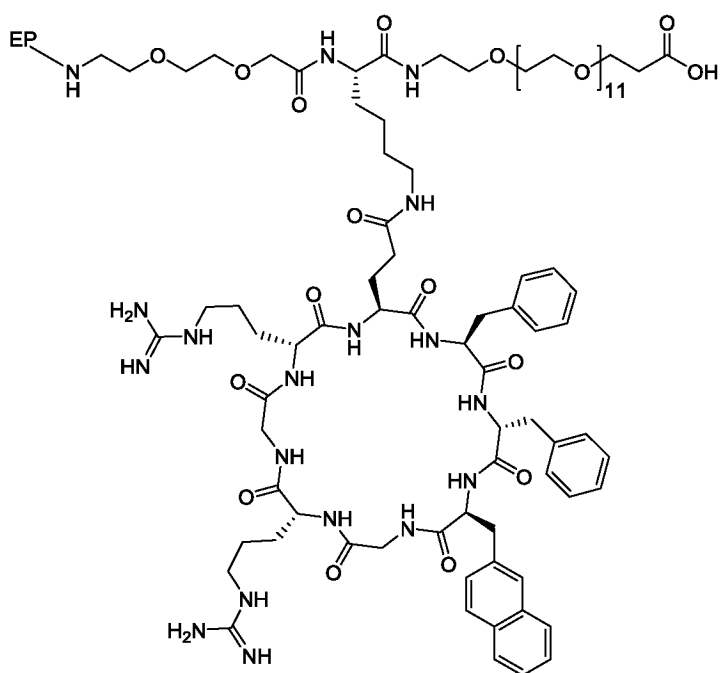
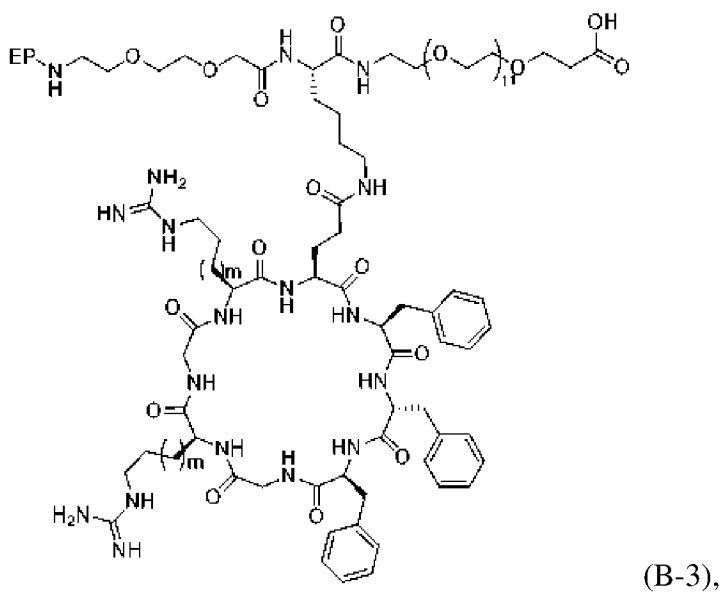
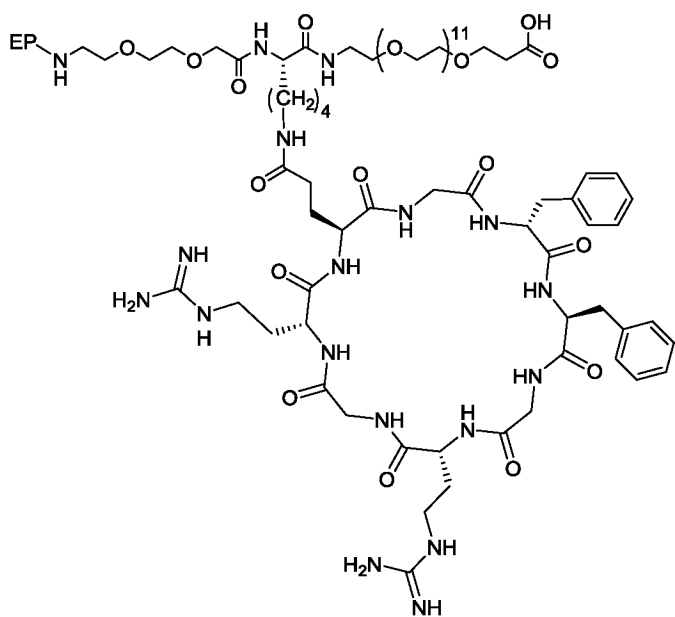
EEV может содержать структуру Формулы (B-c):



или ее протонированную форму, где EP, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> и m определены выше в Формуле (B); AA представляет собой аминокислоту, определенную в настоящем документе; M определен в настоящем документе; n представляет собой целое число от 0 до 2; x представляет собой целое число от 1 до 10; y представляет собой целое число от 1 до 5; и z представляет собой целое число от 1 до 10.

EEV может иметь структуру Формулы (B-1), (B-2), (B-3) или (B-4):



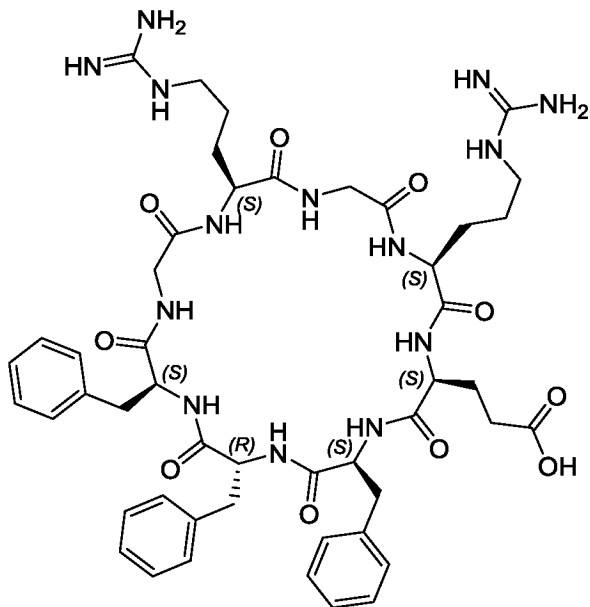


или ее

протонированную форму, где EP определен выше в Формуле (B).

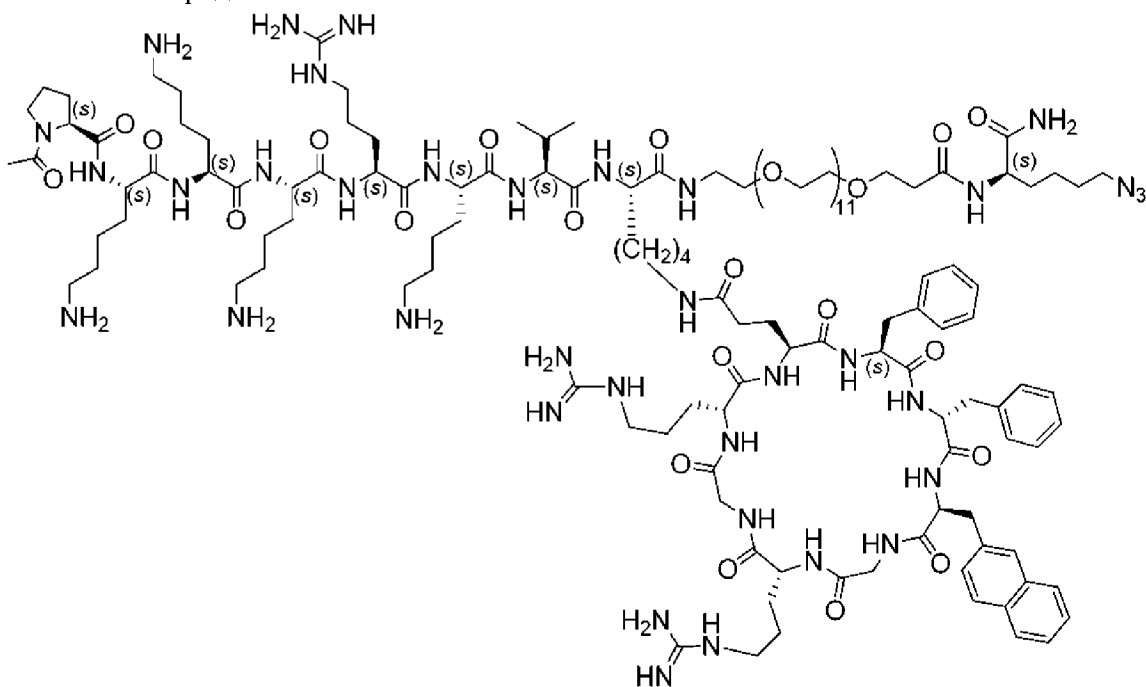
EEV может содержать Формулу (B) и может иметь структуру: Ас-РКККРКVAЕЕА-К(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH (Ас-SEQ ID NO:132- К(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH) или Ас-РК-ККР-КV-АЕЕА-К(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH (Ас- SEQ ID NO:133-К(цикло[SEQ ID NO:83])-PEG<sub>12</sub>-OH).

EEV может содержать сСРР формулы:

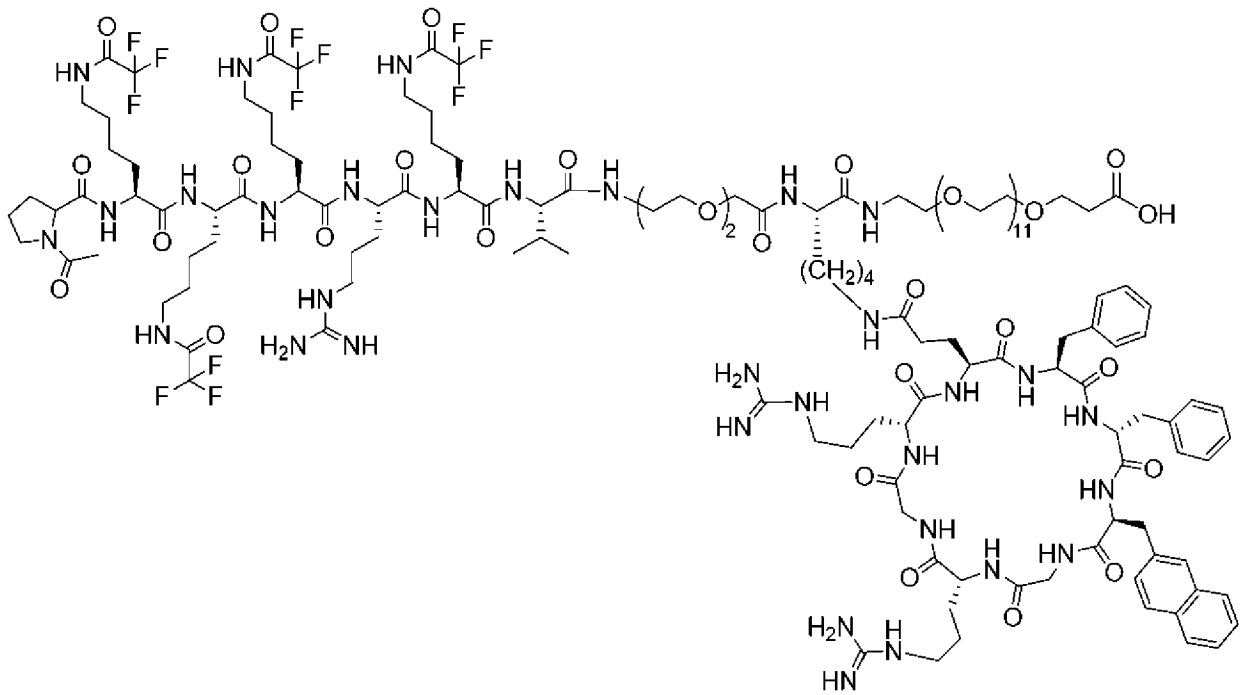


EEV может содержать формулу: Ас-РКККРКV-миниPEG<sub>2</sub>-Lys(цикло(FfFGRGRQ))-PEG<sub>2</sub>-K(N<sub>3</sub>) (Ас-SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(цикло(SEQ ID NO:81))-PEG<sub>2</sub>-K(N<sub>3</sub>)).

EEV может представлять собой:

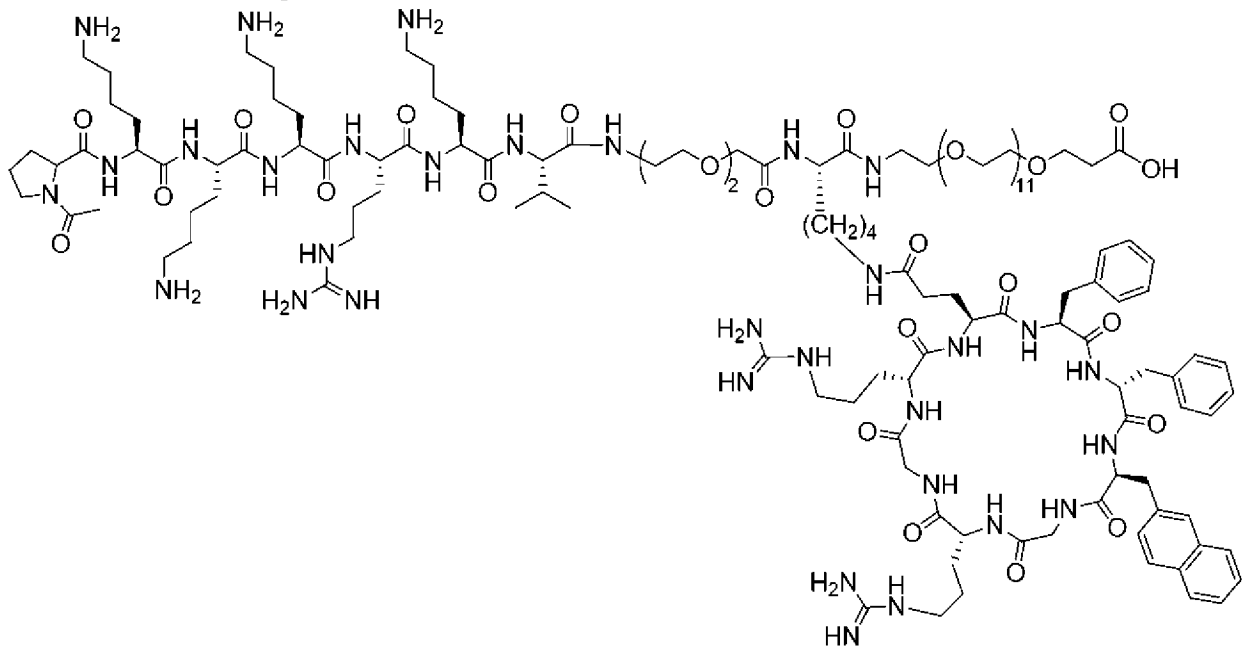


EEV может представлять собой



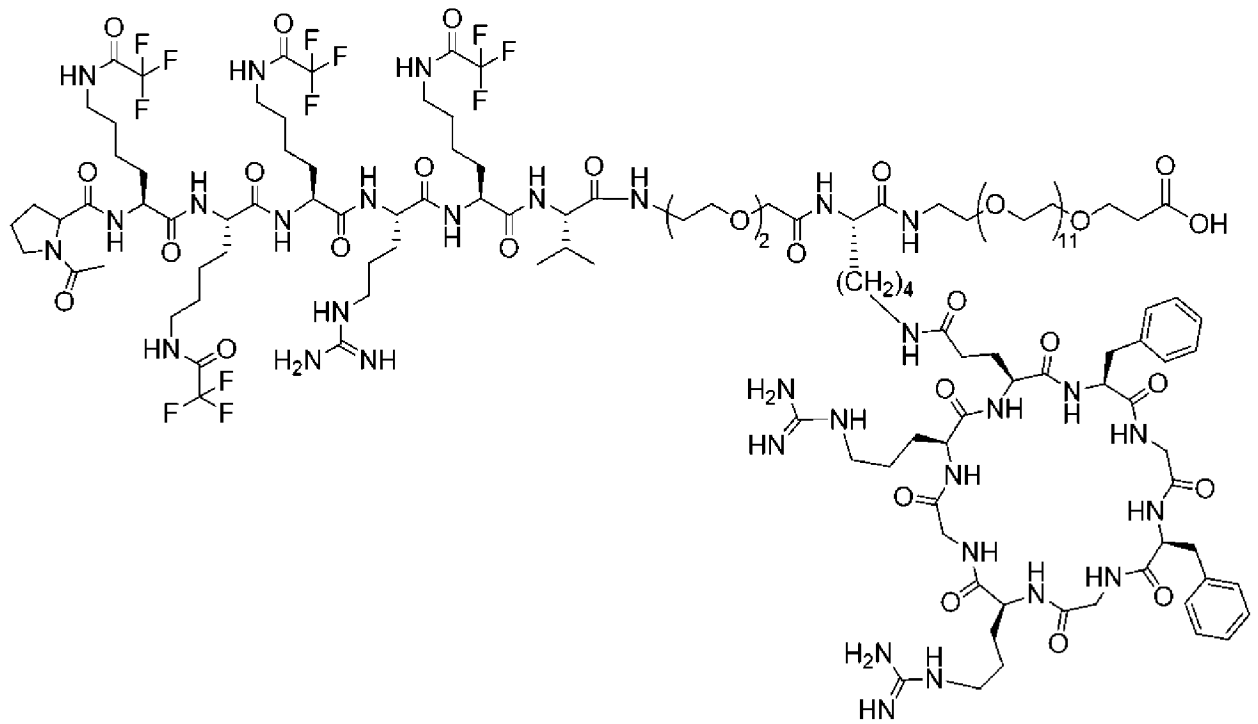
EEV может представлять собой Ac-P-K(Tfa)-K(Tfa)-K(Tfa)-R-K(Tfa)-V-миниPEG<sub>2</sub>-K(цикло(Ff-Nal-GrGrQ))-PEG<sub>12</sub>-OH (Ac-SEQ ID NO:134-миниPEG<sub>2</sub>-K(цикло(SEQ ID NO:135))-PEG<sub>12</sub>-OH).

EEV может представлять собой

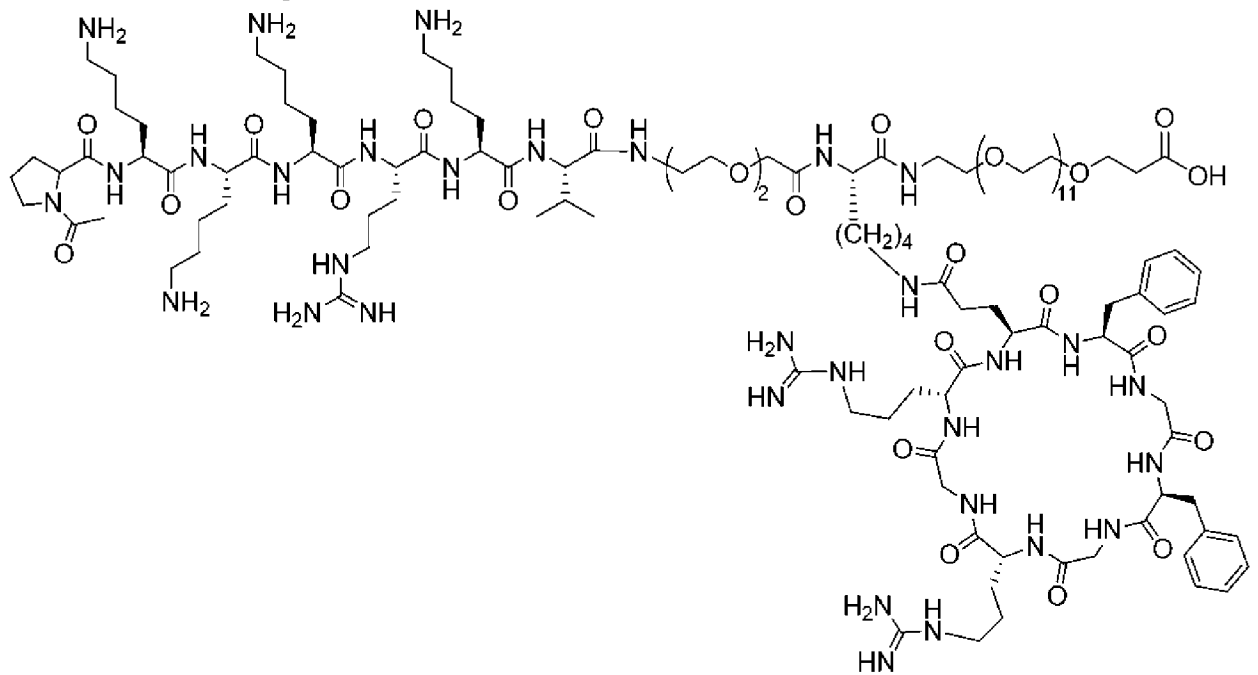


EEV может представлять собой Ac-P-K-K-K-R-K-V-миниPEG<sub>2</sub>-K(цикло(Ff-Nal-GrGrQ))-PEG<sub>12</sub>-OH (Ac-SEQ ID NO:42-миниPEG<sub>2</sub>-K(цикло(SEQ ID NO:135))-PEG<sub>12</sub>-OH).

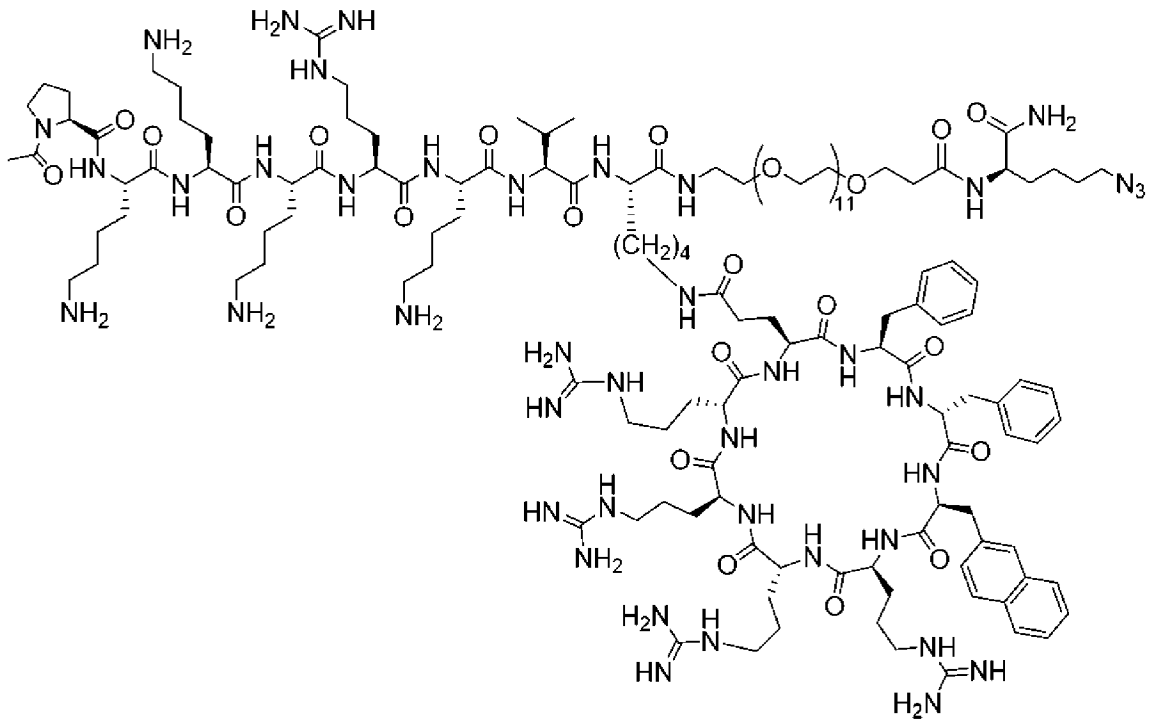
EEV может представлять собой



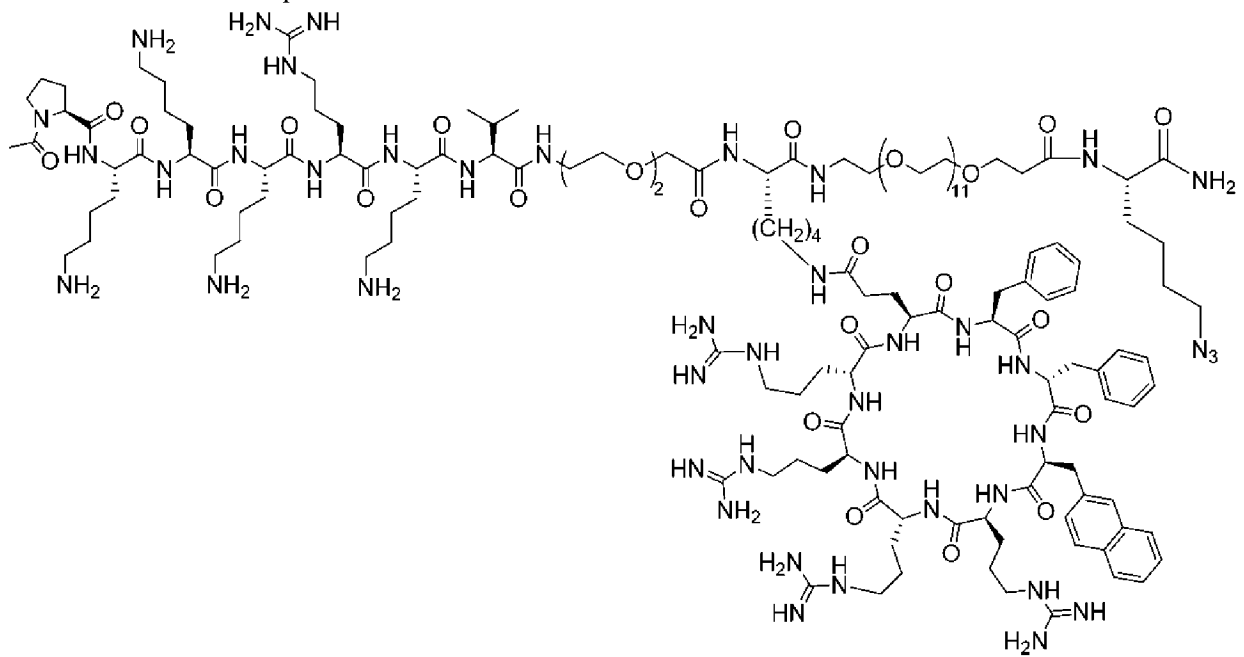
EEV может представлять собой



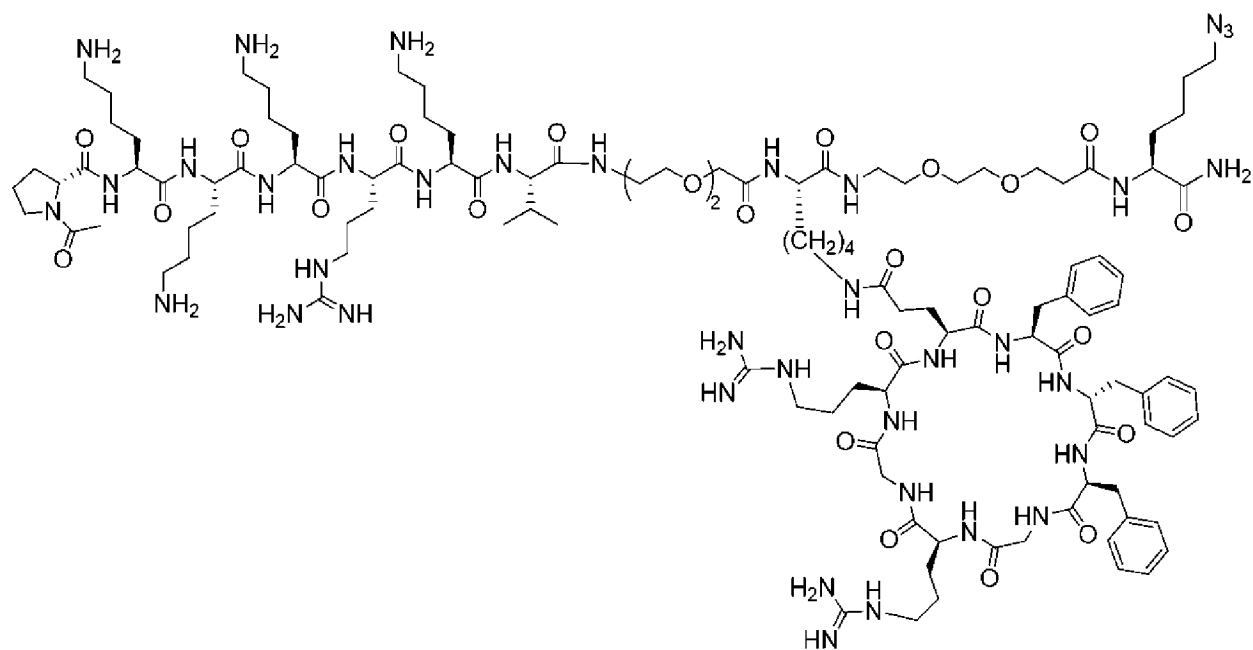
EEV может представлять собой



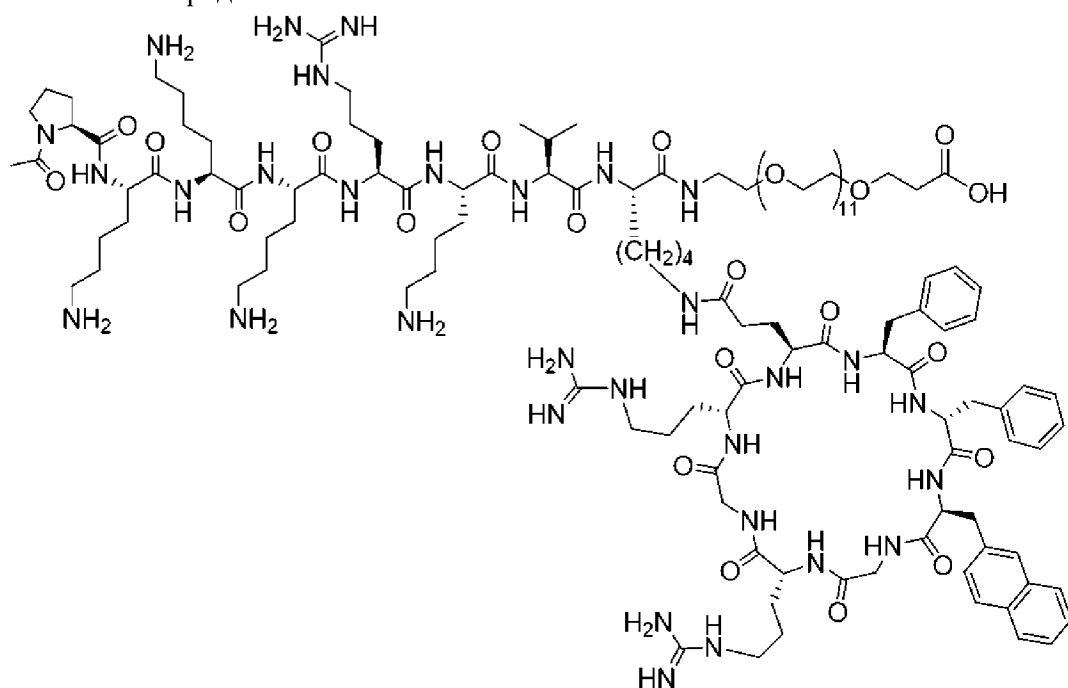
EEV может представлять собой



EEV может представлять собой

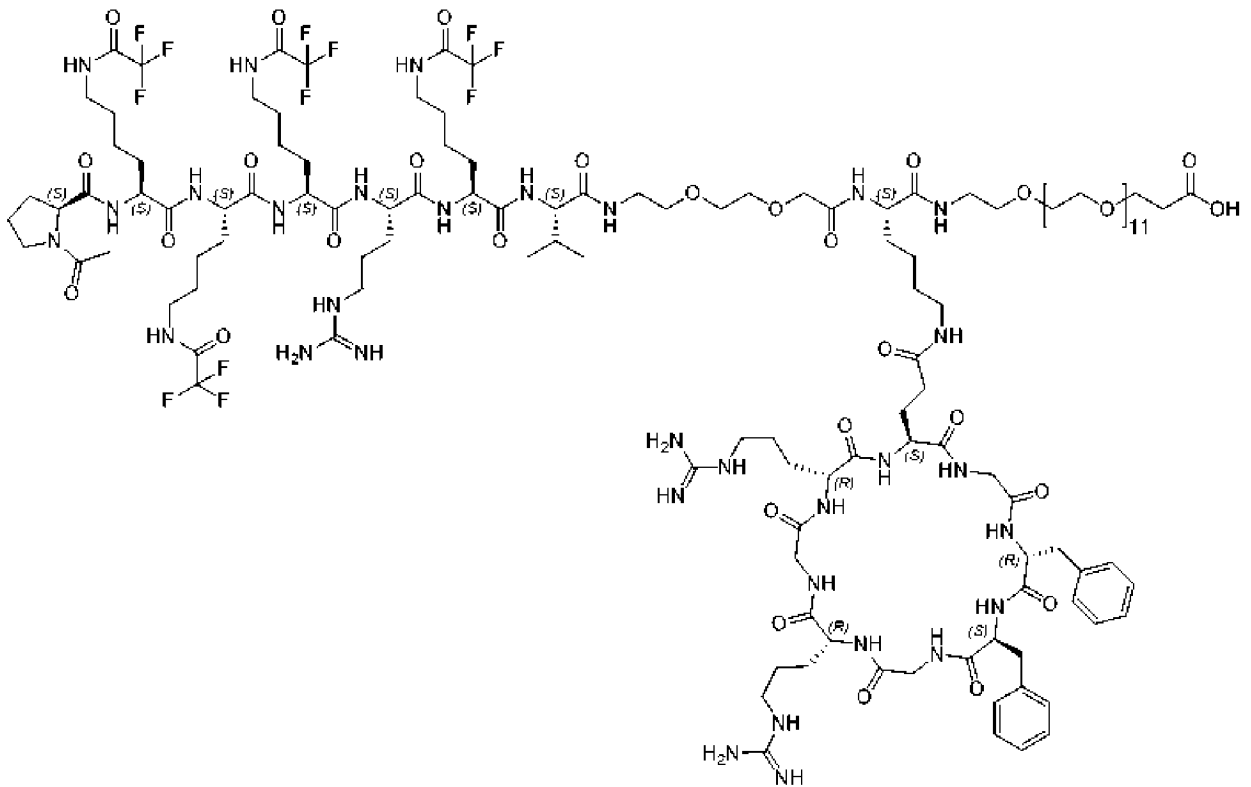


EEV может представлять собой

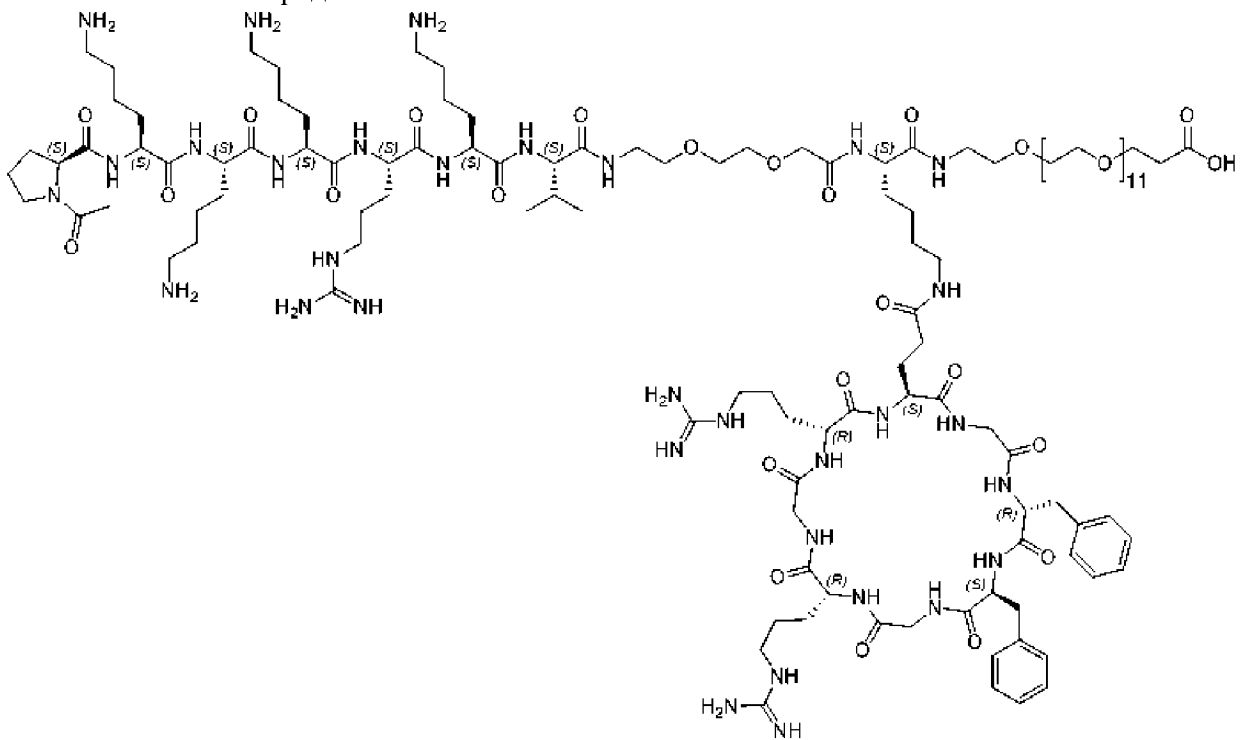


EEV может представлять собой:

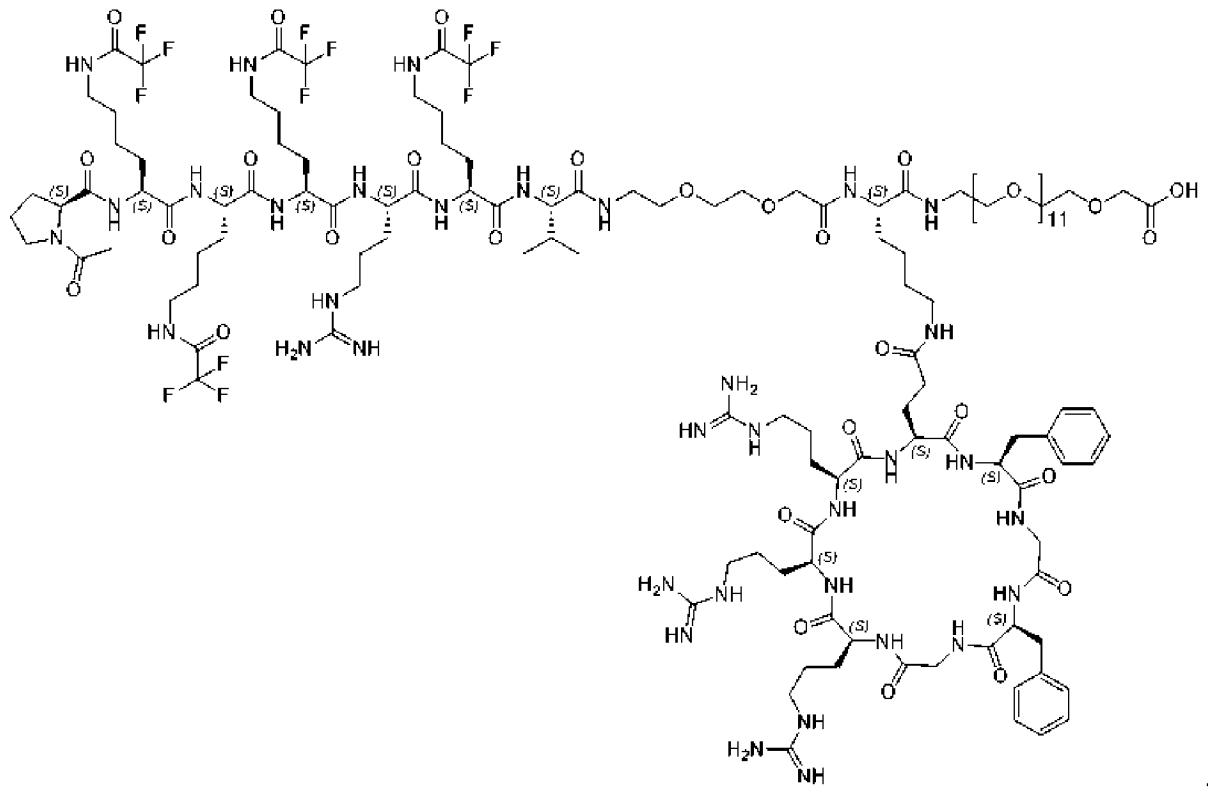




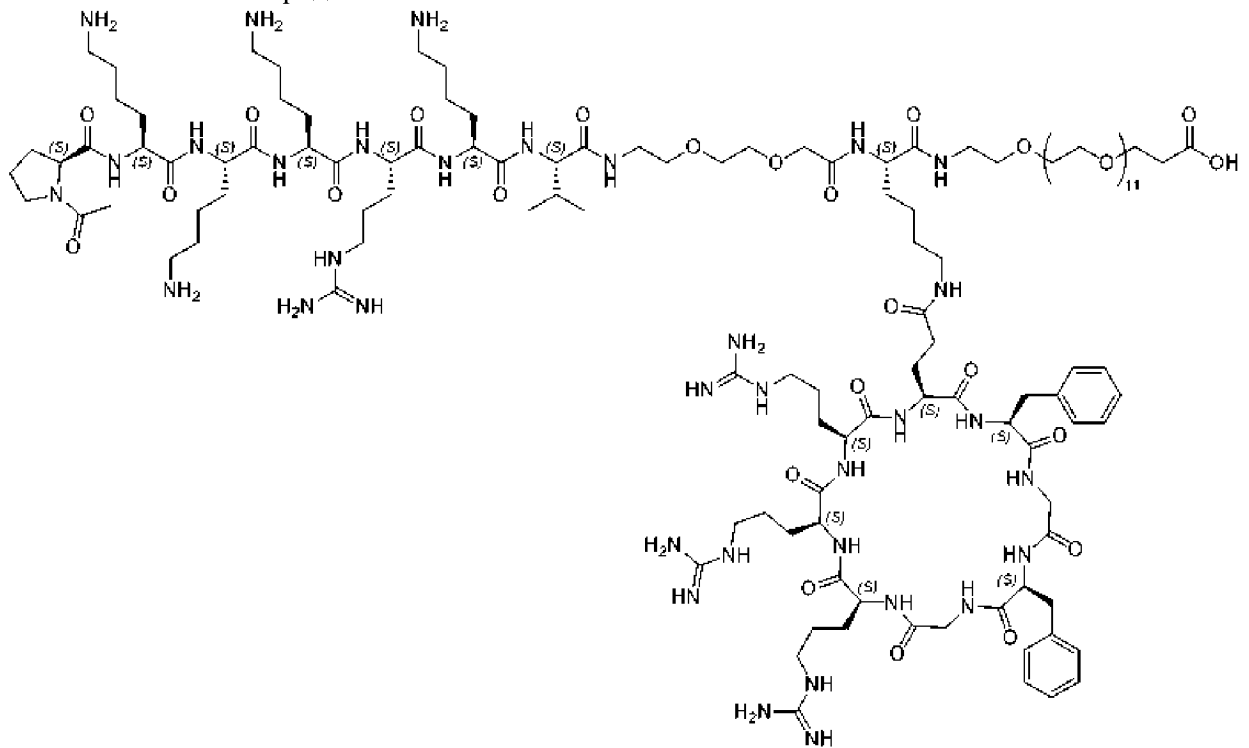
EEV может представлять собой



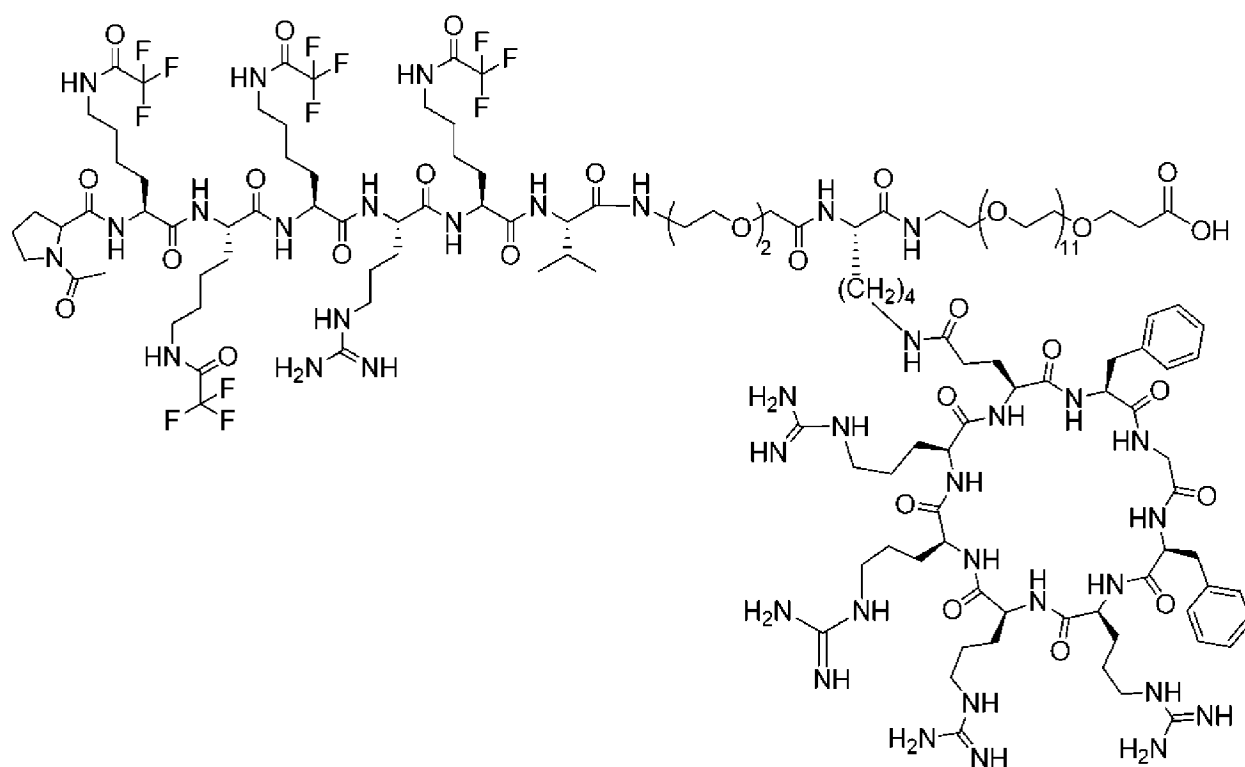
EEV может представлять собой



EEV может представлять собой



EEV может представлять собой



ЕЕV может быть выбран из

Ac-rf-миниPEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-rf-миниPEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-frr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-frr-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rfr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-rfr-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rbfbr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-SEQ ID NO:137-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rrr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-rrr-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rbr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-rbr-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rbrbr-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-SEQ ID NO:138-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-hh-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-hh-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-hbh-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-hbh-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-hbhbh-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-SEQ ID NO: 139-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH

PEG <sub>12</sub> -OH	Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rhbhb-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac- SEQ ID NO: 140-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-hbrbh-PEG <sub>2</sub> -Dap(цикло[FfΦ-Cit-r-Cit-rQ])-PEG <sub>12</sub> -OH	Ac-SEQ ID NO:141-PEG <sub>2</sub> -Dap(цикло[SEQ ID NO:136])-PEG <sub>12</sub> -OH
Ac-rr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rr-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-frr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-frr-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rfr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rfr-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rbfr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:137-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rrr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rrr-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rbr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-rbr-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rbrbr-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:138-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-hh-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-hh-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-hbh-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-hbh-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-hhbhb-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO: 139-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-rhbhb-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac-SEQ ID NO: 140-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-hbrbh-Dap(цикло[FfΦ-Cit-r-Cit-rQ])-b-OH	Ac- SEQ ID NO:141-Dap(цикло[SEQ ID NO:136])-b-OH
Ac-KKKK-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:7-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:Z80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-KGKK-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:13-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:Z80])-

	миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-ККГК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:14-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:Z80])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-ККК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-ККК-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-КК-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КГК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-КГК-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КВК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-КВК-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КВКВК-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:24-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КР-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-КР-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-КВР-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас-КВР-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-РКККРКV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:42-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-РКККРКV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:42-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-РГККРКV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:43-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-РКГКРКV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:44-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ас-РККГРКV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ас- SEQ ID NO:45-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>

Ac-PKKKGV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:46-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-PKKKRGV-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:47-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-PKKKRKG-миниPEG <sub>2</sub> - Lys(цикло[FfΦGrGrQ])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:48-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-KKKRK-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub>	Ac- SEQ ID NO:19-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub>
Ac-KRK-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub> and	Ac- SEQ ID NO:8-миниPEG <sub>2</sub> - Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> - K(N <sub>3</sub> )-NH <sub>2</sub> and
Ac-KRK-миниPEG <sub>2</sub> -Lys(цикло[FfΦGrGrQ])- миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub> .	Ac-KRK-миниPEG <sub>2</sub> -Lys(цикло[SEQ ID NO:80])-миниPEG <sub>2</sub> -K(N <sub>3</sub> )-NH <sub>2</sub> .

EEV может быть выбран из: Ac-PKKKRKV-Lys(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-Lys(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKV-миниPEG<sub>2</sub>-Lys(цикло[FfΦGrGrQ])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(цикло[SEQ ID NO:80])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKV-миниPEG<sub>2</sub>-Lys(цикло[FGFGRGRQ])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(цикло[SEQ ID NO:82])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-KR-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac-KR-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKGV-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:46-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKG-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:48-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-KKKRK-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:19-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKV-миниPEG<sub>2</sub>-Lys(цикло[FFΦGRGRQ])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(цикло[SEQ ID NO:80])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKV-миниPEG<sub>2</sub>-Lys(цикло[βhFfΦGrGrQ])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(цикло[SEQ ID NO:142])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-PKKKRKV-миниPEG<sub>2</sub>-Lys(цикло[FfΦSrSrQ])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:143])-миниPEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>).

EEV может быть выбран из: Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*(GfFGrGrQ))-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*(SEQ ID NO:133))-PEG<sub>12</sub>-ОН) Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFKRKRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:144])-PEG<sub>12</sub>-ОН) Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFRGRGQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:145])-PEG<sub>12</sub>-ОН) Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFGRGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:146])-PEG<sub>12</sub>-ОН) Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFGRrRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:147])-PEG<sub>12</sub>-ОН) Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFGRRRRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:84])-PEG<sub>12</sub>-ОН) и Ac-РКККРКV-миниPEG<sub>2</sub>-Lys(*цикло*[FGFRRRRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-миниPEG<sub>2</sub>-Lys(*цикло*[SEQ ID NO:85])-PEG<sub>12</sub>-ОН).

EEV может быть выбран из: Ac-К-К-К-R-К-G-миниPEG<sub>2</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac-SEQ ID NO:148-миниPEG<sub>2</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-К-К-К-R-К-миниPEG<sub>2</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:19-миниPEG<sub>2</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-К-К-R-К-К-PEG<sub>4</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:22-PEG<sub>4</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-К-R-К-К-К-PEG<sub>4</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:21-PEG<sub>4</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-К-К-К-К-R-PEG<sub>4</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:23-PEG<sub>4</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-R-К-К-К-К-PEG<sub>4</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:20-PEG<sub>4</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) и Ac-К-К-К-R-К-PEG<sub>4</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:19-PEG<sub>4</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН).

EEV может быть выбран из: Ac-РКККРКV-PEG<sub>2</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) Ac-РКККРКV-PEG<sub>2</sub>-K(*цикло*[FGFGRGRQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(*цикло*[SEQ ID NO:82])-PEG<sub>12</sub>-ОН) Ac-РКККРКV-PEG<sub>2</sub>-K(*цикло*[GfFGrGrQ])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>

(Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(*цикло*[SEQ ID NO:133])-PEG<sub>2</sub>-K(N<sub>3</sub>)-NH<sub>2</sub>) и Ac-РКККРКV-PEG<sub>2</sub>-K(*цикло*[GfFGrGrQ])-PEG<sub>12</sub>-ОН

(Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(*цикло*[SEQ ID NO:133])-PEG<sub>12</sub>-ОН).

Груз может представлять собой АС, и ЕЕV может быть выбран из:

Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
 Ac-PKKKRKV-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:42-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FfF-GRGRQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
 Ac-rr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
 (Ac-rr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
 (Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-rrr-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH



(Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
Ac-rrr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
(Ac-rrr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
Ac-rhr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
(Ac-rhr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
Ac-rbr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
(Ac-rbr-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
(Ac-SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
(Ac-SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
(Ac-SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
Ac-rbrbr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH

(Ac- SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
 Ac-rbrbr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:138-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
 Ac-rbhbr-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:149-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FfΦGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:80])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FfΦCit-r-Cit-rQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:79])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FfFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:81])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FGFGRGRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:82])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[GfFGrGrQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:133])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FGFGRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:84])-PEG<sub>12</sub>-OH)  
 Ac-hbrbh-PEG<sub>2</sub>-K(цикло[FGFRRRRQ])-PEG<sub>12</sub>-OH  
 (Ac- SEQ ID NO:141-PEG<sub>2</sub>-K(цикло[SEQ ID NO:85])-PEG<sub>12</sub>-OH),

где b представляет собой бета-аланин, и экзоциклическая последовательность может иметь D- или L-стереохимию.

### **Груз**

[001] Проникающий в клетку пептид (CPP), такой как проникающий в клетку циклический пептид (например, cCPP), может быть конъюгирован с грузом. В настоящем документе «груз» представляет собой соединение или фрагмент, для которого желательна доставка в клетку. Груз может быть конъюгирован с концевой карбонильной группой

линкера. По меньшей мере один атом циклического пептида может быть заменен грузом, или по меньшей мере одна одиночная пара может образовывать связь с грузом. Груз может быть конъюгирован с сСРР с помощью линкера. Груз может быть конъюгирован с AA<sub>SC</sub> с помощью линкера. По меньшей мере один атом сСРР может быть заменен грузом, или по меньшей мере одна одиночная пара сСРР образует связь с грузом. Гидроксильная группа на боковой цепи аминокислоты сСРР может быть заменена связью с грузом. Гидроксильная группа на боковой цепи глутамина сСРР может быть заменена связью с грузом. Груз может быть конъюгирован с сСРР с помощью линкера. Груз может быть конъюгирован с AA<sub>SC</sub> с помощью линкера.

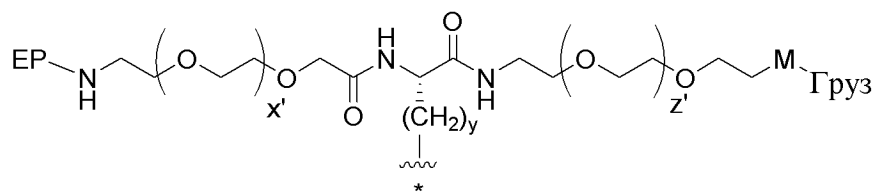
[002] Согласно вариантам реализации боковая цепь аминокислоты содержит химически реакционноспособную группу, с которой конъюгируют линкер или груз. Химически реакционноспособная группа может содержать аминогруппу, карбоновую кислоту, амид, гидроксильную группу, сульфгидрильную группу, гуанидинильную группу, фенольную группу, тиоэфирную группу, имидазольную группу или индолильную группу. Согласно вариантам реализации аминокислота сСРР, с которой конъюгирован груз, содержит лизин, аргинин, аспарагиновую кислоту, глутаминовую кислоту, аспарагин, глутамин, гомоглутамин, серин, треонин, тирозин, цистеин, аргинин, тирозин, метионин, гистидин или триптофан.

Груз может содержать один или более детектируемых фрагментов, один или более терапевтических фрагментов (ТМ), один или более нацеливающих фрагментов или любую их комбинацию. Согласно вариантам реализации груз содержит ТМ. Согласно вариантам реализации груз содержит АС.

### **Проникающие в клетку циклические пептиды (сСРР), конъюгированные с фрагментом груза**

Проникающий в клетку циклический пептид (сСРР) может быть конъюгирован с фрагментом груза

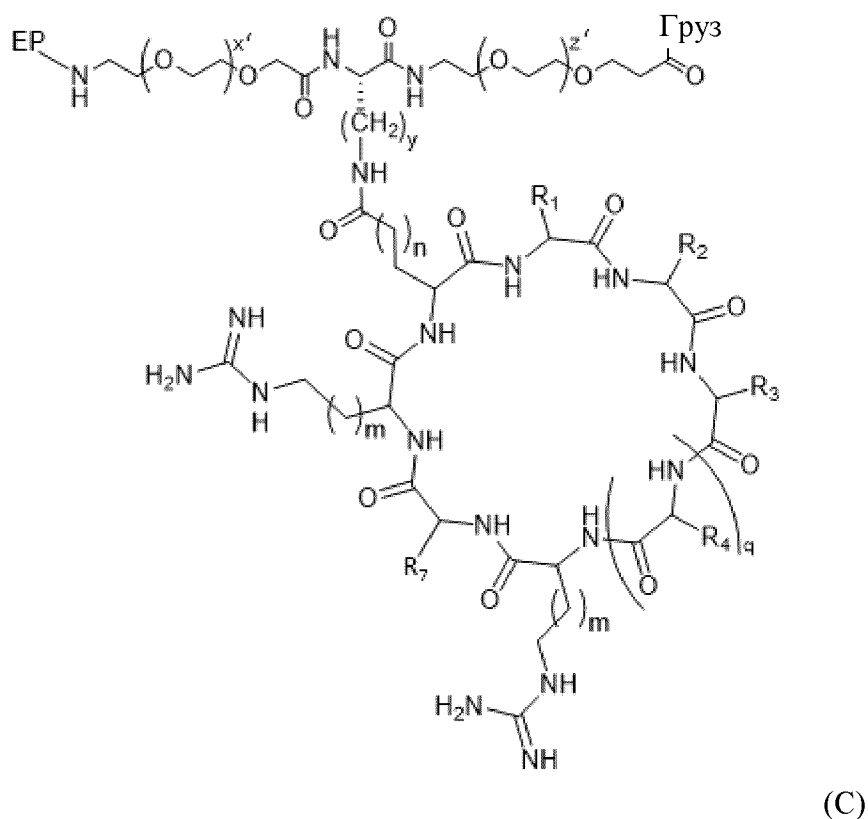
Фрагмент груза может быть конъюгирован с линкером в концевой карбонильной группе с обеспечением следующей структуры:



-, где:

EP представляет собой экзоциклический пептид, и M, AA<sub>SC</sub>, груз, x', y и z' определены выше, \* представляет собой положение присоединения к AA<sub>SC</sub>. x' может представлять собой 1. y может представлять собой 4. z' может представлять собой 11. -(OCH<sub>2</sub>CH<sub>2</sub>)<sub>x</sub>- и/или -(OCH<sub>2</sub>CH<sub>2</sub>)<sub>z</sub>- может быть независимо заменен одной или более аминокислотами, включая, например, глицин, бета-аланин, 4-аминомасляную кислоту, 5-аминопентановую кислоту, 6-аминогексановую кислоту или их комбинации.

Носитель для ускользания из эндосом (EEV) может содержать проникающий в клетку циклический пептид (сCPP), экзоциклический пептид (EP) и линкер и может быть конъюгирован с грузом с образованием EEV-конъюгата, содержащего структуру Формулы (С):



или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  может независимо представлять собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

$R_4$  представляет собой H или боковую цепь аминокислоты;

EP представляет собой экзоциклический пептид, определенный в настоящем документе;

Груз представляет собой фрагмент, определенный в настоящем документе;

каждый  $m$  независимо представляет собой целое число от 0 до 3;

$n$  представляет собой целое число от 0 до 2;

$x'$  представляет собой целое число от 2 до 20;

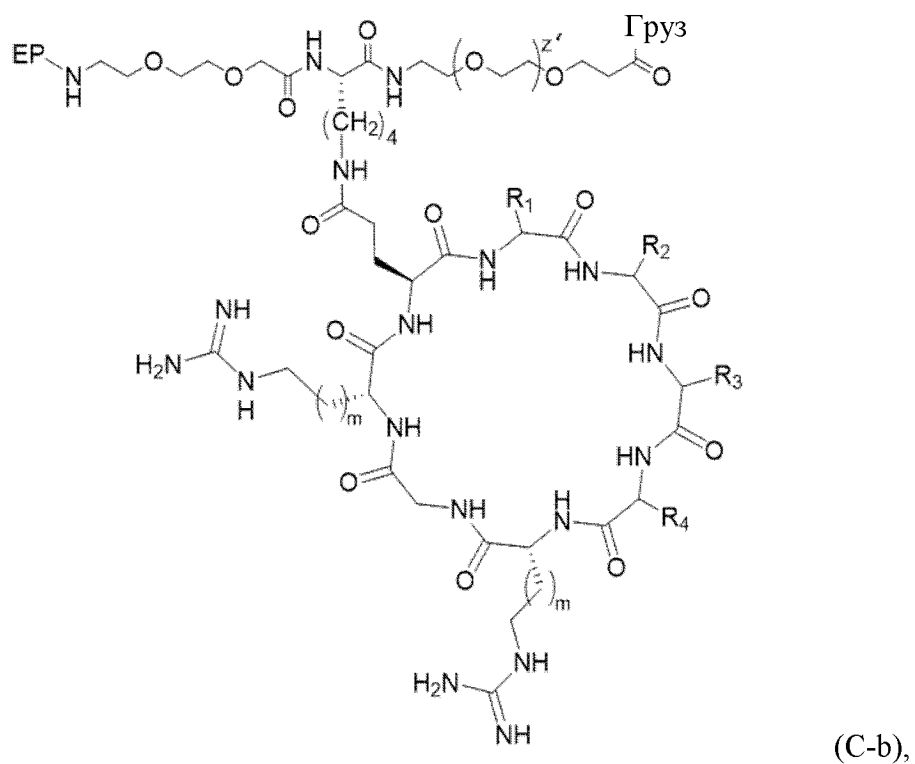
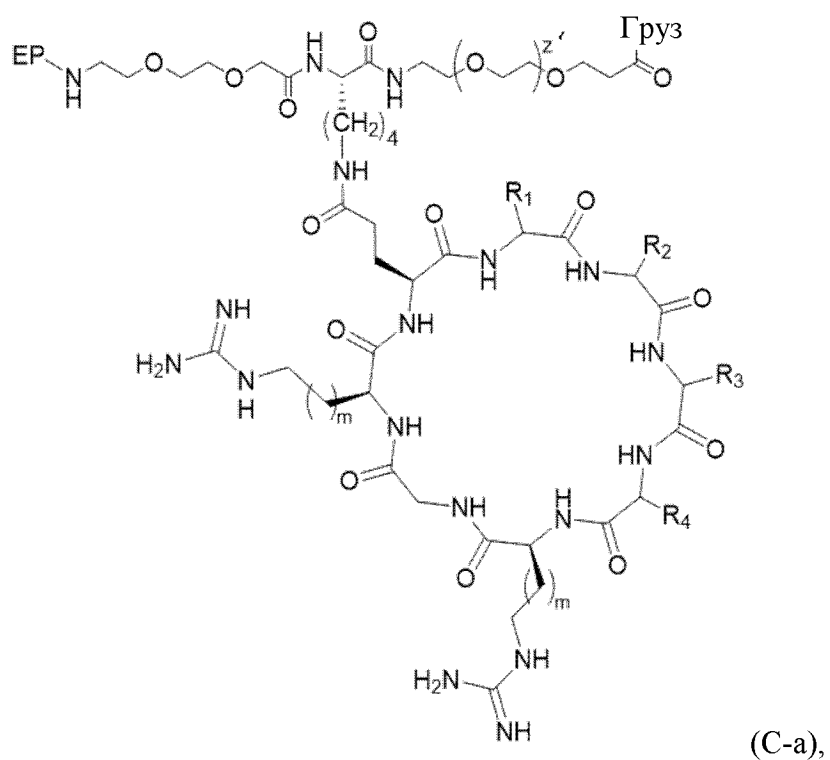
$y$  представляет собой целое число от 1 до 5;

$q$  представляет собой целое число от 1 до 4; и

$z'$  представляет собой целое число от 2 до 20.

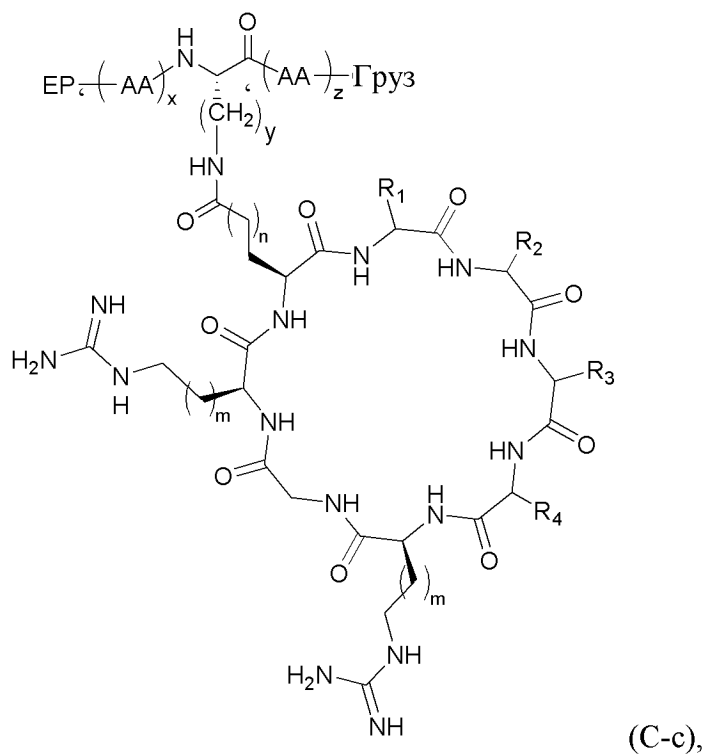
$R_1$ ,  $R_2$ ,  $R_3$ ,  $R_4$ , EP, груз,  $m$ ,  $n$ ,  $x'$ ,  $y$ ,  $q$  и  $z'$  определены в настоящем документе.

EEV может быть конъюгирован с грузом, и EEV-конъюгат может содержать структуру Формулы (С-а) или (С-б):



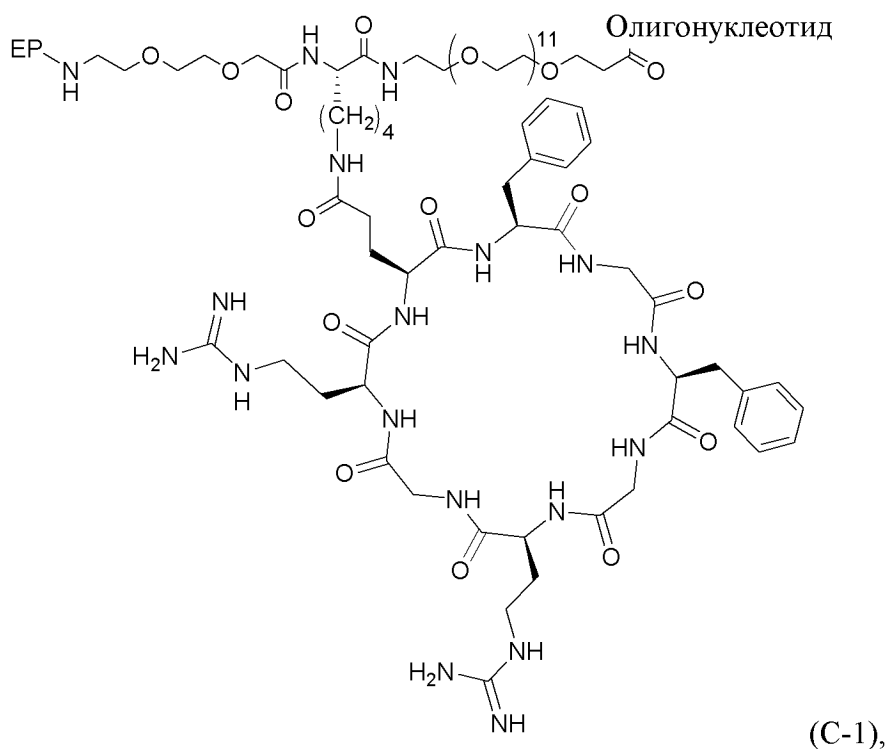
или ее протонированную форму, где EP m и z определены выше в Формуле (C).

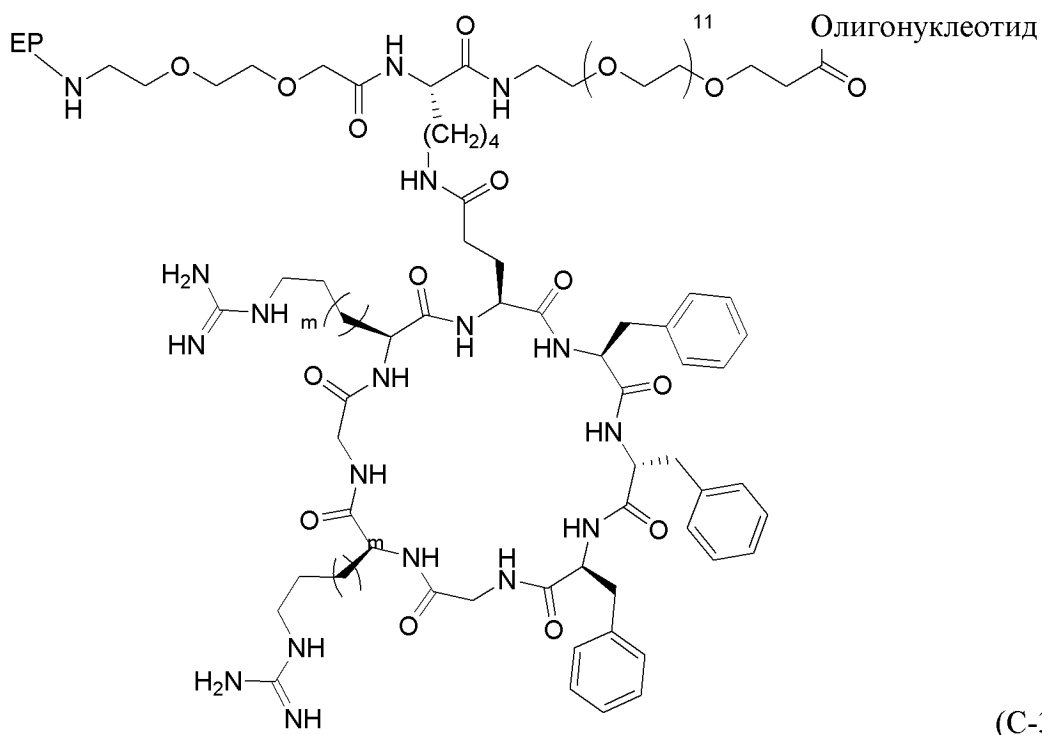
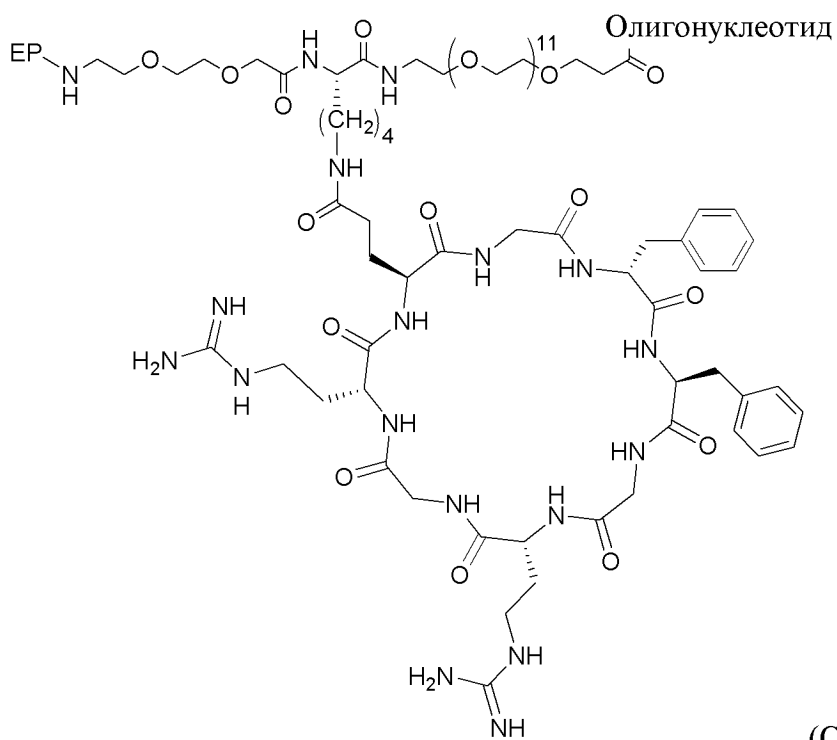
EEV может быть конъюгирован с грузом, и EEV-конъюгат может содержать структуру Формулы (C-c):

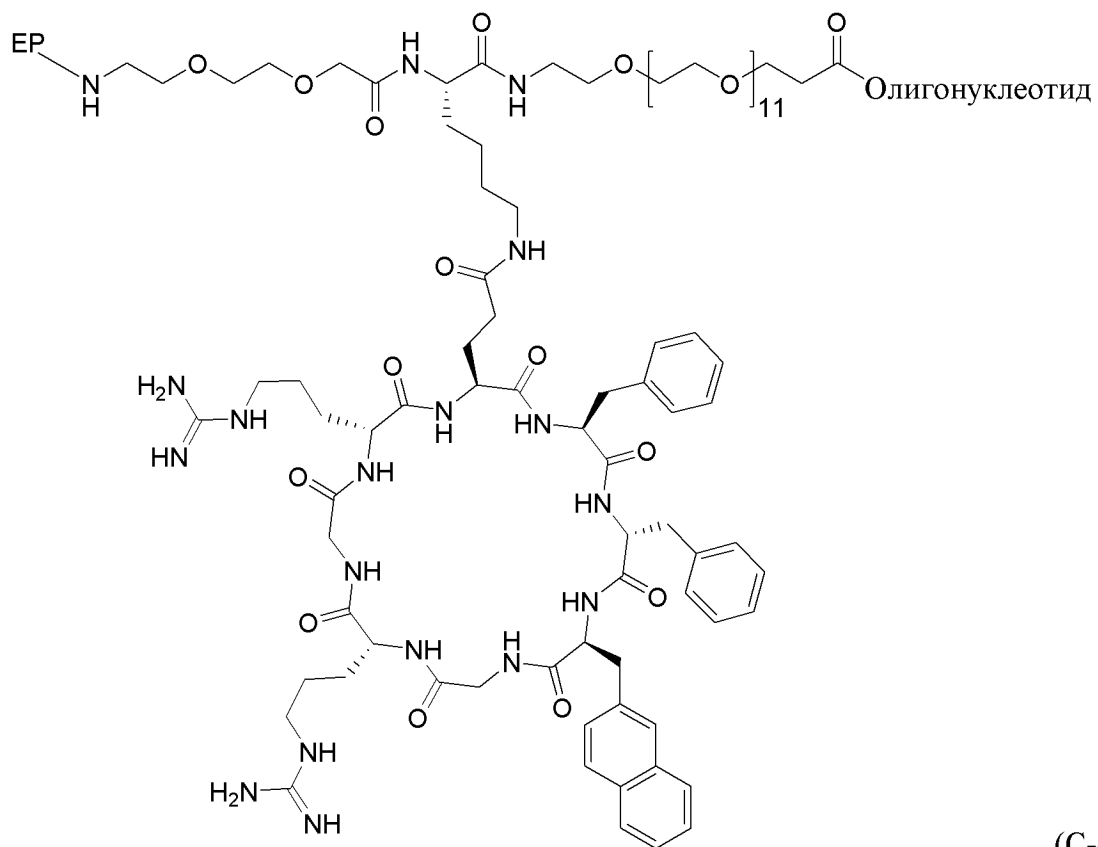


или ее протонированную форму, где EP, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> и m определены выше в Формуле (III); AA может представлять собой аминокислоту, определенную в настоящем документе; n может представлять собой целое число от 0 до 2; x может представлять собой целое число от 1 до 10; y может представлять собой целое число от 1 до 5; и z может представлять собой целое число от 1 до 10.

EEV может быть конъюгирован с олигонуклеотидным грузом, и конъюгат EEV-олигонуклеотид может содержать структуру Формулы (C-1), (C-2), (C-3) или (C-4):



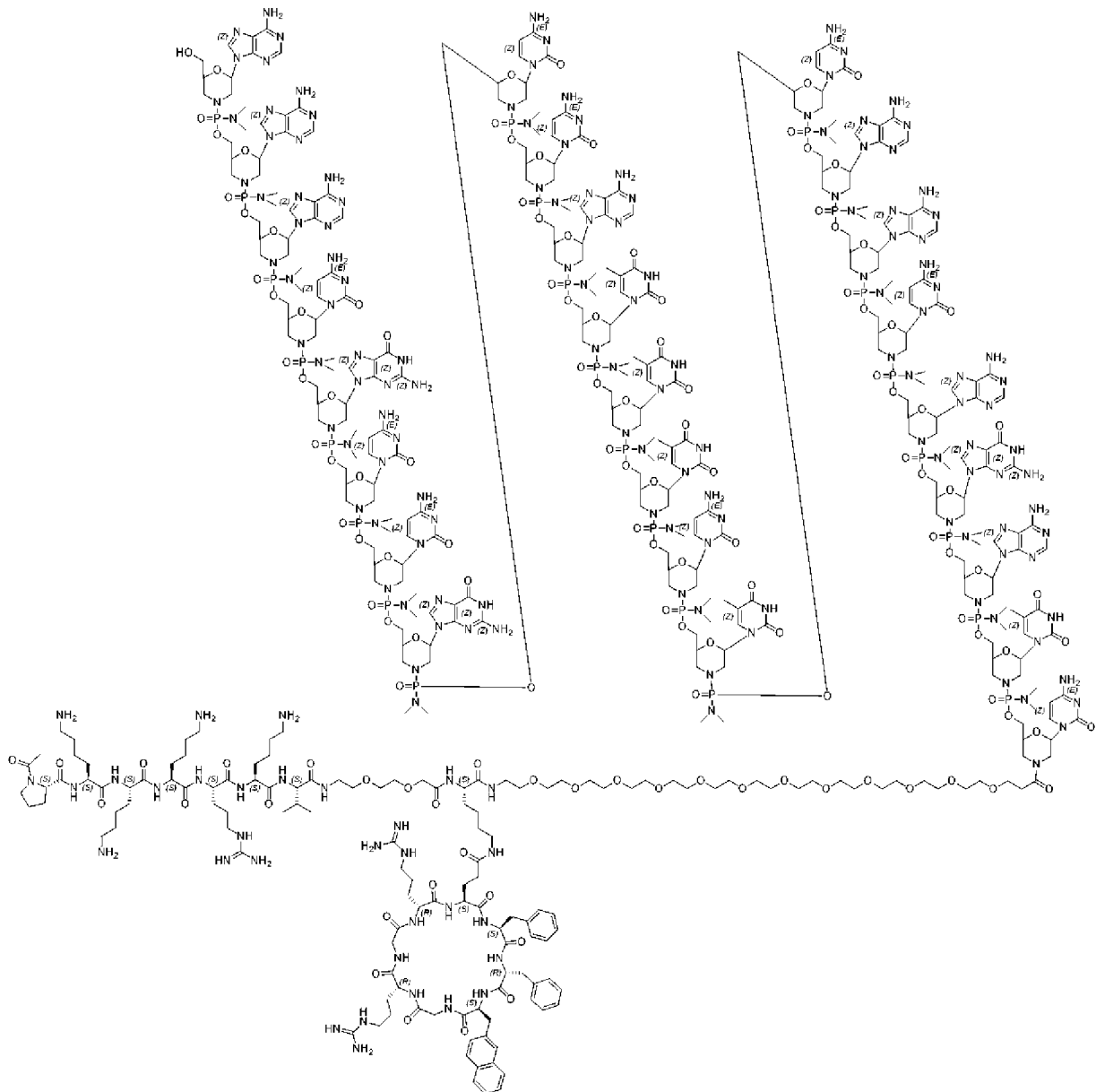




(C-4)

EEV может быть конъюгирован с олигонуклеотидным грузом, и EEV-конъюгат может содержать структуру:





### Эффективность цитозольной доставки

Модификации проникающего в клетку циклического пептида (сСРР) могут улучшить эффективность цитозольной доставки. Улучшенная эффективность цитозольного поглощения может быть измерена путем сравнения эффективности цитозольной доставки сСРР, имеющего модифицированную последовательность, с контрольной последовательностью. Контрольная последовательность не включает конкретный заменяющий аминокислотный остаток в модифицированной последовательности (включая, но не ограничиваясь перечисленными, аргинин, фенилаланин и/или глицин), но в остальном она идентична.

В настоящем документе эффективность цитозольной доставки относится к способности сСРР пересекать клеточную мембрану и поступать в цитозоль клетки. Эффективность цитозольной доставки сСРР не обязательно зависит от рецептора или типа клетки. Эффективность цитозольной доставки может относиться к абсолютной

эффективности цитозольной доставки или относительной эффективности цитозольной доставки.

Абсолютная эффективность цитозольной доставки представляет собой соотношение цитозольной концентрации сСРР (или конъюгата сСРР-груз) и концентрации сСРР (или конъюгата сСРР-груз) в ростовой среде. Относительная эффективность цитозольной доставки относится к концентрации сСРР в цитозоле по сравнению с концентрацией контрольного сСРР в цитозоле. Количественное определение может быть достигнуто путем флуоресцентного мечения сСРР (например, красителем FITC) и измерения интенсивности флуоресценции с использованием методик, хорошо известных в данной области техники.

Относительную эффективность цитозольной доставки определяют путем сравнения (i) количества сСРР согласно настоящему изобретению, интернализированного некоторым типом клеток (например, клетками HeLa), с (ii) количеством контрольного сСРР, интернализированного этим же типом клеток. Для измерения относительной эффективности цитозольной доставки тип клеток можно инкубировать в присутствии сСРР в течение определенного периода времени (например, 30 минут, 1 час, 2 часа и т. д.), после чего количество сСРР, интернализированного клеткой, количественно определяют с использованием способов, известных в данной области техники, например, флуоресцентной микроскопии. Отдельно такую же концентрацию контрольного сСРР инкубируют в присутствии типа клеток в течение аналогичного периода времени, и количество контрольного сСРР, интернализированного клеткой, количественно определяют.

Относительную эффективность цитозольной доставки можно определить путем измерения  $IC_{50}$  сСРР, имеющего модифицированную последовательность для внутриклеточной мишени, и сравнения  $IC_{50}$  сСРР, имеющего модифицированную последовательность, с контрольной последовательностью (как описано в настоящем документе).

Относительная эффективность цитозольной доставки сСРР может находиться в диапазоне от примерно 50% до примерно 450% по сравнению с цикло(FfΦRrRrQ, SEQ ID NO:150), например, примерно 60%, примерно 70%, примерно 80%, примерно 90%, примерно 100%, примерно 110%, примерно 120%, примерно 130%, примерно 140%, примерно 150%, примерно 160%, примерно 170%, примерно 180%, примерно 190%, примерно 200%, примерно 210%, примерно 220%, примерно 230%, примерно 240%, примерно 250%, примерно 260%, примерно 270%, примерно 280%, примерно 290%, примерно 300%, примерно 310%, примерно 320%, примерно 330%, примерно 340%, примерно 350%, примерно 360%, примерно 370%, примерно 380%, примерно 390%, примерно 400%, примерно 410%, примерно 420%, примерно 430%, примерно 440%, примерно 450%, примерно 460%, примерно 470%, примерно 480%, примерно 490%, примерно 500%, примерно 510%, примерно 520%, примерно 530%, примерно 540%, примерно 550%, примерно 560%, примерно 570%, примерно 580% или примерно 590%, включая все значения и поддиапазоны между ними. Относительная эффективность

цитозольной доставки сСРР может быть улучшена более чем на примерно 600% по сравнению с циклическим пептидом, содержащим цикло(FfFRrRrQ, SEQ ID NO:150).

Абсолютная эффективность цитозольной доставки составляет от примерно 40% до примерно 100%, например, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 91%, примерно 92%, примерно 93%, примерно 94%, примерно 95%, примерно 96%, примерно 97%, примерно 98%, примерно 99%, включая все значения и поддиапазоны между ними.

сСРР согласно настоящему изобретению могут улучшить эффективность цитозольной доставки в примерно 1,1-30 раз по сравнению с идентичной в остальном последовательностью, например, примерно 1,2, примерно 1,3, примерно 1,4, примерно 1,5, примерно 1,6, примерно 1,7, примерно 1,8, примерно 1,9, примерно 2,0, примерно 2,5, примерно 3,0, примерно 3,5, примерно 4,0, примерно 4,5, примерно 5,0, примерно 5,5, примерно 6,0, примерно 6,5, примерно 7,0, примерно 7,5, примерно 8,0, примерно 8,5, примерно 9,0, примерно 10, примерно 10,5, примерно 11,0, примерно 11,5, примерно 12,0, примерно 12,5, примерно 13,0, примерно 13,5, примерно 14,0, примерно 14,5, примерно 15,0, примерно 15,5, примерно 16,0, примерно 16,5, примерно 17,0, примерно 17,5, примерно 18,0, примерно 18,5, примерно 19,0, примерно 19,5, примерно 20, примерно 20,5, примерно 21,0, примерно 21,5, примерно 22,0, примерно 22,5, примерно 23,0, примерно 23,5, примерно 24,0, примерно 24,5, примерно 25,0, примерно 25,5, примерно 26,0, примерно 26,5, примерно 27,0, примерно 27,5, примерно 28,0, примерно 28,5, примерно 29,0 или примерно 29,5 раза, включая все значения и поддиапазоны между ними.

### **Детектируемый фрагмент**

Согласно вариантам реализации соединение, раскрытое в настоящем документе, включает детектируемый фрагмент. Согласно вариантам реализации детектируемый фрагмент присоединен к проникающему в клетку пептиду в аминогруппе, карбоксилатной группе или боковой цепи любой из аминокислот проникающего в клетку пептидного фрагмента (например, в аминогруппе, карбоксилатной группе или боковой цепи любой аминокислоты в СРР). Согласно вариантам реализации терапевтический фрагмент включает детектируемый фрагмент. Детектируемый фрагмент может включать любую детектируемую метку. Примеры подходящих детектируемых меток включают, но не ограничиваются перечисленными, метку УФ-видимого диапазона, метку ближнего инфракрасного диапазона, люминесцентную группу, фосфоресцентную группу, метку магнитного спинового резонанса, фотосенсибилизатор, фоторасщепляемый фрагмент, хелатирующий центр, тяжелый атом, радиоактивный изотоп, изотопную детектируемую метку спинового резонанса, парамагнитный фрагмент, хромофор или любую их комбинацию. Согласно вариантам реализации метка является детектируемой без добавления дополнительных реагентов.

Согласно вариантам реализации детектируемый фрагмент представляет собой биологически совместимый детектируемый фрагмент так, что соединения могут подходить для применения в различных биологических приложениях. «Биосовместимый» и «биологически совместимый» в контексте настоящего документа обычно относятся к соединениям, которые, вместе с любыми метаболитами или продуктами их разложения, обычно являются нетоксичными для клеток и тканей и которые не вызывают каких-либо значительных нежелательных эффектов в отношении клеток и тканей при инкубации (например, культивировании) клеток и тканей в их присутствии.

Детектируемый фрагмент может содержать люминофор, такой как флуоресцентная метка или метка ближнего инфракрасного диапазона. Примеры подходящих люминофоров включают, но не ограничиваются перечисленными, металлосодержащие порфирины; бензопорфирины; азабензопорфирин; нафтопорфирин; фталоцианин; полициклические ароматические углеводороды, такие как перилендиимин, пирены; азокрасители; ксантоновые красители; дипирометен бора, дипирометен азобора, цианиновые красители, комплекс металл-лиганд, такой как бипиридин, бипиридилы, фенантролин, кумарин и ацетилацетонаты рутения и иридия; акридин, производные оксазина, такие как бензофеноксазин; аза-аннулен, скварен; 8-гидроксихинолин, полиметины, люминесцентные наночастицы, такие как квантовые точки, нанокристаллы; карбостирил; тербиевый комплекс; неорганический фосфор; ионофоры, такие как краун-эфиры-связанные или производные красители; или их комбинации. Конкретные примеры подходящих люминофоров включают, но не ограничиваются перечисленными, Pd(II) октаэтилпорфирин; Pt(II) октаэтилпорфирин; Pd(II) тетрафенилпорфирин; Pt(II) тетрафенилпорфирин; Pd(II) мезо-тетрафенилпорфиринтетрабензопорфин; Pt(II) мезо-тетрафенилметрилбензопорфирин; Pd(II) октаэтилпорфиринкетон; Pt(II) октаэтилпорфиринкетон; Pd(II) мезо-тетра(пентафторфенил)порфирин; Pt(II) мезо-тетра(пентафторфенил)порфирин; Ru(II) трис(4,7-дифенил-1,10-фенантролин) (Ru(dpp)<sub>3</sub>); Ru(II) трис(1,10-фенантролин) (Ru(phen)<sub>3</sub>), гексагидрат хлорида трис(2,2'-бипиридин)рутения (II) (Ru(bpy)<sub>3</sub>); эритрозин В; флуоресцеин; флуоресцеинизотиоцианат (FITC); эозин; иридий (III) ((N-метилбензимидазол-2-ил)-7-(диэтиламино)-кумарин)); 174энзотиазол) ((бензотиазол-2-ил)-7-(диэтиламино)-кумарин))-2-(ацетилацетонат); люмогеновые красители; Макрофлекс флуоресцентный красный; Макролекс флуоресцентный желтый; Техасский красный; родамин В; родамин 6G; модифицированный серой родамин; м-крезол; тимоловый синий; ксиленоловый синий; крезоловый красный; хлорфеноловый синий; бромкрезоловый зеленый; бромкрезоловый красный; бромтимоловый синий; Cu<sub>2</sub>; Cu<sub>3</sub>; Cu<sub>5</sub>; Cu<sub>5,5</sub>; Cu<sub>7</sub>; 4-нитирофенол; ализарин; фенолфталеин; о-крезолфталеин; хлорфеноловый красный; кальмагит; бромксиленол; феноловый красный; нейтральный красный; нитразин; 3,4,5,6-тетрабромфенолфталеин; конго красный; флуоресцеин; эозин; 2',7'-дихлорфлуоресцеин; 5(6)-карбоксифлуоресцеин; карбоксинафтофлуоресцеин; 8-гидроксипирен-1,3,6-трисульфоновую кислоту; семинафтородафтор; семинафтофлуоресцеин; дихлорид трис(4,7-дифенил-1,10-

фенантролин)рутения (II); (4,7-дифенил-1,10-фенантролин)рутений (II) тетрафенилбор; октаэтилпорфин платины (II); диалкилкарбоцианин; диоктадецилциклоксакарбоцианин; флуоренилметилоксикарбонилхлорид; 7-амино-4-метилкурмарин (Amc); зеленый флуоресцентный белок (GFP); и их производные или комбинации.

В некоторых примерах детектируемый фрагмент может включать родамин В (Rho), флуоресцеинизотиоцианат (FITC), 7-амино-4-метилкурмарин (Amc), зеленый флуоресцентный белок (GFP) или их производные или комбинации.

#### **Способы изготовления**

Соединения, описанные в настоящем документе, могут быть получены различными способами, известными специалисту в области органического синтеза, или их вариации, как будет понятно специалистам в данной области техники. Соединения, описанные в настоящем документе, можно получить из легкодоступных исходных материалов. Оптимальные условия реакции могут варьироваться в зависимости от конкретных используемых реагентов или растворителей, но такие условия могут быть определены специалистом в данной области техники.

Вариации соединений, описанных в настоящем документе, включают добавление, удаление или перемещение различных составляющих, описанных для каждого соединения. Аналогичным образом, если в молекуле присутствуют один или более хиральных центров, хиральность молекулы может быть изменена. Кроме того, синтез соединения может включать защиту и удаление защиты различных химических групп. Применение защиты и удаления защиты, а также выбор подходящих защитных групп может быть определен специалистом в данной области техники. Химические структуры защитных групп можно найти, например, в руководстве Wuts and Greene, *Protective Groups in Organic Synthesis*, 4th Ed., Wiley & Sons, 2006, которое полностью включено в настоящий документ посредством ссылки.

Исходные материалы и реагенты, используемые при приготовлении раскрытых соединений и композиций, доступны от коммерческих поставщиков, таких как Aldrich Chemical Co., (Милуоки, Висконсин), Acros Organics (Моррис Плейнс, Нью-Джерси), Fisher Scientific (Питтсбург, Пенсильвания), Sigma (Сент-Луис, Миссури), Pfizer (Нью-Йорк, Нью-Йорк), GlaxoSmithKline (Роли, Северная Каролина), Merck (Уайтхаус Стейшн, Нью-Джерси), Johnson & Johnson (Нью-Брансуик, Нью-Джерси), Aventis (Бриджуотер, Нью-Джерси), AstraZeneca (Уилмингтон, Делавэр), Novartis (Базель, Швейцария), Wyeth (Мэдисон, Нью-Джерси), Bristol-Myers-Squibb (Нью-Йорк, Нью-Йорк), Roche (Базель, Швейцария), Lilly (Индианаполис, Индиана), Abbott (Эбботт-Парк, Иллинойс), Schering Plough (Кенилворт, Нью-Джерси) или Boehringer Ingelheim (Ингельхайм, Германия), или получены методами, известными специалистам в данной области техники, в соответствии с процедурами, изложенными в таких источниках как Fieser and Fieser's *Reagents for Organic Synthesis*, тома 1-17 (John Wiley and Sons, 1991); *Rodd's Chemistry of Carbon Compounds*, тома 1-5 и дополнительные материалы (Elsevier Science Publishers, 1989); *Organic Reactions*, тома 1-40 (John Wiley and Sons, 1991); *March's Advanced Organic*

Chemistry, (John Wiley and Sons, 4th Edition); и Larock's Comprehensive Organic Transformations (VCH Publishers Inc., 1989). Другие материалы, такие как фармацевтические носители, раскрытые в настоящем документе, можно получить из коммерческих источников.

Реакции для получения соединений, описанных в настоящем документе, можно проводить в растворителях, которые могут быть выбраны специалистом в области органического синтеза. Растворители могут быть по существу не вступающими в реакцию с исходными материалами (реагирующими веществами), промежуточными соединениями или продуктами в условиях, при которых осуществляют реакции, т. е. температура и давление. Реакции можно проводить в одном растворителе или смеси более чем одного растворителя. Образование продукта или промежуточного соединения можно отслеживать в соответствии с любым подходящим способом, известным в данной области техники. Например, образование продукта можно отслеживать с помощью спектроскопических средств, таких как спектроскопия ядерного магнитного резонанса (например,  $^1\text{H}$  или  $^{13}\text{C}$ ), инфракрасная спектроскопия, спектрофотометрия (например, УФ-видимый диапазон) или масс-спектрометрия, или с помощью хроматографии, такой как высокоэффективная жидкостная хроматография (ВЭЖХ) или тонкослойная хроматография.

Раскрытые соединения могут быть приготовлены с помощью твердофазного пептидного синтеза, при котором  $\alpha$ -N-конец аминокислоты защищен кислотной или основной защитной группой. Такие защитные группы должны обладать устойчивостью к условиям образования пептидных связей и при этом легко поддаваться удалению без разрушения растущей пептидной цепи или рацемизации любого из хиральных центров, содержащихся в них. Подходящие защитные группы представляют собой 9-флуоренилметилоксикарбонил (Fmoc), трет-бутилоксикарбонил (Boc), бензилоксикарбонил (Cbz), бифенилизопропилоксикарбонил, трет-амилоксикарбонил, изоборнилоксикарбонил,  $\alpha,\alpha$ -диметил-3,5-диметоксибензилоксикарбонил, о-нитрофенилсульфонил, 2-циано-трет-бутилоксикарбонил и т. п. 9-флуоренилметилоксикарбонильная (Fmoc) защитная группа является особенно предпочтительной для синтеза раскрытых соединений. Другие предпочтительные защитные группы боковой цепи представляют собой для аминокислот боковой цепи, таких как у лизина и аргинина, 2,2,5,7,8-пентаметилхроман-6-сульфонил (pmc), нитро, п-толуолсульфонил, 4-метоксибензолсульфонил, Cbz, Boc и адамантилоксикарбонил; для тирозина, бензил, о-бромбензилоксикарбонил, 2,6-дихлорбензил, изопропил, трет-бутил (трет-Bu), циклогексил, циклопентил и ацетил (Ac); для серина, трет-бутил, бензил и тетрагидропиранил; для гистидина, тритил, бензил, Cbz, п-толуолсульфонил и 2,4-динитрофенил; для триптофана, формил; для аспарагиновой кислоты и глутаминовой кислоты, бензил и трет-бутил, и для цистеина, трифенилметил (тритил).

В способе твердофазного пептидного синтеза  $\alpha$ -C-концевую аминокислоту присоединяют к подходящей твердой подложке или смоле. Подходящими твердыми

подложками, применимыми для вышеуказанного синтеза, являются те материалы, которые являются инертными по отношению к реагентам и условиям ступенчатых реакций конденсации-удаления защитной группы, а также являются нерастворимыми в используемых средах. Твердые подложки для синтеза содержащих  $\alpha$ -С-концевую карбоксильную группу пептидов представляют собой 4-гидроксиметилфеноксиметил-сополи(стирол-1% дивинилбензол) или 4-(2',4'-диметоксифенил-Fmoc-аминометил)феноксиацетиловую смолу, доступную от Applied Biosystems (Фостер-Сити, Калифорния).  $\alpha$ -С-концевую аминокислоту связывают со смолой посредством N,N'-дициклогексилкарбодиимида (DCC), N,N'-диизопропилкарбодиимида (DIC) или гексафторфосфата О-бензотриазол-1-ил-N,N,N',N'-тетраметилуруния (HBTU), с опосредуемым 4-диметиламинопиридином (DMAP), 1-гидроксibenзотриазолом (HOBT), гексафторфосфатом бензотриазол-1-илокси-трис(диметиламино)фосфония (BOP) или бис-(2-оксо-3-оксазолидинил)фосфинхлоридом (BOPCl), или без него, связыванием в течение от примерно 1 до примерно 24 часов при температуре от 10°C до 50°C в растворителе, таком как дихлорметан или ДМФА. Если твердая подложка представляет собой 4-(2',4'-диметоксифенил-Fmoc-аминометил)феноксиацетиловую смолу, группу Fmoc отщепляют с помощью вторичного амина, предпочтительно пиперидина, перед связыванием с  $\alpha$ -С-концевой аминокислотой, как описано выше. Одним из методов связывания с 4-(2',4'-диметоксифенил-Fmoc-аминометил)феноксиацетиловую смолу со снятой защитой является гексафторфосфат О-бензотриазол-1-ил-N,N,N',N'-тетраметилуруния (HBTU, 1 экв.) и 1-гидроксibenзотриазол (HOBT, 1 экв.) в ДМФА. Связывание последовательных защищенных аминокислот можно осуществлять в автоматическом синтезаторе полипептидов. В одном примере  $\alpha$ -N-конец в аминокислотах растущей пептидной цепи защищен Fmoc. Удаление защитной группы Fmoc с  $\alpha$ -N-концевой стороны растущего пептида осуществляют путем обработки вторичным амином, предпочтительно пиперидином. Затем каждую защищенную аминокислоту вводят в примерно 3-кратном молярном избытке и связывание предпочтительно проводят в ДМФА. Связывающий агент может представлять собой гексафторфосфат О-бензотриазол-1-ил-N,N,N',N'-тетраметилуруния (HBTU, 1 экв.) и 1-гидроксibenзотриазол (HOBT, 1 экв.). В конце твердофазного синтеза полипептид удаляют из смолы и снимают защиту либо последовательно, либо за одну операцию. Удаление полипептида и снятие защиты можно осуществлять за одну операцию путем обработки связанного со смолой полипептида реагентом для отщепления, содержащим тианисол, воду, этандитиол и трифторуксусную кислоту. В тех случаях, когда  $\alpha$ -С-конец полипептида представляет собой алкиламид, смолу отщепляют путем аминолиза алкиламином. В качестве альтернативы, пептид можно удалить путем трансэтерификации, например, с помощью метанола, с последующим аминолизом или путем прямого трансамидирования. Защищенный пептид можно очистить на данном этапе или непосредственно перейти к следующему этапу. Удаление защитных групп боковой цепи может быть выполнено с применением смеси для расщепления, описанной выше. Пептид с полностью удаленной

защитой может быть очищен с помощью последовательных хроматографических этапов с применением любого или всех из следующих типов: ионный обмен на слабоосновной смоле (ацетатная форма); гидрофобная адсорбционная хроматография на недериватизированном полистиролдивинилбензоле (например, Amberlite XAD); адсорбционная хроматография на силикагеле; ионообменная хроматография на карбоксиметилцеллюлозе; распределительная хроматография, например, на Sephadex G-25, LH-20 или распределение противотока; высокоэффективная жидкостная хроматография (ВЭЖХ), в частности, ВЭЖХ с обращенной фазой на колонке, наполненной фазой, связанной с октил- или октадецилсилил-диоксидом кремния.

Упомянутые выше полимеры, такие как группы PEG, могут быть присоединены к олигонуклеотиду, такому как АС, в любых подходящих условиях. Можно применять любые средства, известные в данной области техники, в том числе путем ацилирования, восстановительного алкилирования, реакции присоединения Михаэля, тиоалкилирования или других способов хемоселективной конъюгации/лигирования посредством реакционноспособной группы на фрагменте PEG (например, альдегидной, амина, сложноэфирной, тиоловой,  $\alpha$ -галогенацетильной, малеимидо или гидразиновой группы) с реакционноспособной группой на АС (например, альдегидной, амина, сложноэфирной, тиоловой,  $\alpha$ -галогенацетильной, малеимидо или гидразиновой группой). Активирующие группы, которые можно применять для связывания водорастворимого полимера с одним или более белками, включают, но не ограничиваются перечисленными, сульфон, малеимид, сульфгидрил, тиол, трифлат, трезилат, азидирин, оксиран, 5-пиридил и альфа-галогенированную ацильную группу (например,  $\alpha$ -йодуксусную кислоту,  $\alpha$ -бромуксусную кислоту,  $\alpha$ -хлоруксусную кислоту). Если выбранный полимер присоединен к АС путем восстановительного алкилирования, то он должен содержать один реакционноспособный альдегид, чтобы контролировать степень полимеризации. См., например, Kinstler et al., *Adv. Drug. Delivery Rev.* (2002), 54: 477-485; Roberts et al., *Adv. Drug Delivery Rev.* (2002), 54: 459-476; и Zalipsky et al., *Adv. Drug Delivery Rev.* (1995), 16: 157-182.

Для управления ковалентным связыванием АС или линкера с СРР можно осуществлять реакцию соответствующих аминокислотных остатков СРР с органическим дериватирующим агентом, который способен вступать в реакцию с выбранной боковой цепью или N- или C-концами аминокислот. Реакционноспособные группы на пептидном или конъюгатном фрагменте включают, например, альдегид, амина, сложный эфир, тиол,  $\alpha$ -галогенацетил, малеимидо или гидразиновую группу. Дериватирующие агенты включают, например, малеимидобензоилсульфосукцинимидный сложный эфир (конъюгация через остатки цистеина), N-гидроксисукцинимид (через остатки лизина), глутаральдегид, ангидрид янтарной кислоты или другие агенты, известные в данной области техники.

Способы получения АС и конъюгации АС с линейным СРР в общих чертах описаны в публикации США № 2018/0298383, которая полностью включена в настоящий



документ посредством ссылки. Способы могут быть применены к циклическим CPP, раскрытым в настоящем документе.

Схемы синтеза приведены на **ФИГ. 3А-3D** и **ФИГ. 4**.

Неограничивающие примеры соединений, которые включают CPP и реакционноспособную группу, пригодную для конъюгации с АС, показаны в **Таблице 9**. Также показаны примеры линкерных групп. Примерные реакционноспособные группы включают тетрафторфениловый сложный эфир (TFP), свободную карбоновую кислоту (COOH) и азид (N<sub>3</sub>). В **Таблице 9** n представляет собой целое число от 0 до 20; P<sub>1</sub>rab представляет собой AcRXRRBRRXRYQFLIRXRBRXRВ, где В представляет собой β-аланин, и X представляет собой аминоксановую кислоту; Dap представляет собой 2,3-диаминопропионовую кислоту; NLS представляет собой последовательность ядерной локализации; βА представляет собой бета-аланин; -ss- представляет собой дисульфид; PABC представляет собой С-концевой домен поли(А)-связывающего белка; C<sub>x</sub>, где x это число, представляет собой алкильную цепь длиной x; и BCN представляет собой бицикло[6.1.0]нонин.

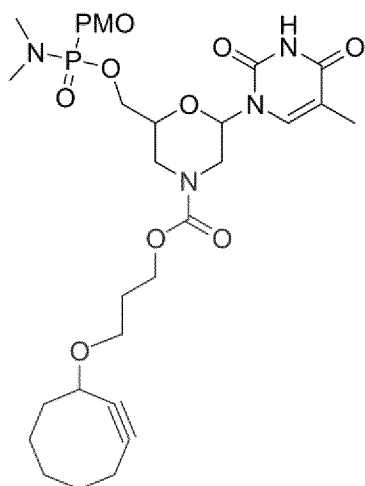
Таблица 9: Соединения, которые включают CPP и реакционноспособную группу

TFP-PEG <sub>n</sub> -K(CPP)
TFP-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -Dap(пальмитоил)
TFP-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -Dap(CPP)
TFP-P <sub>1</sub> rab
CPP-PEG <sub>n</sub> -TFP
CPP-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -TFP
CPP-PEG <sub>n</sub> -Lys(N <sub>3</sub> )
CPP-K(CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(PEG <sub>n</sub> -CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(PEG <sub>n</sub> -CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-K(CPP)-K(CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(PEG <sub>n</sub> -CPP)-K(PEG <sub>n</sub> -CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(PEG <sub>n</sub> -CPP)-K(PEG <sub>n</sub> -CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
Ac-NLS-Lys(CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
K(N <sub>3</sub> )-PEG <sub>n</sub> -NLS-ss-PEG <sub>n</sub> -CPP
BCN-NLS-ss-CPP
CPP-PEG <sub>n</sub> -Val-Cit-PABC-K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -Cys-ss-Cys-K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -Cys-ss-Cys-K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -TFP

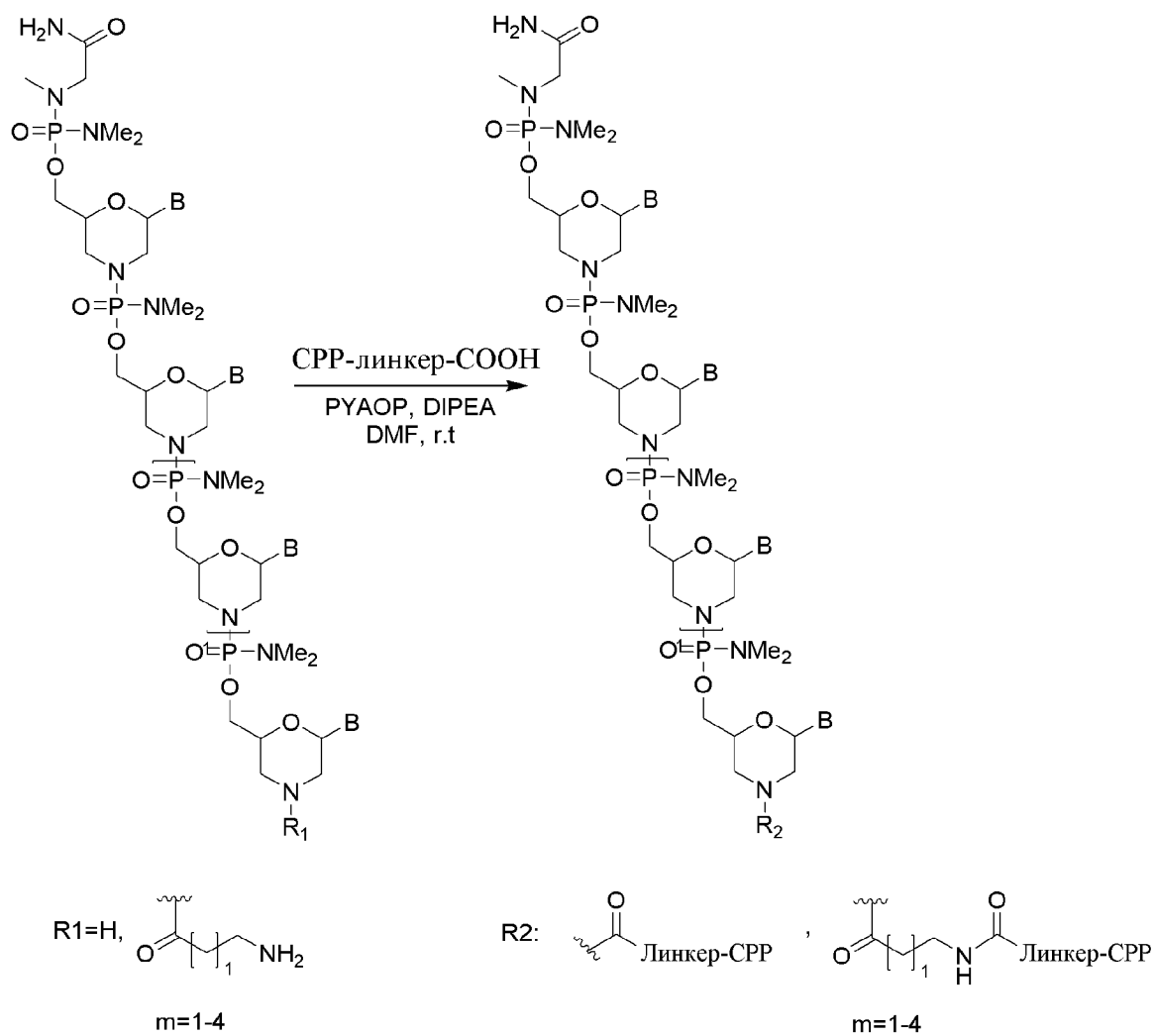
CPP-PEG <sub>n</sub> -Lys(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -Cys-продисульфид-K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-K(CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -TFP
CPP-C <sub>6</sub> -TFP
CPP-PEG <sub>n</sub> -K(PEG <sub>n</sub> -CPP)PEG <sub>n</sub> -K(N <sub>3</sub> )
Ac-T9-PEG <sub>n</sub> -Lys(CPP-PEG <sub>n</sub> )-K(N <sub>3</sub> )
Ac-MSP-PEG <sub>n</sub> -K(CPP-PEG <sub>n</sub> )-K(N <sub>3</sub> )
CPP-PEG <sub>n</sub> -TFP (ENTRD 802)
CPP-C <sub>6</sub> -TFP (ENTRD 696)
CPP-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -TFP (ENTRD-344)
CPP-PEG <sub>n</sub> -COOH
CPP-C <sub>12</sub> -TFP (ENTD-695)
пальмитоил-PEG <sub>n</sub> -K(CPP)-PEG <sub>n</sub> -TFP (ENTD-343)
CPP-PEG <sub>n</sub> -K(N <sub>3</sub> ) (ENTRD-617)
Ac-T9-PEG <sub>n</sub> -K(CPP)-K(N <sub>3</sub> ) (ENTRD 673)
Ac-MSP-PEG <sub>n</sub> -K(CPP-PEG <sub>n</sub> )-K(N <sub>3</sub> ) (ENTRD 675)
Ac-NLS-K(CPP)-PEG <sub>n</sub> -K(N <sub>3</sub> ) (ENTRD 684)
K(N <sub>3</sub> )-PEG <sub>n</sub> -NLS-ss-PEG <sub>n</sub> -CPP (ETRD-681)
K(N <sub>3</sub> )-PEG <sub>n</sub> -NLS-K-βA-βA-CPP (ETRD-682)

Согласно вариантам реализации CPP содержат свободные карбоксильные группы, которые могут быть использованы для конъюгации с АС. Согласно вариантам реализации EEV содержат свободные карбоксильные группы, которые могут быть использованы для конъюгации с АС.

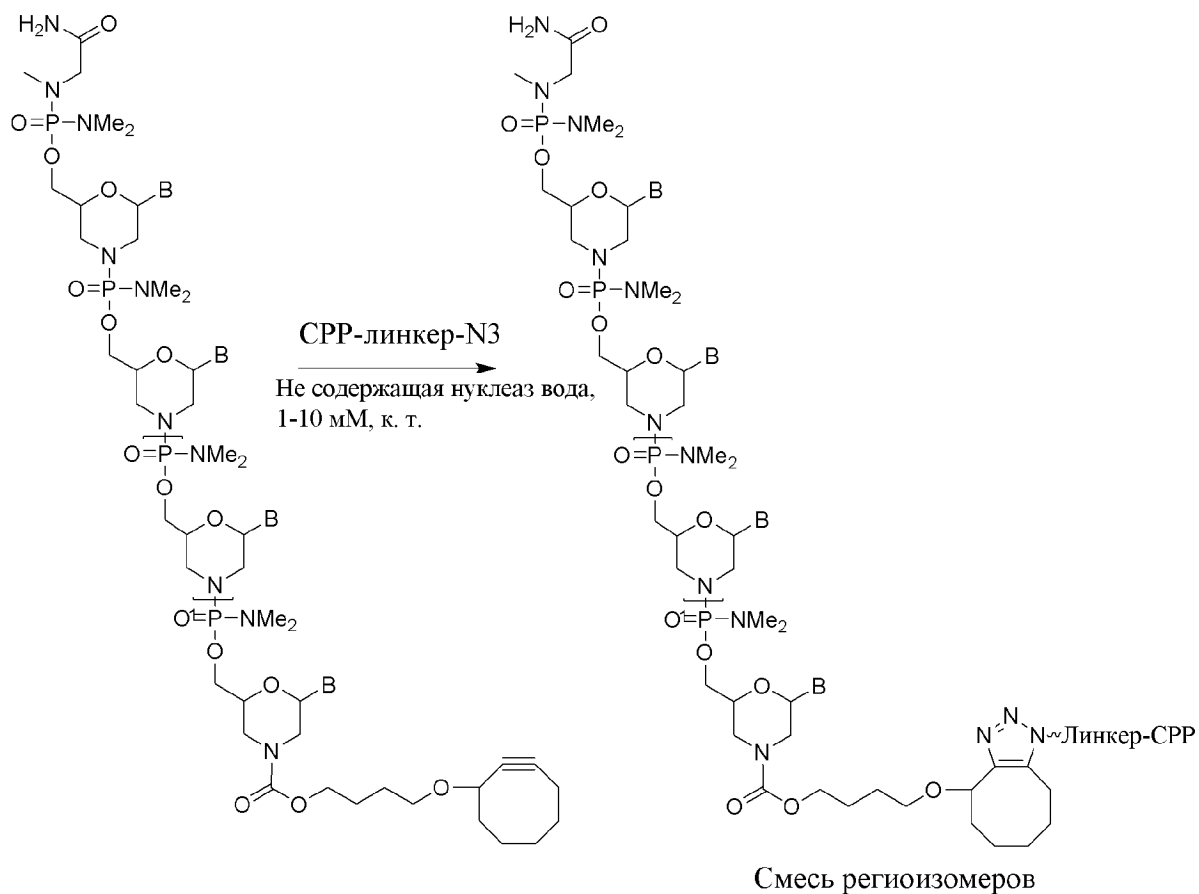
Структура ниже представляет собой 3'-циклооктин-модифицированный РМО, применяемый для клик-реакции с соединением, которое включает азид:



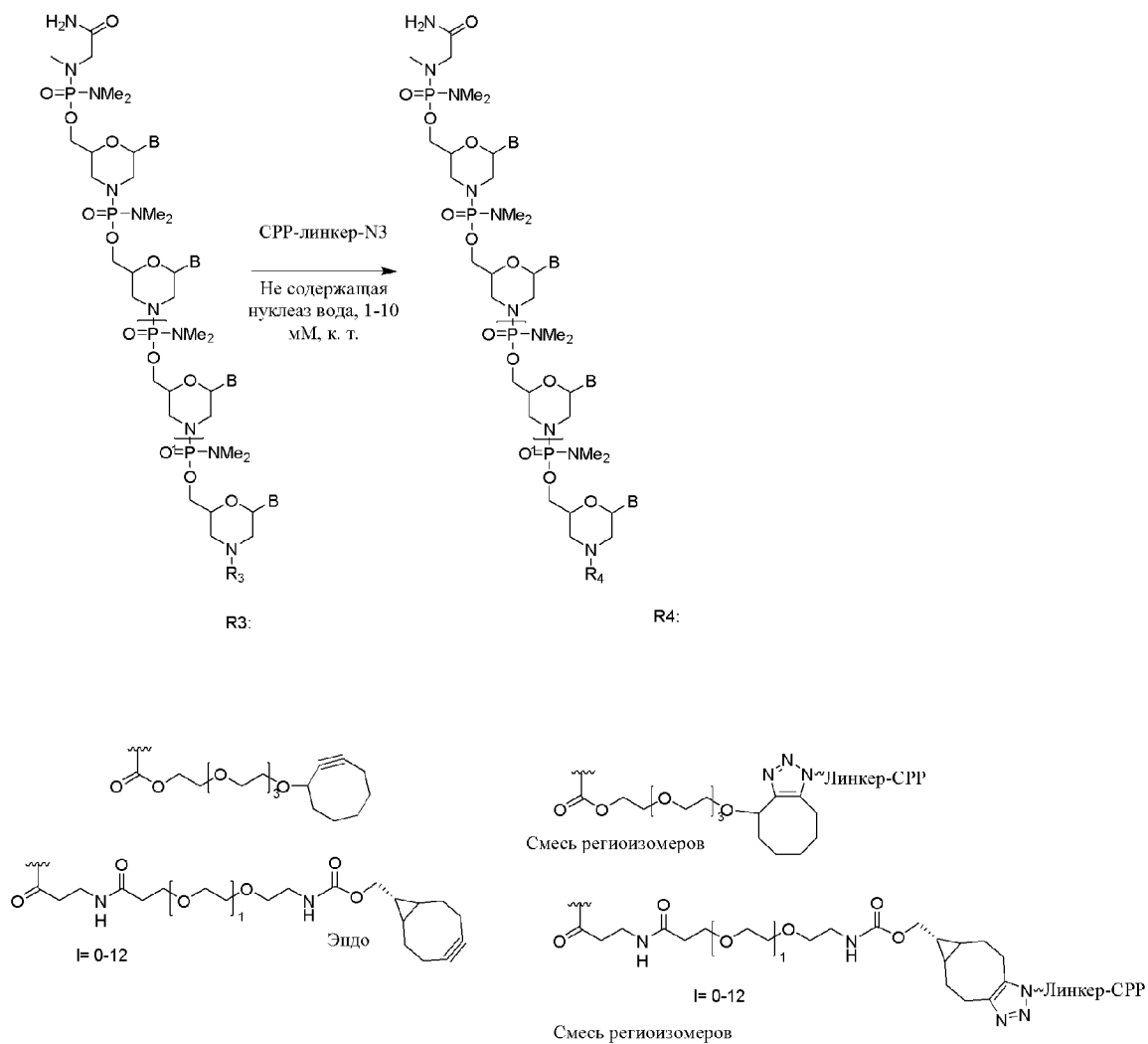
Примерная схема конъюгации CPP и линкера с 3'-концом АС посредством амидной связи показана ниже.



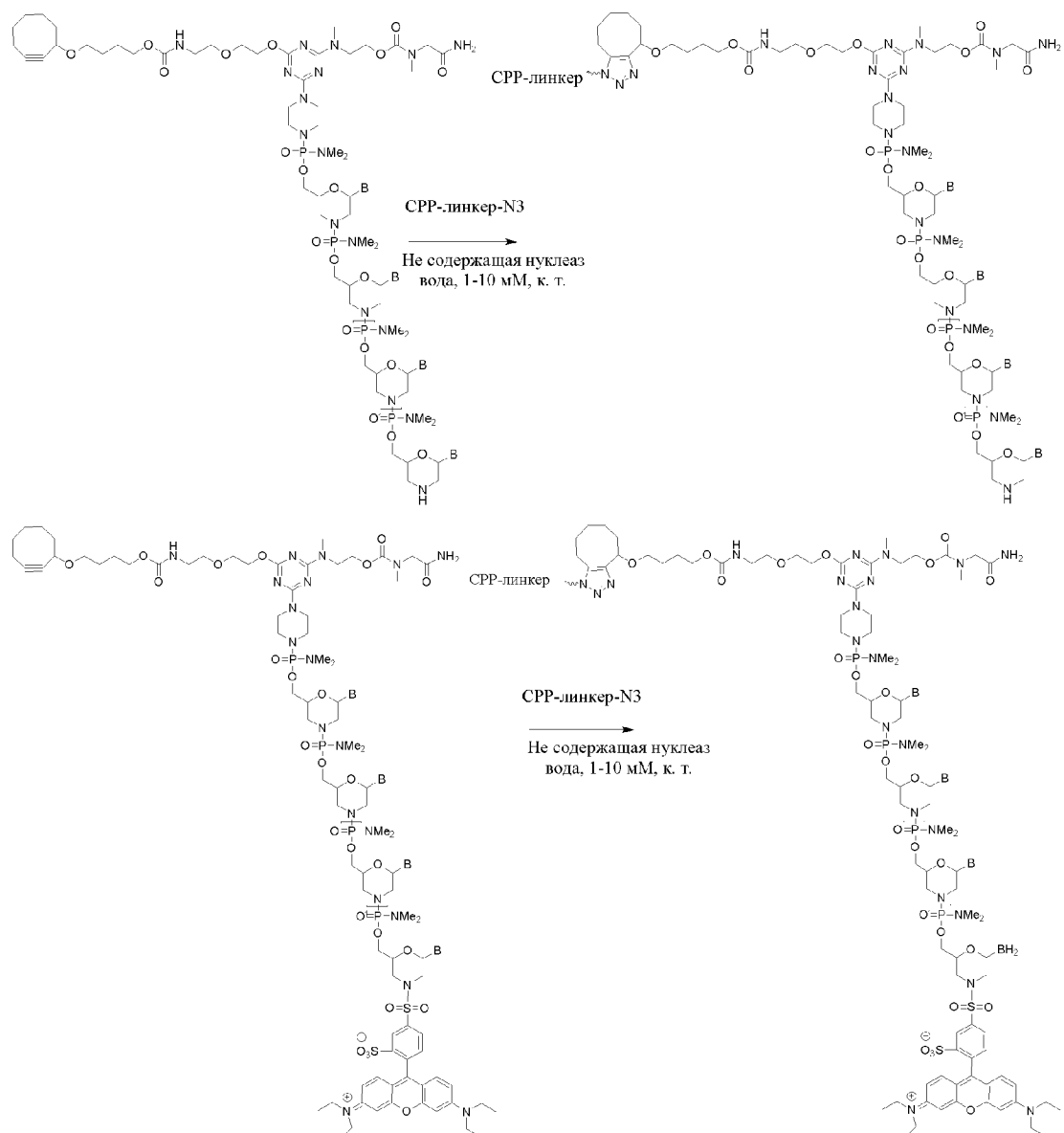
Примерная схема конъюгации CPP и линкера с 3'-циклооктин-модифицированным РМО посредством стимулированного деформацией азид-алкинового циклоприсоединения показана ниже:



Пример химической структуры для конъюгации, используемой для соединения АС и CPP с дополнительным линкером, содержащим фрагмент полиэтиленгликоля, показан ниже:



Пример конъюгации CRR-линкера с 5'-циклооктин-модифицированным РМО посредством стимулированного деформацией азид-алкинового циклоприсоединения (клик-химия) показан ниже:



Способы синтеза олигомерных антисмысловых соединений известны в данной области техники. Настоящее раскрытие не ограничено способом синтеза АС. Согласно вариантам реализации настоящего изобретения предложены соединения, имеющие реакционноспособные фосфорные группы, подходящие для образования межнуклеозидных связей, включая, например, фосфодиэфирные и фосфотиоатные межнуклеозидные связи. Способы получения и/или очистки предшественников или антисмысловых соединений не являются ограничением композиций или способов, предложенных в настоящем документе. Способы синтеза и очистки ДНК, РНК и антисмысловых соединений хорошо известны специалистам в данной области техники.

Олигомеризацию модифицированных и немодифицированных нуклеозидов можно рутинно выполнять в соответствии с процедурами, описанными в литературе, для ДНК (Protocols for Oligonucleotides and Analogs, Ed. Agrawal (1993), Humana Press) и/или РНК (Scaringe, Methods (2001), 23, 206-217. Gait et al., Applications of Chemically synthesized

RNA in RNA: Protein Interactions, Ed. Smith (1998), 1-36. Gallo et al., Tetrahedron (2001), 57, 5707-5713).

Антисмысловые соединения, предложенные в настоящем документе, можно удобно и рутинно получать с помощью хорошо известной методики твердофазного синтеза. Оборудование для такого синтеза продается несколькими поставщиками, включая, например, Applied Biosystems (Фостер-Сити, Калифорния). Для такого синтеза можно дополнительно или альтернативно использовать любые другие средства, известные в данной области техники. Хорошо известно, что подобные методики применяют для получения олигонуклеотидов, таких как фосфотиоаты и алкилированные производные. Настоящее изобретение не ограничено способом синтеза антисмыслового соединения.

Способы очистки и анализа олигонуклеотидов известны специалистам в данной области техники. Способы анализа включают капиллярный электрофорез (СЕ) и масс-спектрометрию с электрораспылением. Такие способы синтеза и анализа можно выполнять в многолуночных планшетах. Способ согласно настоящему изобретению не ограничивается способом очистки олигомера.

В соединениях, раскрытых в настоящем документе, АС связан с СРР (например, циклическим пептидом). В настоящем документе термин «связанный» может относиться к ковалентной или нековалентной ассоциации между СРР и АС, включая слияние СРР с АС и химическую конъюгацию СРР (например, циклического пептида) с АС. Неограничивающий пример средства для нековалентного присоединения СРР к АС представляет собой взаимодействие стрептавидин/биотин, например, с помощью конъюгации биотина с СРР и слияния АС со стрептавидином. В полученном соединении СРР связан с АС посредством нековалентной ассоциации между биотином и стрептавидином.

Согласно вариантам реализации СРР (например, циклический пептид) конъюгирован, прямо или опосредованно, с АС с образованием тем самым конъюгата СРР-АС. Конъюгация АС с СРР может происходить в любом подходящем сайте на этих фрагментах. Например, в вариантах реализации 5'- или 3'-конец АС может быть конъюгирован с С-концом, N-концом или боковой цепью аминокислоты в СРР.

Согласно вариантам реализации АС ковалентно связан с СРР (например, циклическим пептидом). Ковалентная связь, в контексте настоящего документа, относится к конструкциям, в которых фрагмент СРР ковалентно связан с 5'- и/или 3'-концом фрагмента АС. В качестве альтернативы, такие конъюгаты могут быть описаны как содержащие фрагмент СРР (например, циклический пептидный фрагмент) и олигонуклеотидный фрагмент. Ковалентно связанный конъюгат АС-СРР или СРР-АС в соответствии с определенными вариантами реализации включает компонент АС и компонент СРР, ассоциированные друг с другом линкером, описанным в настоящем документе.

Согласно вариантам реализации АС может быть конъюгирован с СРР (например, циклическим пептидом) посредством боковой цепи аминокислоты на СРР. Любая боковая

цепь аминокислоты на СРР, которая способна образовывать ковалентную связь или которая может быть модифицирована таким образом, может быть использована для связывания АС с СРР. Аминокислота на СРР может представлять собой природную или неприродную аминокислоту. Согласно вариантам реализации аминокислота на СРР, используемая для конъюгации АС, представляет собой аспарагиновую кислоту, глутаминовую кислоту, глутамин, аспарагин, лизин, орнитин, 2,3-диаминопропионовую кислоту или их аналоги, в которых боковая цепь замещена связью с АС или линкером. Согласно вариантам реализации аминокислота представляет собой лизин или его аналог. Согласно вариантам реализации аминокислота представляет собой глутаминовую кислоту или ее аналог. Согласно вариантам реализации аминокислота представляет собой аспарагиновую кислоту или ее аналог.

Согласно вариантам реализации СРР является циклическим. Существует множество возможных конфигураций соединений, раскрытых в настоящем документе. Согласно вариантам реализации соединения согласно настоящему изобретению включают соединения, в которых АС конъюгирован с боковой цепью аминокислоты в циклическом пептиде. Согласно вариантам реализации соединения, раскрытые в настоящем документе, имеют структуру (т.е. экзоциклическую) в соответствии с Формулой I-A:

**СРР-L-АС**

**(I-A)** ,

в которой линкер ковалентно связан с боковой цепью аминокислоты на СРР и с 5'-концом АС, остовом АС или 3'-концом АС.

#### **Заболевания и гены-мишени**

Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с одним или более генами, содержащими экспансированные нуклеотидные повторы (например, экспансированные тринуклеотидные повторы, такие как экспансированные тринуклеотидные повторы). Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с одним или более генами, содержащими экспансированный тринуклеотидный повтор CTG·CUG. Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с одним или более генами, содержащими экспансированный тринуклеотидный повтор CTG·CUG в 3'-НТО гена. Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с геном, который содержит экспансированный CTG·CUG в 3'-НТО, таким как DMPK, ATXN8OS ATXN8 и/или JPH3. Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с одним или более генами, содержащими экспансированный тринуклеотидный повтор CTG·CUG в интроне гена. Согласно вариантам реализации предложены соединения и способы для лечения заболевания или нарушения, ассоциированного с экспансированным тринуклеотидным повтором CTG·CUG в интроне



TCF4. Согласно вариантам реализации предложены соединения и способы для лечения миотонической дистрофии 1 типа (DM1), эндотелиальной дистрофии роговицы Фукса (FECD), спиноцереbellарной атаксии-8 (SCA8) и/или болезни, подобной болезни Хантингтона (HDL2).

#### *Миотоническая дистрофия 1 типа (DM1)*

Согласно вариантам реализации предложены соединения, композиции и способы для лечения миотонической дистрофии (DM1 или болезни Штейнерта). DM1 представляет собой мультисистемное нарушение, часто характеризующееся дегенерацией мышц и миотонией или замедленной мышечной релаксацией из-за повторяющихся потенциалов действия в мышечных волокнах. Миотоническая дистрофия 1 типа (DM1) является наиболее распространенной формой мышечной дистрофии, поражающей примерно 1 из 8000 человек. DM1 представляет собой пример генетических нарушений, вызванных экспансиями CTG·CUG. DM1 представляет собой нейромышечное нарушение, вызванное экспансией повтора CTG·CUG в 3'-нетранслируемой области (НТО) гена протеинкиназы мышечной дистрофии (DMPK). На уровне РНК транскрипт DMPK (например, экспансированный повтор CUG) секвестрирует регулирующие сплайсинг белки, например, muscleblind-подобный белок (MBNL), что приводит к ошибочному сплайсингу ряда нисходящих пре-мРНК (пре-мРНК, которые не содержат экспансированного повтора CUG), которые регулируются MBNL1. Это приобретение функции является причиной DM1.

Чрезмерное число повторов CUG придает токсическую активность, называемую токсическим приобретением функции. Множество ключевых белков подвергаются неправильному процессингу, и это вносит вклад в мультисистемный характер заболевания, который включает генерализованную слабость конечностей, нарушение работы респираторных мышц, нарушения сердечной деятельности, утомляемость, желудочно-кишечные осложнения, катаракту, недержание мочи и чрезмерную дневную сонливость.

Пациенты, страдающие DM1, характеризующейся экспансиями CTG·CUG в 3'-нетранслируемой области гена DMPK, подвержены повышенному риску развития FECD и у них образуются очаги CUGexp-MBNL1 в эндотелии роговицы. (Mootha et al., *Investigative ophthalmology & visual science*, 2017; 58, 4579-4585). Ассоциация MBNL1 с мутантной РНК влияет на клеточный пул свободного MBNL1 и запускает ошибочный сплайсинг некоторых генов-мишеней MBNL1 (например, регулируемых MBNL1) в пораженных тканях головного мозга, мышц и сердца (Jiang et al. *Hum Mol Genet.* 2004; 13: 3079-3088). Gattey et al. (*Cornea.* 2014; 33: 96-98) сообщили о FECD у четырех субъектов, страдающих DM1, в том числе у пары мать-дочь. Таким образом, вероятно, будет наблюдаться ассоциация между DM1 и FECD (FECD более подробно описывается в другом месте настоящего документа).

Не ограничиваясь какой-либо теорией, существует по меньшей мере две гипотезы, объясняющие патогенез DM1. Одна из них заключается в том, что экспансированные

повторы CTG·CUG ингибируют продуцирование мРНК или белка DMPK, что приводит к гаплонедостаточности DMPK. Это подтверждалось исследованиями, демонстрирующими уменьшенную экспрессию мРНК и белка DMPK в мышце при DM1 (Fu, Y.H. et al. (1993) Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235-238). Согласно вариантам реализации соединения и способы, описанные в настоящем документе, уменьшают гаплонедостаточность DMPK. Другая гипотеза о приобретении функции РНК предполагает, что мутантная РНК, транскрибируемая из экспансированного аллеля, достаточна для индукции симптомов заболевания. Об этом свидетельствовали следующие наблюдения: (i) экспансированные повторы CTG транскрибируются в повторы CUG, которые накапливаются в дискретных ядерных очагах, (ii) экспрессии только 3'-НТО DMPK с 200 повторами CTG достаточно для ингибирования миогенеза (Davis, B.M., et al. (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, 7388-7393; Amack, J.D. et al., (1999) Cis and trans effects of the myotonic dystrophy (DM) mutation in a cell culture model. *Hum. Mol. Genet.* 8, 1975-1984). Согласно вариантам реализации соединения и способы, описанные в настоящем документе, снижают транскрипцию мутантной РНК, которая ассоциирована с экспансированным аллелем.

Экспансированные тринуклеотидные повторы CTG·CUG в 3'-нетранслируемой области мРНК DMPK образуют дефектные стабильные шпилечные структуры, которые накапливаются в ядре клетки в небольших рибонуклеарных комплексах или микроскопически видимых включениях и нарушают функцию белков, вовлеченных в транскрипцию, сплайсинг или экспорт РНК. Несмотря на то, что гены DMPK с повторами CUG транскрибируются в мРНК, мутантные транскрипты секвестрируются в ядре в виде агрегатов (очагов), что приводит к уменьшению цитоплазматических уровней мРНК DMPK. Эти агрегаты приводят к нарушению регуляции альтернативного сплайсинга многих различных транскриптов вследствие секвестрации двух РНК-связывающих белков: MBNL1 (подобный muscleblind 1) и CUGBP1 (CUG-связывающий белок 1), что приводит к утрате функции MBNL1 и положительной регуляции CUGBP1 (Lee and Cooper. (2009) «Pathogenic mechanisms of myotonic dystrophy», *Biochem Soc Trans.* 37(06): 1281-1286).

При DM1 РНК-связывающий белок MBNL1 секвестрирован в двухцепочечной шпилечной структуре, образованной повторами CUG, что приводит к его истощению из нуклеоплазмы. Затем повторы CUG, с которыми связан MBNL1, стимулируют активацию протеинкиназы С (PKC) посредством неизвестного механизма, который индуцирует гиперфосфорилирование и стабилизацию CUGBP1. Последующие эффекты включают нарушение альтернативного сплайсинга, трансляции мРНК и распада мРНК нисходящих генов. Важным молекулярным признаком DM1 является нарушение регуляции альтернативного сплайсинга из-за секвестрации MBNL1 в повторах CUG с двухцепочечной шпилечной структурой. Из более чем двух десятков событий сплайсинга,

регуляция которых нарушена при DM1, ошибочный сплайсинг ClC-1, специфического для скелетных мышц (канал для ионов хлора 1), как известно, является одной из причин миотонии. Повышенное включение экзонов, содержащих преждевременные стоп-кодоны, приводит к отрицательной регуляции мРНК и белка ClC-1, что является достаточным для того, чтобы вызвать миотонию (Charlet-B et al. (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. et al. (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). Согласно вариантам реализации соединения и способы, описанные в настоящем документе, уменьшают последующие эффекты, включая нарушение альтернативного сплайсинга, трансляции мРНК и распада мРНК нисходящих генов. Согласно вариантам реализации соединения и способы, описанные в настоящем документе, снижают число событий сплайсинга, ошибочно регулируемого при DM1, по сравнению с субъектом, страдающим DM1, который не получал лечения соединениями или способами согласно настоящему изобретению. Например, в некоторых вариантах реализации соединения и способы, описанные в настоящем документе, могут снижать число событий ошибочно регулируемого сплайсинга в одном или более нисходящих генах, таких как 4833439L19Rik, Abcc9, Atp2a1, Arhgef10, Arhgap28, Armcx6, Angell, Best3, Bin1, Brd2, Casna1s, Casna2d1, Cpd, Creb3, Ccpg1, Clasp1, Clcn1, Clk4, Creb2, 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, Klcl, 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, Pdim3, 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 или Vsp39. Согласно вариантам реализации соединения и способы, описанные в настоящем документе, снижают число экзонов, содержащих преждевременные стоп-кодоны, что приводит к отрицательной регуляции мРНК ClC-1 по сравнению с субъектом, страдающим DM1, который не получал лечения соединениями или способами согласно настоящему изобретению.

Уровни MBNL1 и CUGBP1 в ядре контролируют подгруппу событий сплайсинга, регулируемых в процессе развития, которые обращаются вспять при DM1. На эмбриональной стадии ядерные уровни MBNL1 являются низкими, а уровни CUGBP1 являются высокими. Во время развития ядерные уровни MBNL1 повышаются, в то время как уровни CUGBP1 уменьшаются, что индуцирует переход от эмбрионального профиля к взрослому для нисходящих мишеней сплайсинга (включая экзон 11 IR, экзоны ClC-1, содержащие стоп-кодоны, и экзон 5 cTNT). Однако при DM1 MBNL1 секвестрируется в повторах CUG, что приводит к уменьшению уровня функционального MBNL1, тогда как уровни CUGBP1 повышаются вследствие фосфорилирования и стабилизации. Это

имитирует эмбриональное состояние и повышает экспрессию эмбриональных изоформ у взрослых, что приводит к множеству симптомов заболевания (Lee and Cooper.; 2009). Согласно вариантам реализации соединения и способы, описанные в настоящем документе, обеспечивают снижение количества секвестрированного MBNL1, увеличение количества функционального MBNL1, уменьшение уровня CUGBP1 по сравнению с субъектом, страдающим DM1, который не получал лечения соединениями или способами согласно настоящему изобретению.

MBNL1 и CELF1 (также называемый «CUGBP1») являются связанными с развитием регуляторами событий сплайсинга во время перехода от эмбриона к взрослому, и модификация их активности при DM1 приводит к экспрессии эмбрионального профиля сплайсинга в тканях взрослых. Последующее влияние низких уровней MBNL1 и высоких уровней CELF1 включает нарушение альтернативного сплайсинга, трансляции мРНК и распада мРНК белков, например, транскриптов сердечного тропонина Т (сТNT), инсулинового рецептора (INSR), специфического для мышц канала для ионов хлора (CLCN1) и кальциевой АТФазы 1 саркоплазматического/эндоплазматического ретикулула (ATP2A1), помимо MBNL1. Konieczny et al. (2017) «Myotonic dystrophy: candidate small molecule therapeutics», *Drug Discovery Today*. 22(11):1740-1748.

Соединения и способы для лечения миотонической дистрофии с применением антисмысловых олигомеров, нацеленных на повторы полиCUG в 3'-НТО гена DMPK, описаны в US10106796B2, US10111962B2, US20150080311A1, каждый из которых полностью включен в настоящий документ посредством ссылки для всех целей. Однако такие РМО или РРМО, нацеленные на повторы CUG, для лечения DM1 могут иметь ограничения по доставке олигонуклеотидов в мышцы, которые представляют собой пораженные заболеванием ткани.

Согласно вариантам реализации настоящее изобретение предусматривает применение разнообразных проникающих в клетку пептидов (СРР) для доставки в цитозоль клетки АС (например, РМО или АСО) и последовательности, обеспечивающей разрушение, описанной в настоящем документе, например, в **Таблицах 2 и 10**. Согласно вариантам реализации СРР или ЕЕV, конъюгированные с АС, обеспечивают доставку представляющего интерес АС в клеточное местоположение, где расположена целевая последовательность на пре-мРНК.

Согласно вариантам реализации заболевание представляет собой форму миотонической дистрофии (например, миотонической дистрофии 1 типа или миотонической дистрофии 2 типа). Согласно вариантам реализации целевой ген представляет собой ген DMPK, который кодирует миотоническую протеинкиназу. Согласно вариантам реализации соединения, предложенные в настоящем документе, содержат АС (например, АСО), который нацелен на DMPK (например, 3'-нетранслируемую область/полиаденилирование гена DMPK) для разрушения гена DMPK. Примерные олигонуклеотиды, которые нацелены на DMPK для разрушения, представлены в **Таблице 10**. Последовательность, обеспечивающую разрушение, можно

применять в комбинации с последовательностью AC, содержащей от 10 до 40 повторов CAG, включая без ограничения AC, представленный в **Таблице 2**.

Таблица 10. Олигонуклеотиды (AC), нацеленные на DMPK для разрушения.

ID олигонуклеотида (AC)	Последовательность (5'-3')	SEQ ID NO:	Мишень
	5'-CAG CAG CAG CAG CAG CAG CAG CAG-3'- клик-K-PEG12-Lys(CPP12)-NLS-AC (все мономеры РМО)		DMPK
DMPK-A-17	GGGCCTTTTATTCGCGAGGGTCGGG	151	DMPK
DMPK-A-18	GAGGGCCTTTTATTCGCGAGGGTCG	152	DMPK
DMPK-A-19	TGGAGGGCCTTTTATTCGCGAGGGT	153	DMPK
DMPK-A-20	GATGGAGGGCCTTTTATTCGCGAGG	154	DMPK
DMPK-A-21	CAGATGGAGGGCCTTTTATTCGCGA	155	DMPK
DMPK-A-22	GGCAGATGGAGGGCCTTTTATTCGC	156	DMPK
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
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
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

*Стинocereбеллярная атаксия-8 (SCA8)*

Согласно вариантам реализации предложены соединения, композиции и способы для лечения спиноцереbellарной атаксии-8 (SCA8). SCA8 представляет собой наследственное нейродегенеративное состояние, которое характеризуется медленно прогрессирующей атаксией. Симптомы обычно возникают в течение третьего-пятого десятилетий жизни. Симптомы включают нарушение движения глаз, сенсорную нейропатию, дисфагию, церебеллярную атаксию и нарушение когнитивных функций.

SCA8 ассоциирована с гетерозиготным патологически экспансированным повтором CTG·CUG в 3'-НТО двух перекрывающихся генов ATXN8OS и ATXN8. Здоровые индивидуумы обычно имеют от 15 до 50 повторов CTG·CUG в генах ATXN8OS и ATXN8. Пациенты с SCA8 имеют более 50 повторов CTG·CUG, иногда до 240 повторов CTG·CUG в генах ATXN8OS и ATXN8.

*Болезнь, подобная болезни Хантингтона, 2 типа (HDL2)*

Согласно вариантам реализации предложены соединения, композиции и способы для лечения болезни, подобной болезни Хантингтона, 2 типа (HDL2). HDL2 представляет собой аутосомно-доминантное нейродегенеративное нарушение, которое фенотипически связано с болезнью Хантингтона. HDL2 характеризуется симптомами, которые включают хорею, дистонию, ригидность, брадикинезию и психиатрические симптомы, такие как деменция. Симптомы HDL2, как правило, возникают в середине жизни и могут привести к преждевременной смерти в течение примерно 10-15 лет.

HDL2 ассоциирована с экспансированным CTG·CUG в 3'-НТО гена джанктофилина 3 (JPH3) (16q24.3). Здоровые индивидуумы обычно имеют от 6 до 27 повторов CTG·CUG в гене JPH3. Пациенты, страдающие HDL2, имеют более 40 повторов CTG·CUG, иногда до 60 или более повторов CTG·CUG в гене JPH3.

*Эндотелиальная дистрофия роговицы Фукса*

Согласно вариантам реализации предложены соединения, композиции и способы для лечения эндотелиальной дистрофии роговицы Фукса (FECD). FECD (MIM 136800) представляет собой возрастное дегенеративное нарушение эндотелия роговицы. FECD характеризуется прогрессирующей потерей эндотелиальных клеток роговицы, утолщением Десцеметовой оболочки и отложением внеклеточного матрикса в виде гутт. Когда число эндотелиальных клеток становится критически низким, роговица набухает и это вызывает потерю зрения (Elhalis et al. Ocul Surf. 2010; 8(4):173-184).

FECD может наследоваться как аутосомно-доминантный признак с генетической гетерогенностью. Редкие гетерозиготные мутации в гене коллагена, VIII типа, альфа 2 (COL8A2, MIM 120252) могут привести к ранней эндотелиальной дистрофии роговицы. Другие гены, такие как член семейства переносчиков растворенных веществ 4, член семейства переносчиков бората натрия 11 (SLC4A11, MIM 610206), транскрипционный фактор 8 (TCF8, MIM 189909), домены гомологии липоксигеназы 1 (LOXHD1, MIM 613267) и белок, подобный АТФ/ГТФ-связывающему белку, 1 (AGBL1, MIM 615523) в совокупности ассоциированы с небольшой долей случаев FECD, развившихся у взрослых. Полногеномные ассоциативные исследования FECD, развившейся у взрослых, показали,

что транскрипционный фактор 4 (TCF4, MIM 602272) и, позднее, белок 4, содержащий мотив KN и домен анкириновых повторов (KANK4, MIM 614612), ламинин гамма-1 (LAMC1, MIM150290), Na<sup>+</sup>/K<sup>+</sup>-транспортирующая АТФаза и полипептид бета-1 (ATP1B1, MIM 182330) с локусом TCF4 оказывают преимущественное влияние на FECD (Mootha et al., *Investigative ophthalmology & visual science*, 2017; 58, 4579-4585).

Экспансированные тринуклеотидные повторы в локусе CTG18.1 в интроне 2 TCF4 ассоциированы с FECD (Wieben et al., *PLoS One*. 2012; 7(11):e49083). Каждая копия экспансированного аллеля CTG18.1 из более чем 40 тринуклеотидных повторов CTG-CUG приводит к значительному риску развития FECD (Mootha et al., *Invest Ophthalmol Vis Sci*. 2014; 55: 33-42). Ядерные очаги РНК, отличительная черта токсичного приобретения функции РНК, были описаны при нейродегенеративных нарушениях, вызванных простыми экспансиями повторов. РНК с экспансированным повтором CUG накапливается в виде ядерных очагов в эндотелии роговицы субъектов, страдающих FECD, характеризующейся экспансией триплетных повторов CTG18.1, при этом она отсутствует в контрольных образцах без экспансии триплетов (Mootha et al., *Invest Ophthalmol Vis Sci*. 2015;56(3):2003-2011). РНК с экспансированным повтором CUG совместно локализуется с подобным muscleblind фактором сплайсинга мРНК 1 (MBNL1), в ядерных очагах в эндотелии в качестве молекулярного отличительного признака. Экспансия триплетных повторов в локусе CTG18.1 может опосредовать дисфункцию эндотелия посредством aberrantного сплайсинга генов в результате секвестрации MBNL1 мутантными РНК-транскриптами с CUG (Du et al., *J Biol Chem*. 2015; 290: 5979-5990). Таким образом, общим признаком для двух отдельных триплетных повторов являются очаги РНК и FECD, и вполне вероятно, что эти очаги могут играть причинную роль для FECD.

Согласно вариантам реализации соединения и способы, которые можно применять при лечении эндотелиальной дистрофии роговицы Фукса (FECD), которые снижают уровень РНК с экспансированным повтором CUG с помощью антисмысловых олигонуклеотидов, описаны в WO2018165541A1, US10760076B2, каждый из которых полностью включен в настоящий документ посредством ссылки для всех целей. Однако такие фосфордиамидат-морфолиновые олигомеры (РМО) или пептид-конъюгированные РМО (РРМО), нацеленные на повторы CUG, такие как те, которые описаны в данной области техники для лечения DM1, могут иметь ограничения по доставке олигонуклеотидов в целевую ткань (например, эндотелиальные слои в роговице), которая представляет собой пораженные заболеванием ткани.

Согласно вариантам реализации настоящее изобретение предусматривает применение разнообразных проникающих в клетку пептидов (СРР) или носителя для ускользания из эндосом (ЕЕV) для доставки АС (например, РМО или АSO), описанного в настоящем документе, например, в Таблице 6, в цитозоль клетки. Согласно вариантам реализации СРР или ЕЕV, конъюгированные с АС, обеспечивают доставку представляющего интерес АС в клеточное местоположение, где расположена целевая последовательность на пре-мРНК.

Согласно вариантам реализации заболевание представляет собой эндотелиальную дистрофию роговицы Фукса (FECD). Согласно вариантам реализации целевой ген представляет собой *TCF4*, который кодирует фактор транскрипции 4 (TCF-4), который также известен как фактор транскрипции иммуноглобулина 2 (ITF-2). Согласно вариантам реализации соединения, предложенные в настоящем документе, содержат антисмысловой олигонуклеотид, нацеленный на *TCF4*. Примерные олигонуклеотиды, которые можно применять для нацеливания на *TCF4*, представлены в **Таблице 2** и **Таблице 11**.

Таблица 11. Примерные олигонуклеотиды, нацеленные на экспансированный триплетный повтор TCF4

Химическая структура олигонуклеотида	Конструкция	Мишень
21-мерный РМО	5'-CAG CAG CAG CAG CAG CAG CAG -3' (все мономеры РМО; SEQ ID NO:146)	TCF4 (CUG) <sub>n</sub>
25-мерный РМО	5'-CAG CAG CAG CAG CAG CAG CAG CAG C-3' (все мономеры РМО; SEQ ID NO:147)	TCF4 (CUG) <sub>n</sub>
21-мерный EEV-NLS-РМО	EEV-NLS_РМО_CAG 21-мерный (конъюгация из SEQ ID NO:146)	TCF4 (CUG) <sub>n</sub>
30-мерный РМО	5'-CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG-3' (все мономеры РМО; SEQ ID NO:148)	TCF4 (CUG) <sub>n</sub>
30-мерный РМО	5'-AGC AGC AGC AGC AGC AGC AGC AGC AGC AGC AGC-3' (все мономеры РМО; SEQ ID NO:149)	TCF4 (CUG) <sub>n</sub>
30-мерный РМО	5'-GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA-3' (все мономеры РМО; SEQ ID NO:150)	TCF4 (CUG) <sub>n</sub>

РМО: Фосфордиамидат-морфолиновый олигомер

Mootha et al. (2017) сообщили, что DM1 и FECD происходят из некодирующих экспансий CTG, даже если не являются идентичными заболеваниями. Экспансия DMPK при DM1 приводит к полиорганному заболеванию, которое затрагивает различные ткани в глазу, включая хрусталик, сетчатку и эндотелий роговицы. Напротив, экспансия повторов TCF4, по-видимому, влияет на эндотелий роговицы без каких-либо клинически очевидных последствий для других тканей глаза или органов организма. Мутантные экспансии в DMPK и TCF4 имеют важное сходство, такое как (i) ядерные очаги, которые содержат экспансированные повторы CUG, (ii) ассоциация очагов с белком MBNL1, и (iii) способность вызывать FECD. Предполагается, что экспансии триплетов как в DMPK, так



и в TCF4 могут вызывать одинаковый фенотип эндотелиальной ткани роговицы при FECD посредством общих молекулярных механизмов.

См. патент США № 10760076B2, международную публикацию заявки № WO2018165541A1, публикацию заявки на патент США № 2016/0355796 и публикацию заявки на патент США № 2018/0344817, каждый из которых включен в настоящий документ посредством ссылки и в которых раскрыты заболевания и соответствующие гены, склонные к образованию и/или экспансии tandemных нуклеотидных повторов.

#### **Композиции и способы введения**

Соединения согласно настоящему изобретению могут быть включены в композиции, подходящие для применения *in vivo*. Соединения и/или композиции можно вводить пациенту, который страдает или у которого подозревают наличие заболевания, ассоциированного с экспансированным тринуклеотидным повтором.

Применение *in vivo* раскрытых соединений и композиций, содержащих указанные соединения, можно осуществлять с помощью любого подходящего способа и методики, известных в настоящее время или проспективно специалистам в данной области техники. Например, раскрытые соединения могут быть включены в состав физиологически или фармацевтически приемлемой композиции и введены любым подходящим путем, известным в данной области техники, включая, например, пероральный и парентеральный пути введения. В настоящем документе термин «парентеральный» включает подкожное, внутрикожное, внутривенное, внутримышечное, внутрибрюшинное, интрастернальное и интратекальное введение, например, путем инъекции. Введение раскрытых соединений или композиций может представлять собой однократное введение или непрерывное введение или введение с различными интервалами, которые могут быть легко определены специалистом в данной области техники.

Соединения, раскрытые в настоящем документе, и композиции, содержащие указанные соединения, также можно вводить с применением технологии липосом, капсул с медленным высвобождением, имплантируемых насосов и биоразлагаемых контейнеров. Эти способы доставки могут предпочтительно обеспечивать единообразную дозировку в течение длительного периода времени. Соединения также можно вводить в виде их солевых производных или кристаллических форм.

Соединения, раскрытые в настоящем документе, могут быть включены в фармацевтические композиции в соответствии с известными способами приготовления фармацевтически приемлемых композиций. Составы подробно описаны в ряде источников, которые хорошо известны и легко доступны специалистам в данной области техники. Например, в *Remington's Pharmaceutical Science* под ред. E.W. Martin (1995) описаны составы, которые можно применять в связи с раскрытыми способами. В целом, соединения, раскрытые в настоящем документе, могут быть изготовлены таким образом, что эффективное количество соединения объединяют с подходящим носителем для облегчения эффективного введения соединения. Применяемые композиции также могут быть в различных формах. Они включают, например, твердые, полутвердые и жидкие

лекарственные формы, такие как таблетки, пилюли, порошки, жидкие растворы или суспензию, суппозитории, инъекционные и инфузионные растворы и спреи. Форма зависит от предполагаемого способа введения и терапевтического применения. Композиции также включают обычные фармацевтически приемлемые носители и разбавители, которые известны специалистам в данной области техники. Примеры носителей или разбавителей для применения с соединениями включают этанол, диметилсульфоксид, глицерин, оксид алюминия, крахмал, солевой раствор и эквивалентные носители и разбавители. Чтобы обеспечить введение таких дозировок для желаемого терапевтического лечения, композиции, раскрытые в настоящем документе, могут преимущественно содержать от примерно 0,1% до 100% по массе от общего количества одного или более рассматриваемых соединений в расчете на массу всей композиции, включая носитель или разбавитель.

Подходящие для введения составы включают, например, водные стерильные инъекционные растворы, которые могут содержать антиоксиданты, буферы, бактериостатические агенты и растворенные вещества, которые делают состав изотоническим с кровью предполагаемого реципиента; и водные и неводные стерильные суспензии, которые могут включать суспендирующие агенты и загущающие агенты. Составы могут быть представлены в однодозовых или многодозовых контейнерах, например, запечатанных ампулах и флаконах, и могут храниться в высушенном сублимацией (лиофилизированном) состоянии, требующем только наличия стерильного жидкого носителя, например, воды для инъекций, перед использованием. Растворы и суспензии для инъекций для немедленного введения могут быть приготовлены из стерильного порошка, гранул, таблеток и т.д. Следует понимать, что в дополнение к ингредиентам, конкретно упомянутым выше, композиции, раскрытые в настоящем документе, могут включать другие агенты, обычные в данной области техники, относящиеся к типу рассматриваемого состава.

Соединения, раскрытые в настоящем документе, и композиции, содержащие указанные соединения, могут быть доставлены в клетку либо посредством прямого контакта с клеткой, либо посредством носителя. Носители для доставки соединений и композиций в клетки известны в данной области техники и включают, например, инкапсулирование композиции в липосомальный фрагмент. Другие средства для доставки соединений и композиций, раскрытых в настоящем документе, в клетку включают присоединение соединений к белку или нуклеиновой кислоте, которая нацелена для доставки в клетку-мишень. В патентах США № 6960648 и публикациях заявок на патент США №№ 20030032594 и 20020120100 раскрыты аминокислотные последовательности, которые могут быть связаны с другой композицией и которые позволяют переносить композицию через биологические мембраны. В публикации заявки на патент США № 20020035243 также описаны композиции для переноса биологических фрагментов через клеточные мембраны для внутриклеточной доставки. Соединения также могут быть включены в полимеры, примеры которых включают сополимер поли-D,L-лактида и

гликолида для внутричерепных опухолей; поли[бис(п-карбоксифенокси)пропан:себаценовая кислота] в молярном соотношении 20:80 (используется в GLIADEL); хондроитин; хитин и хитозан.

Соединения и композиции, раскрытые в настоящем документе, в том числе фармацевтически приемлемые соли или пролекарства на их основе, можно вводить внутривенно, внутримышечно или внутривнутрибрюшинно путем инфузии или инъекции. Растворы активного агента или его солей могут быть приготовлены в воде, необязательно смешаны с нетоксичным поверхностно-активным веществом. Дисперсии также могут быть приготовлены в глицерине, жидких полиэтиленгликолях, триацетине или их смесях и в маслах. В обычных условиях хранения и использования эти препараты могут содержать консервант, чтобы предотвратить рост микроорганизмов.

Лекарственные формы, подходящие для инъекции или инфузии, могут включать стерильные водные растворы или дисперсии, или стерильные порошки, содержащие активный ингредиент, которые приспособлены для приготовления стерильных инъекционных или инфузионных растворов или дисперсий, необязательно инкапсулированных в липосомы, для немедленного применения. Конечная лекарственная форма должна быть стерильной, жидкой и стабильной в условиях изготовления и хранения. Жидкий носитель или среда может представлять собой растворитель или жидкую дисперсионную среду, содержащую, например, воду, этанол, полиол (например, глицерин, пропиленгликоль, жидкие полиэтиленгликоли и т. п.), растительные масла, нетоксичные глицероловые сложные эфиры и их подходящие смеси. Надлежащую текучесть можно поддерживать, например, путем образования липосом, путем поддержания требуемого размера частиц в случае дисперсий или путем применения поверхностно-активных веществ. Необязательно, предотвращение действия микроорганизмов может быть достигнуто с помощью различных других антибактериальных и противогрибковых агентов, например, парабенов, хлорбутанола, фенола, сорбиновой кислоты, тимеросала и т.п. Во многих случаях могут быть включены изотонические агенты, например, сахара, буферы или хлорид натрия. Пролонгированное всасывание инъекционных композиций может быть обеспечено включением агентов, которые замедляют всасывание, например, моностеарата алюминия и желатина.

Стерильные инъекционные растворы готовят путем включения соединения и/или агента, раскрытого в настоящем документе, в требуемом количестве в соответствующий растворитель с различными другими ингредиентами, перечисленными выше, при необходимости с последующей стерилизацией путем фильтрации. В случае стерильных порошков для приготовления стерильных инъекционных растворов способы приготовления включают вакуумную сушку и методики сублимационной сушки, которые приводят к получению порошка активного ингредиента плюс любого дополнительного желаемого ингредиента, присутствующего в растворах, предварительно стерилизованных фильтрованием.

Полезные дозировки соединений и агентов, а также фармацевтических композиций, раскрытых в настоящем документе, могут быть определены путем сравнения их активности *in vitro* и активности *in vivo* в моделях на животных. Способы экстраполяции эффективных дозировок у мышей и других животных на людей известны в данной области техники.

Диапазоны дозировок для введения композиций являются достаточно большими для обеспечения желаемого эффекта, при котором имеется влияние на симптомы или нарушение. Дозировка не должна быть настолько большой, чтобы вызвать нежелательные побочные эффекты, такие как нежелательные перекрестные реакции, анафилактические реакции и т. п. Обычно дозировка будет варьироваться в зависимости от возраста, состояния, пола и интенсивности заболевания у пациента и может быть определена специалистом в данной области техники. Дозировка может быть скорректирована врачом индивидуума в случае любых противопоказаний. Дозировка может варьироваться и может быть введена в виде одного или более введений дозы ежедневно, в течение одного или нескольких дней.

Также раскрыты фармацевтические композиции, которые содержат соединение, раскрытое в настоящем документе, в комбинации с фармацевтически приемлемым носителем. В настоящем документе раскрыты фармацевтические композиции, адаптированные для перорального, местного или парентерального введения, содержащие определенное количество соединения. Доза, вводимая пациенту, в частности, человеку, должна быть достаточной для достижения терапевтического ответа у пациента в течение приемлемого периода времени без летальной токсичности и вызывать не более чем приемлемый уровень побочных эффектов или заболеваемости. Специалисту в данной области техники будет понятно, что дозировка будет зависеть от ряда факторов, включая состояние (здоровье) субъекта, массу тела субъекта, вид сопутствующего лечения, если таковой имеется, частоту лечения, терапевтическое соотношение, а также тяжесть и стадию патологического состояния.

Также раскрыты наборы, которые содержат соединение, раскрытое в настоящем документе, и/или фармацевтические композиции, содержащие указанное соединение, в одном или более контейнерах. Раскрытые наборы могут необязательно включать фармацевтически приемлемые носители и/или разбавители. Согласно одному варианту реализации набор включает один или более других компонентов, добавок или адьювантов, описанных в настоящем документе. Согласно одному варианту реализации набор включает инструкции или упаковочные материалы, в которых описано, как вводить соединение или композицию набора. Контейнеры набора могут быть изготовлены из любого подходящего материала, например, стекла, пластика, металла и т.д., и могут быть любого подходящего размера, формы или конфигурации. Согласно одному варианту реализации соединение и/или агент, раскрытое(ый) в настоящем документе, обеспечено(обеспечен) в наборе в твердой форме, такой как таблетка, пилюля, или порошковой форме. Согласно другому варианту реализации соединение и/или агент,

раскрытое(ый) в настоящем документе, обеспечено(обеспечен) в наборе в виде жидкости или раствора. Согласно одному варианту реализации набор содержит ампулу или шприц, содержащий соединение и/или агент, раскрытое(ый) в настоящем документе, в виде жидкости или раствора.

Согласно вариантам реализации соединение и/или композицию согласно настоящему изобретению вводят пациенту, у которого диагностировано заболевание, ассоциированное с экспансией нуклеотидных повторов, в дозе от примерно 0,1 мг/кг до примерно 1000 мг/кг, например, примерно 0,1 мг/кг, примерно 0,2 мг/кг, примерно 0,3 мг/кг, примерно 0,4 мг/кг, примерно 0,5 мг/кг, примерно 0,6 мг/кг, примерно 0,7 мг/кг, примерно 0,8 мг/кг, примерно 0,9 мг/кг, примерно 1 мг/кг, примерно 2 мг/кг, примерно 3 мг/кг, примерно 4 мг/кг, примерно 5 мг/кг, примерно 6 мг/кг, примерно 7 мг/кг, примерно 8 мг/кг, примерно 9 мг/кг, примерно 10 мг/кг, примерно 11 мг/кг, примерно 12 мг/кг, примерно 13 мг/кг, примерно 14 мг/кг, примерно 15 мг/кг, примерно 16 мг/кг, примерно 17 мг/кг, примерно 18 мг/кг, примерно 19 мг/кг, примерно 20 мг/кг, примерно 21 мг/кг, примерно 22 мг/кг, примерно 23 мг/кг, примерно 24 мг/кг, примерно 25 мг/кг, примерно 26 мг/кг, примерно 27 мг/кг, примерно 28 мг/кг, примерно 29 мг/кг, примерно 30 мг/кг, примерно 31 мг/кг, примерно 32 мг/кг, примерно 33 мг/кг, примерно 34 мг/кг, примерно 35 мг/кг, примерно 36 мг/кг, примерно 37 мг/кг, примерно 38 мг/кг, примерно 39 мг/кг, примерно 40 мг/кг, примерно 41 мг/кг, примерно 42 мг/кг, примерно 43 мг/кг, примерно 44 мг/кг, примерно 45 мг/кг, примерно 46 мг/кг, примерно 47 мг/кг, примерно 48 мг/кг, примерно 49 мг/кг, примерно 50 мг/кг, примерно 51 мг/кг, примерно 52 мг/кг, примерно 53 мг/кг, примерно 54 мг/кг, примерно 55 мг/кг, примерно 56 мг/кг, примерно 57 мг/кг, примерно 58 мг/кг, примерно 59 мг/кг, примерно 60 мг/кг, примерно 61 мг/кг, примерно 62 мг/кг, примерно 63 мг/кг, примерно 64 мг/кг, примерно 65 мг/кг, примерно 66 мг/кг, примерно 67 мг/кг, примерно 68 мг/кг, примерно 69 мг/кг, примерно 70 мг/кг, примерно 71 мг/кг, примерно 72 мг/кг, примерно 73 мг/кг, примерно 74 мг/кг, примерно 75 мг/кг, примерно 76 мг/кг, примерно 77 мг/кг, примерно 78 мг/кг, примерно 79 мг/кг, примерно 80 мг/кг, примерно 81 мг/кг, примерно 82 мг/кг, примерно 83 мг/кг, примерно 84 мг/кг, примерно 85 мг/кг, примерно 86 мг/кг, примерно 87 мг/кг, примерно 88 мг/кг, примерно 89 мг/кг, примерно 90 мг/кг, примерно 91 мг/кг, примерно 92 мг/кг, примерно 93 мг/кг, примерно 94 мг/кг, примерно 95 мг/кг, примерно 96 мг/кг, примерно 97 мг/кг, примерно 98 мг/кг, примерно 99 мг/кг, примерно 100 мг/кг, примерно 110 мг/кг, примерно 120 мг/кг, примерно 130 мг/кг, примерно 140 мг/кг, примерно 150 мг/кг, примерно 160 мг/кг, примерно 170 мг/кг, примерно 180 мг/кг, примерно 190 мг/кг, примерно 200 мг/кг, примерно 210 мг/кг, примерно 220 мг/кг, примерно 230 мг/кг, примерно 240 мг/кг, примерно 250 мг/кг, примерно 260 мг/кг, примерно 270 мг/кг, примерно 280 мг/кг, примерно 290 мг/кг, примерно 300 мг/кг, примерно 310 мг/кг, примерно 320 мг/кг, примерно 330 мг/кг, примерно 340 мг/кг, примерно 350 мг/кг, примерно 360 мг/кг, примерно 370 мг/кг, примерно 380 мг/кг, примерно 390 мг/кг, примерно 400 мг/кг, примерно 410 мг/кг, примерно 420 мг/кг, примерно 430 мг/кг, примерно 440 мг/кг,

примерно 450 мг/кг, примерно 460 мг/кг, примерно 470 мг/кг, примерно 480 мг/кг, примерно 490 мг/кг, примерно 500 мг/кг, примерно 510 мг/кг, примерно 520 мг/кг, примерно 530 мг/кг, примерно 540 мг/кг, примерно 550 мг/кг, примерно 560 мг/кг, примерно 570 мг/кг, примерно 580 мг/кг, примерно 590 мг/кг, примерно 600 мг/кг, примерно 610 мг/кг, примерно 620 мг/кг, примерно 630 мг/кг, примерно 640 мг/кг, примерно 650 мг/кг, примерно 660 мг/кг, примерно 670 мг/кг, примерно 680 мг/кг, примерно 690 мг/кг, примерно 700 мг/кг, примерно 710 мг/кг, примерно 720 мг/кг, примерно 730 мг/кг, примерно 740 мг/кг, примерно 750 мг/кг, примерно 760 мг/кг, примерно 770 мг/кг, примерно 780 мг/кг, примерно 790 мг/кг, примерно 800 мг/кг, примерно 810 мг/кг, примерно 820 мг/кг, примерно 830 мг/кг, примерно 840 мг/кг, примерно 850 мг/кг, примерно 860 мг/кг, примерно 870 мг/кг, примерно 880 мг/кг, примерно 890 мг/кг, примерно 900 мг/кг, примерно 910 мг/кг, примерно 920 мг/кг, примерно 930 мг/кг, примерно 940 мг/кг, примерно 950 мг/кг, примерно 960 мг/кг, примерно 970 мг/кг, примерно 980 мг/кг, примерно 990 мг/кг или примерно 1000 мг/кг, включая все указанные значения и диапазоны между ними.

### **Способы лечения**

Согласно настоящему изобретению предложен способ лечения заболевания у субъекта, нуждающегося в этом, включающий введение соединения и/или композиции, содержащей соединение, раскрытое в настоящем документе. Согласно вариантам реализации заболевание представляет собой любое из заболеваний, приведенных в настоящем раскрытии. Согласно вариантам реализации целевой ген или генный транскрипт представляет собой любой из целевых генов или генных транскриптов, приведенных в настоящем раскрытии.

Согласно вариантам реализации пациент идентифицирован, как имеющий или подверженный риску развития любого заболевания, описанного в настоящем документе. Согласно вариантам реализации предложен способ лечения заболевания, ассоциированного с повтором CTG·CUG в 3'-нетранслируемой области гена/транскрипта. Согласно вариантам реализации предложен способ лечения миотонической дистрофии. Согласно вариантам реализации предложен способ лечения миотонической дистрофии 1 типа (DM1). Согласно вариантам реализации предложен способ лечения SCA8. Согласно вариантам реализации предложен способ лечения HDL2. Согласно вариантам реализации предложен способ лечения FECD.

Согласно вариантам реализации лечение относится к частичному или полному ослаблению, уменьшению, облегчению, ингибированию, задержке начала, снижению тяжести и/или частоты возникновения одного или более симптомов у субъекта.

Лечение заболевания и/или симптомов заболевания может происходить с помощью различных молекулярных механизмов, таких как описанные в настоящем документе.

Согласно вариантам реализации предложен способ изменения экспрессии и/или активности целевого гена у субъекта, нуждающегося в этом, включающий введение соединения, раскрытого в настоящем документе. Согласно вариантам реализации лечение

приводит к снижению экспрессии целевого белка из целевого транскрипта. Согласно вариантам реализации лечение приводит к снижению уровней целевого транскрипта. Согласно вариантам реализации лечение приводит к модуляции сплайсинга транскриптов нисходящих генов, которые регулируются целевым транскриптом и/или белками, которые связываются с целевым транскриптом. Согласно вариантам реализации модуляция сплайсинга транскриптов нисходящих генов приводит к увеличению уровня нисходящих транскриптов и/или изоформ нисходящих белков, которые ассоциированы со здоровыми фенотипами. Согласно вариантам реализации альтернативный сплайсинг приводит к уменьшению уровня нисходящих транскриптов и/или изоформ нисходящих белков, которые ассоциированы с патологическими фенотипами.

Согласно вариантам реализации предложен способ лечения DM1 путем снижения секвестрации по меньшей мере одного РНК-связывающего белка в пре-мРНК, содержащей по меньшей мере один экспансированный повтор CUG. Согласно вариантам реализации предложен способ лечения DM1 путем снижения накопления пре-мРНК, содержащей по меньшей мере один экспансированный повтор CUG. Согласно вариантам реализации предложен способ лечения DM1 путем коррекции дефектов сплайсинга транскриптов нисходящих генов.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению уровня целевого транскрипта и/или экспрессии транскрипта целевых генов (например, DMPK, TCF4, JPH3, ATXN8OS и/или ATXN8) на более чем примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем белка у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим CPP, раскрытым в настоящем документе. Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению уровня целевого транскрипта (например, DMPK, TCF4, JPH3, ATXN8OS и/или ATXN8) и/или экспрессии целевого транскрипта на от примерно 5% до примерно 100%, от примерно 10% до примерно 100%, от примерно 20% до примерно 100%, от примерно 50% до примерно 100%, от примерно 70% до примерно 100%, от примерно 80% до примерно 100%, от примерно 90% до примерно 100%, от примерно 95% до примерно 100%, от примерно 40% до примерно 95%, от примерно 50% до примерно 95%, от примерно 70% до примерно 95% или от примерно 90% до примерно 95% по сравнению со средним уровнем транскрипта и/или белка у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим CPP, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению числа ядерных очагов РНК с повторами CUG целевого гена (например, DMPK, TCF4, JPH3, ATXN8OS и/или ATXN8) на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем очагов у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим CPP, раскрытым в настоящем документе. Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению числа ядерных очагов РНК с повторами CUG целевого гена (например, DMPK, TCF4, JPH3, ATXN8OS и/или ATXN8) на от примерно 5% до примерно 100%, от примерно 10% до примерно 100%, от примерно 20% до примерно 100%, от примерно 50% до примерно 100%, от примерно 70% до примерно 100%, от примерно 80% до примерно 100%, от примерно 90% до примерно 100%, от примерно 95% до примерно 100%, от примерно 40% до примерно 95%, от примерно 50% до примерно 95%, от примерно 70% до примерно 95% или от примерно 90% до примерно 95% по сравнению со средним уровнем очагов у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим CPP, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению уровня нисходящего транскрипта и/или экспрессии нисходящего генного продукта, который ассоциирован с патологическим фенотипом, на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем белка у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим CPP, раскрытым в настоящем документе. Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению уровня нисходящего транскрипта и/или экспрессии нисходящего генного продукта, который ассоциирован с патологическим фенотипом, на от примерно 5% до примерно 100%, от примерно 10% до примерно 100%, от примерно 20% до примерно 100%, от примерно 50% до примерно 100%, от примерно 70% до примерно 100%, от примерно 80% до примерно 100%, от примерно 90% до примерно 100%, от примерно 95% до примерно 100%, от примерно 40% до примерно 95%, от примерно 50% до примерно 95%, от примерно 70% до примерно 95% или от примерно 90% до примерно 95% по сравнению со средним уровнем белка у субъекта до лечения или



у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к повышению уровня нисходящего транскрипта и/или экспрессии нисходящего генного продукта, который ассоциирован со здоровым фенотипом, на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем белка у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе. Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к повышению уровня нисходящего транскрипта и/или экспрессии нисходящего генного продукта, который ассоциирован со здоровым фенотипом, на от примерно 5% до примерно 100%, от примерно 10% до примерно 100%, от примерно 20% до примерно 100%, от примерно 50% до примерно 100%, от примерно 70% до примерно 100%, от примерно 80% до примерно 100%, от примерно 90% до примерно 100%, от примерно 95% до примерно 100%, от примерно 40% до примерно 95%, от примерно 50% до примерно 95%, от примерно 70% до примерно 95% или от примерно 90% до примерно 95% по сравнению со средним уровнем белка у субъекта до лечения или у одного или более контрольных индивидуумов с подобным заболеванием, не получавших лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению экспрессии изоформы белка, ассоциированной с патологическим фенотипом, в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце, трехглавой мышце, передней большеберцовой мышце, икроножной мышце или сердце субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем белка в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце, трехглавой мышце, передней большеберцовой мышце, икроножной мышце или сердце субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе. Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению экспрессии

изоформы белка, ассоциированной с патологическим фенотипом, в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце, трехглавой мышце, передней большеберцовой мышце, икроножной мышце или сердце субъекта на более чем от примерно 5% до примерно 100%, от примерно 10% до примерно 100%, от примерно 20% до примерно 100%, от примерно 50% до примерно 100%, от примерно 70% до примерно 100%, от примерно 80% до примерно 100%, от примерно 90% до примерно 100%, от примерно 95% до примерно 100%, от примерно 40% до примерно 95%, от примерно 50% до примерно 95%, от примерно 70% до примерно 95% или от примерно 90% до примерно 95% по сравнению со средним уровнем белка в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце, трехглавой мышце, передней большеберцовой мышце, икроножной мышце или сердце субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к увеличению экспрессии альтернативно сплайсированного нисходящего белка в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце или сердце субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95%, примерно 100%, примерно 150%, примерно 200%, примерно 250%, примерно 300%, примерно 350%, примерно 400%, примерно 450%, примерно 500%, примерно 550%, примерно 600%, примерно 650%, примерно 700%, примерно 750%, примерно 800, примерно 850%, примерно 900%, примерно 950% или примерно 1000% или более по сравнению со средним уровнем нисходящего белка в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце или сердце субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к увеличению или уменьшению экспрессии изомера белка дикого типа в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце или сердце субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем изомера белка дикого типа в ткани роговицы, мышечной ткани, ткани диафрагмы, четырехглавой мышце или сердце субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с

подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к уменьшению экспрессии белка в представляющей интерес ткани субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем белка в представляющей интерес ткани субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к увеличению экспрессии альтернативно сплайсированного нисходящего белка в представляющей интерес ткани субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95%, примерно 100%, примерно 150%, примерно 200%, примерно 250%, примерно 300%, примерно 350%, примерно 400%, примерно 450%, примерно 500%, примерно 550%, примерно 600%, примерно 650%, примерно 700%, примерно 750%, примерно 800, примерно 850%, примерно 900%, примерно 950% или примерно 1000% или более по сравнению со средним уровнем нисходящего белка в представляющей интерес ткани субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации лечение в соответствии с настоящим раскрытием приводит к увеличению или уменьшению экспрессии изомера нисходящего белка дикого типа в представляющей интерес ткани субъекта на более чем примерно 5%, например, примерно 5%, примерно 10%, примерно 15%, примерно 20%, примерно 25%, примерно 30%, примерно 35%, примерно 40%, примерно 45%, примерно 50%, примерно 55%, примерно 60%, примерно 65%, примерно 70%, примерно 75%, примерно 80%, примерно 85%, примерно 90%, примерно 95% и примерно 100% по сравнению со средним уровнем нисходящего белка в представляющей интерес ткани субъекта до лечения, по сравнению с одним или более контрольными индивидуумами с подобным заболеванием, не получавшими лечения, или по сравнению с лечением АС, не конъюгированным с циклическим СРР, раскрытым в настоящем документе.

Согласно вариантам реализации представляющая интерес ткань субъекта представляет собой ткань роговицы или мышечную ткань.

Термины «улучшать», «повышать», «снижать», «уменьшать» и т. п. в контексте настоящего документа указывают на значения относительно контроля. Согласно вариантам реализации подходящим контролем является измерение в начальных условиях, такое как измерение у одного и того же индивидуума до начала лечения, описанного в настоящем документе, или измерение у контрольного индивидуума (или нескольких контрольных индивидуумов) в отсутствие лечения, описанного в настоящем документе. «Контрольный индивидуум» представляет собой индивидуума, пораженного тем же заболеванием, который имеет примерно такой же возраст и/или пол, что и индивидуум, которого лечат (чтобы обеспечить сопоставимость стадий заболевания у получавшего лечение индивидуума и контрольного индивидуума(ов)).

Индивидуум (также называемый «пациентом» или «субъектом»), которого лечат, представляет собой индивидуума (плод, младенца, ребенка, подростка или взрослого человека), страдающего заболеванием или у которого заболевание может развиваться. Индивидуум может страдать заболеванием, опосредуемым aberrантной экспрессией гена или aberrантным сплайсингом гена. Согласно различным вариантам реализации индивидуум, страдающий заболеванием, может иметь уровни экспрессии или активности нисходящего белка, которые составляют менее примерно 1-99% от нормальных уровней экспрессии или активности белка дикого типа у индивидуума, не пораженного этим заболеванием. Согласно вариантам реализации диапазон включает, но не ограничивается этим, менее примерно 80-99%, менее примерно 65-80%, менее примерно 50-65%, менее примерно 30-50%, менее примерно 25-30%, менее примерно 20-25%, менее примерно 15-20%, менее примерно 10-15%, менее примерно 5-10%, менее примерно 1-5% от нормального уровня экспрессии или активности белка дикого типа. Согласно вариантам реализации индивидуум может иметь уровни экспрессии или активности нисходящего белка, которые на 1-500% выше, чем нормальные уровни экспрессии или активности целевого белка дикого типа у индивидуума, не пораженного заболеванием. Согласно вариантам реализации диапазон включает, но не ограничивается этим, более примерно 1-10%, примерно 10-50%, примерно 50-100%, примерно 100-200%, примерно 200-300%, примерно 300-400%, примерно 400-500% или примерно 500-1000% от нормального уровня экспрессии или активности целевого белка дикого типа.

Согласно вариантам реализации индивидуум представляет собой индивидуума, у которого недавно было диагностировано заболевание. Как правило, раннее лечение (лечение, начинающееся как можно скорее после постановки диагноза) важно для минимизации эффектов заболевания и максимизации пользы лечения.

### **Некоторые определения**

В настоящем описании и прилагаемой формуле изобретения термины в единственном числе (соотв. «a», «an» и «the» в исходном тексте на английском языке) включают их эквиваленты во множественном числе, если иное явным образом не следует из контекста. Таким образом, например, упоминание «композиции» включает смеси двух или более таких композиций, упоминание «агента» включает смеси двух или более таких

агентов, упоминание «компонента» включает смеси двух или более таких компонентов и т. п.

Термин «примерно», когда он непосредственно предшествует численному значению, означает диапазон (например, плюс или минус 20%, 10% или 5% от этого значения). Например, «примерно 50» может означать от 45 до 55, «примерно 25000» может означать от 22500 до 27500 и т. д., если контекст настоящего изобретения не указывает иное или не противоречит такой интерпретации. Например, в перечне численных значений, таких как «примерно 49, примерно 50, примерно 55, ...», «примерно 50» означает диапазон, проходящий до менее чем половины интервала (интервалов) между предыдущим и последующим значениями, например, от более чем 49,5 до менее чем 52,5. Кроме того, выражения «менее чем примерно» значения или «более чем примерно» значения следует понимать с учетом определения термина «примерно», представленного в настоящем документе. Аналогичным образом, термин «примерно», когда он предшествует ряду численных значений или диапазону значений (например, «примерно 10, 20, 30» или «примерно 10-30»), относится, соответственно, ко всем значениям в ряду или конечным точкам диапазона.

В настоящем документе термин «проникающий в клетку пептид» или «СРР» относится к пептиду, который облегчает доставку груза, например, терапевтического фрагмента (ТМ), в клетку. Согласно вариантам реализации СРР является циклическим и представлен как «сСРР». Согласно вариантам реализации сСРР способен направлять терапевтический фрагмент для проникновения через мембрану клетки. Согласно вариантам реализации с помощью сСРР доставляют терапевтический фрагмент в цитозоль клетки. Согласно вариантам реализации с помощью сСРР доставляют антисмысловое соединение (АС) в клеточное местоположение, где расположена пре-мРНК.

В настоящем документе термин «носитель для ускользания из эндосом» (EEV) относится к сСРР, который конъюгирован с помощью химической связи (т. е. ковалентной связи или нековалентного взаимодействия) с линкером и/или экзоциклическим пептидом (EP). EEV может представлять собой EEV Формулы (B).

В настоящем документе термин «конъюгат EEP» относится к носителю для ускользания из эндосом, определенному в настоящем документе, конъюгированному с помощью химической связи (т. е. ковалентной связи или нековалентного взаимодействия) с грузом. Груз может представлять собой терапевтический фрагмент (например, олигонуклеотид, пептид или малую молекулу), который может быть доставлен в клетку с помощью EEV. Конъюгат EEV может представлять собой конъюгат EEV Формулы (C).

В настоящем документе термины «экзоциклический пептид» (EP) и «модулирующий пептид» (MP) могут использоваться взаимозаменяемо для обозначения двух или более аминокислотных остатков, связанных пептидной связью, которые могут быть конъюгированы с проникающим в клетку циклическим пептидом (сСРР), раскрытым в настоящем документе. EP при конъюгировании с циклическим пептидом, раскрытым в настоящем документе, может изменять распределение и/или удержание соединения в

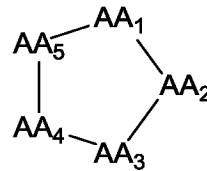
тканях. Как правило, EP содержит по меньшей мере один положительно заряженный аминокислотный остаток, например, по меньшей мере один остаток лизина и/или по меньшей мере один остаток аргинина. В настоящем документе описаны неограничивающие примеры EP. EP может представлять собой пептид, который был идентифицирован в данной области техники как «последовательность ядерной локализации» (NLS). Неограничивающие примеры последовательностей ядерной локализации включают последовательность ядерной локализации большого Т-антигена вируса SV40, минимальным функциональным звеном которой является последовательность из семи аминокислот PKKKRKV (SEQ ID NO:42), двусоставную NLS нуклеоплазмينا с последовательностью NLSKRPAAIKKAGQAKKKK (SEQ ID NO:52), с-тус последовательность ядерной локализации, имеющую аминокислотную последовательность PAAKRVKLD (SEQ ID NO:53) или RQRRNELKRSF (SEQ ID NO:54), последовательность RMRKFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO:50) домена IBV из импортина-альфа, последовательности VSRKRPRP (SEQ ID NO:57) и PPKKARED (SEQ ID NO:58) белка Т миомы, последовательность PPKKKPL (SEQ ID NO:59) человеческого p53, последовательность SALIKKKKKMAP (SEQ ID NO:60) мышинового c-abl IV, последовательности DRLRR (SEQ ID NO:61) и PKQKKRK (SEQ ID NO:62) NS1 вируса гриппа, последовательность RKLKKIKKL (SEQ ID NO:63) дельта-антигена вируса гепатита, последовательность REKKKFLKRR (SEQ ID NO:64) мышинового белка Mx1, последовательность KRKGDEV DGVDEVAKKKSKK (SEQ ID NO:65) человеческой поли(АДФ-рибоза)-полимеразы и последовательность RKCLQAGMNLEARKTKK (SEQ ID NO:66) рецепторов глюкокортикоидных стероидных гормонов (человеческих). В международной публикации № 2001/038547 описаны дополнительные примеры NLS, и она полностью включена в настоящий документ посредством ссылки.

В настоящем документе термин «линкер» или «L» относится к фрагменту, который ковалентно связывает один или более фрагментов (например, экзоциклический пептид (EP) и груз, например, олигонуклеотид, пептид или малую молекулу) с проникающим в клетку циклическим пептидом (сСРР). Линкер может содержать природную или неприродную аминокислоту или полипептид. Линкер может представлять собой синтетическое соединение, содержащее две или более подходящих функциональных групп, подходящих для связывания сСРР с фрагментом груза с образованием тем самым соединений, раскрытых в настоящем документе. Линкер может содержать фрагмент полиэтиленгликоля (PEG). Линкер может содержать одну или более аминокислот. сСРР может быть ковалентно связан с грузом посредством линкера.

Термины «пептид», «белок» и «полипептид» используются взаимозаменяемо для обозначения природной или синтетической молекулы, содержащей две или более аминокислот, связанных карбоксильной группой одной аминокислоты с альфа-аминогруппой другой. Два или более аминокислотных остатка могут быть связаны карбоксильной группой одной аминокислоты с альфа-аминогруппой. Две или более

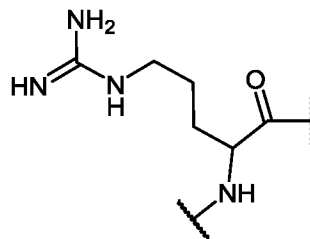
аминокислоты полипептида могут быть соединены пептидной связью. Полипептид может включать модификацию пептидного остова, в которой две или более аминокислот ковалентно присоединены связью, отличной от пептидной связи. Полипептид может включать одну или более не природных аминокислот, аналогов аминокислот или других синтетических молекул, которые способны интегрироваться в полипептид. Термин «полипептид» включает встречающиеся в природе и искусственные аминокислоты. Термин «полипептид» включает, например, пептиды, которые включают от примерно 2 до примерно 100 аминокислотных остатков, а также белки, которые включают более примерно 100 аминокислотных остатков или более примерно 1000 аминокислотных остатков, включая, но не ограничиваясь этим, терапевтические белки, такие как антитела, ферменты, рецепторы, растворимые белки и т.п.

В настоящем документе термин «смежный» относится к двум аминокислотам, которые соединены ковалентной связью. Например, в контексте типичного проникающего



в клетку циклического пептида (сСРР), такого как  $AA_1/AA_2$ ,  $AA_2/AA_3$ ,  $AA_3/AA_4$ , и  $AA_5/AA_1$  представляют собой примерные пары смежных аминокислот.

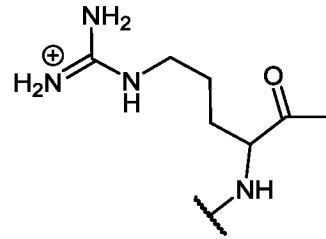
Остаток химического соединения, в контексте настоящего документа, относится к производному химического соединения, которое присутствует в конкретном продукте. Для образования продукта по меньшей мере один атом соединения заменяют связью с другим фрагментом так, что продукт содержит производное, или остаток, химического соединения. Например, проникающие в клетку циклические пептиды (сСРР), описанные в настоящем документе, содержат аминокислоты (например, аргинин), включенные в них посредством образования одной или более пептидных связей. Аминокислоты, включенные в сСРР, можно называть остатками или просто аминокислотами. Например,



аргинин или остаток аргинина относится к

Термин «ее протонированная форма» относится к протонированной форме аминокислоты или боковой цепи. Например, гуанидиновая группа на боковой цепи

аргинина может быть протонирована с образованием группы гуанидиния. Структура



протонированной формы аргинина представляет собой

В настоящем документе термин «хиральность» относится к молекуле, которая имеет более одного стереоизомера, который отличается трехмерным пространственным расположением атомов, при этом один стереоизомер представляет собой неналагаемое зеркальное изображение другого. Аминокислоты, за исключением глицина, содержат хиральный атом углерода, смежный с карбоксильной группой. Термин «энантиомер» относится к стереоизомерам, которые являются хиральными. Хиральная молекула может представлять собой аминокислотный остаток, имеющий энантиомеры «D» и «L». Молекулы без хирального центра, такие как глицин, могут называться «ахиральными».

В настоящем документе термин «гидрофобный» относится к фрагменту, который является нерастворимым в воде или имеет минимальную растворимость в воде. Обычно нейтральные фрагменты и/или неполярные фрагменты или фрагменты, которые являются преимущественно нейтральными и/или неполярными, являются гидрофобными. Гидрофобность может быть измерена одним из способов, раскрытых в настоящем документе.

В настоящем документе «ароматический» относится к ненасыщенной циклической молекуле, содержащей  $4n+2$   $\pi$  электронов, где  $n$  представляет собой любое целое число. «Гетероароматический», определенный ниже, представляет собой подгруппу ароматических молекул. Примеры ароматических аминокислот включают фенилаланин и нафтилаланин. Термин «неароматический» относится к любой молекуле, которая не подпадает под определение ароматической. Например, любая линейная, разветвленная или циклическая молекула, которая не подпадает под определение ароматической, является неароматической. Примеры неароматических аминокислот включают, но не ограничиваются перечисленными, глицин и цитруллин.

«Алкил», «алкильная цепь» или «алкильная группа» относятся к полностью насыщенному углеводородному радикалу с прямой или разветвленной цепью, содержащему от одного до сорока атомов углерода и который присоединен к остальной части молекулы одной связью. Включены алкилы, содержащие любое число атомов углерода от 1 до 40. Алкил, содержащий до 40 атомов углерода, представляет собой C<sub>1</sub>-C<sub>40</sub> алкил, алкил, содержащий до 10 атомов углерода, представляет собой C<sub>1</sub>-C<sub>10</sub> алкил, алкил, содержащий до 6 атомов углерода, представляет собой C<sub>1</sub>-C<sub>6</sub> алкил, и алкил, содержащий до 5 атомов углерода, представляет собой C<sub>1</sub>-C<sub>5</sub> алкил. C<sub>1</sub>-C<sub>5</sub> алкил включает C<sub>5</sub> алкилы, C<sub>4</sub> алкилы, C<sub>3</sub> алкилы, C<sub>2</sub> алкилы и C<sub>1</sub> алкил (т. е. метил). C<sub>1</sub>-C<sub>6</sub> алкил включает все фрагменты, описанные выше для C<sub>1</sub>-C<sub>5</sub> алкилов, но также включает C<sub>6</sub> алкилы. C<sub>1</sub>-C<sub>10</sub>



алкил включает все фрагменты, описанные выше для  $C_1$ - $C_5$  алкилов и  $C_1$ - $C_6$  алкилов, но также включает  $C_7$ ,  $C_8$ ,  $C_9$  и  $C_{10}$  алкилы. Аналогичным образом,  $C_1$ - $C_{12}$  алкил включает все вышеуказанные фрагменты, но также включает  $C_{11}$  и  $C_{12}$  алкилы. Неограничивающие примеры  $C_1$ - $C_{12}$  алкила включают метил, этил, *n*-пропил, изопропил, *втор*-пропил, *n*-бутил, изобутил, *втор*-бутил, *трет*-бутил, *n*-пентил, *трет*-амил, *n*-гексил, *n*-гептил, *n*-октил, *n*-нонил, *n*-децил, *n*-ундецил и *n*-додецил. Если в описании конкретно не указано иное, алкильная группа может быть необязательно замещенной.

«Алкилен», «алкиленовая цепь» или «алкиленовая группа» относятся к полностью насыщенному двухвалентному углеводородному радикалу с прямой или разветвленной цепью, содержащему от одного до сорок атомов углерода. Неограничивающие примеры  $C_2$ - $C_{40}$  алкилена включают этилен, пропилен, *n*-бутилен, этенилен, пропенилен, *n*-бутенилен, пропинилен, *n*-бутинилен и т. п. Если в описании конкретно не указано иное, алкиленовая цепь может быть необязательно замещенной.

«Алкенил», «алкенильная цепь» или «алкенильная группа» относится к углеводородному радикалу с прямой или разветвленной цепью, содержащему от двух до сорок атомов углерода и содержащему одну или более углерод-углеродных двойных связей. Каждая алкенильная группа присоединена к остальной части молекулы одной связью. Включены алкенильные группы, содержащие любое число атомов углерода от 2 до 40. Алкенильная группа, содержащая до 40 атомов углерода, представляет собой  $C_2$ - $C_{40}$  алкенил, алкенил, содержащий до 10 атомов углерода, представляет собой  $C_2$ - $C_{10}$  алкенил, алкенильная группа, содержащая до 6 атомов углерода, представляет собой  $C_2$ - $C_6$  алкенил, и алкенил, содержащий до 5 атомов углерода, представляет собой  $C_2$ - $C_5$  алкенил.  $C_2$ - $C_5$  алкенил включает  $C_5$  алкенилы,  $C_4$  алкенилы,  $C_3$  алкенилы и  $C_2$  алкенилы.  $C_2$ - $C_6$  алкенил включает все фрагменты, описанные выше для  $C_2$ - $C_5$  алкенилов, но также включает  $C_6$  алкенилы.  $C_2$ - $C_{10}$  алкенил включает все фрагменты, описанные выше для  $C_2$ - $C_5$  алкенилов и  $C_2$ - $C_6$  алкенилов, но также включает  $C_7$ ,  $C_8$ ,  $C_9$  и  $C_{10}$  алкенилы. Аналогичным образом,  $C_2$ - $C_{12}$  алкенил включает все вышеуказанные фрагменты, но также включает  $C_{11}$  и  $C_{12}$  алкенилы. Неограничивающие примеры  $C_2$ - $C_{12}$  алкенила включают этенил (винил), 1-пропенил, 2-пропенил (аллил), изопропенил, 2-метил-1-пропенил, 1-бутенил, 2-бутенил, 3-бутенил, 1-пентенил, 2-пентенил, 3-пентенил, 4-пентенил, 1-гексенил, 2-гексенил, 3-гексенил, 4-гексенил, 5-гексенил, 1-гептенил, 2-гептенил, 3-гептенил, 4-гептенил, 5-гептенил, 6-гептенил, 1-октенил, 2-октенил, 3-октенил, 4-октенил, 5-октенил, 6-октенил, 7-октенил, 1-ноненил, 2-ноненил, 3-ноненил, 4-ноненил, 5-ноненил, 6-ноненил, 7-ноненил, 8-ноненил, 1-деценил, 2-деценил, 3-деценил, 4-деценил, 5-деценил, 6-деценил, 7-деценил, 8-деценил, 9-деценил, 1-ундеценил, 2-ундеценил, 3-ундеценил, 4-ундеценил, 5-ундеценил, 6-ундеценил, 7-ундеценил, 8-ундеценил, 9-ундеценил, 10-ундеценил, 1-додеценил, 2-додеценил, 3-додеценил, 4-додеценил, 5-додеценил, 6-додеценил, 7-додеценил, 8-додеценил, 9-додеценил, 10-додеценил и 11-додеценил. Если в описании конкретно не указано иное, алкильная группа может быть необязательно замещенной.

«Алкенилен», «алкениленовая цепь» или «алкениленовая группа» относится к двухвалентному углеводородному радикалу с прямой или разветвленной цепью, содержащему от двух до сорока атомов углерода и содержащему одну или более углерод-углеродных двойных связей. Неограничивающие примеры  $C_2$ - $C_{40}$  алкенилена включают этен, пропен, бутен и т. п. Если в описании конкретно не указано иное, алкениленовая цепь может быть необязательно.

«Алкокси» или «алкоксигруппа» относится к группе  $-OR$ , где  $R$  представляет собой алкил, алкенил, алкинил, циклоалкил или гетероциклил, определенные в настоящем документе. Если в описании конкретно не указано иное, алкоксигруппа может быть необязательно замещенной.

«Ацил» или «ацильная группа» относится к группам  $-C(O)R$ , где  $R$  представляет собой водород, алкил, алкенил, алкинил, карбоциклил или гетероциклил, определенные в настоящем документе. Если в описании конкретно не указано иное, ацил может быть необязательно замещенным.

«Алкилкарбамоил» или «алкилкарбамоильная группа» относится к группе  $-O-C(O)-NR_aR_b$ , где  $R_a$  и  $R_b$  являются одинаковыми или разными и независимо представляют собой алкил, алкенил, алкинил, арил, гетероарил, определенные в настоящем документе, или  $R_aR_b$  могут быть взяты вместе с образованием циклоалкильной группы или гетероциклильной группы, определенной в настоящем документе. Если в описании конкретно не указано иное, алкилкарбамоильная группа может быть необязательно замещенной.


«Алкилкарбоксамидил» или «алкилкарбоксамидильная группа» относится к группе  $-C(O)-NR_aR_b$ , где  $R_a$  и  $R_b$  являются одинаковыми или разными и независимо представляют собой алкильную, алкенильную, алкинильную, арильную, гетероарильную, циклоалкильную, циклоалкенильную, циклоалкинильную или гетероциклильную группу, определенную в настоящем документе, или  $R_aR_b$  могут быть взяты вместе с образованием циклоалкильной группы, определенной в настоящем документе. Если в описании конкретно не указано иное, алкилкарбоксамидильная группа может быть необязательно замещенной.

«Арил» относится к углеводородному кольцевому радикалу, содержащему водород, от 6 до 18 атомов углерода и по меньшей мере одно ароматическое кольцо. Для целей настоящего изобретения арильный радикал может представлять собой моноциклическую, бициклическую, трициклическую или тетрациклическую кольцевую систему, которая может включать конденсированные или мостиковые кольцевые системы. Арильные радикалы включают, но не ограничиваются перечисленными, арильные радикалы, полученные из ацеантрилена, аценафтилена, ацефенантрилена, антрацена, азулена, бензола, хризена, флуорантена, флуорена, *as*-индацена, *s*-индацена, индана, индена, нафталина, феналена, фенантрена, плеядена, пирена и трифенилена. Если в описании конкретно не указано иное, термин «арил» подразумевает включение арильных радикалов, которые необязательно замещены.

«Гетероарил» относится к 5-20-членному кольцевому радикалу, содержащему атомы водорода, от одного до тринадцати атомов углерода, от одного до шести гетероатомов, выбранных из азота, кислорода и серы, и по меньшей мере одно ароматическое кольцо. Для целей настоящего изобретения гетероарильный радикал может представлять собой моноциклическую, бициклическую, трициклическую или тетрациклическую кольцевую систему, которая может включать конденсированные или мостиковые кольцевые системы; и атомы азота, углерода или серы в гетероарильном радикале могут быть необязательно окислены; атом азота может быть необязательно кватернизован. Примеры включают, но не ограничиваются перечисленными, азепинил, акридинил, бензимидазолил, бензотиазолил, бензиндолил, бензодиоксолил, бензофуранил, бензооксазолил, бензотиазолил, бензотиадиазолил, бензо[*b*][1,4]диоксепинил, 1,4-бензодиоксанил, бензонафтофуранил, бензоксазолил, бензодиоксолил, бензодиоксинил, бензопиранил, бензопиранонил, бензофуранил, бензофуранонил, бензотиенил (бензотиофенил), бензотриазолил, бензо[4,6]имидазо[1,2-а]пиридинил, карбазолил, циннолинил, дибензофуранил, дибензотиофенил, фуранил, фуранонил, изотиазолил, имидазолил, индазолил, индолил, индазолил, изоиндолил, индолинил, изоиндолинил, изохинолил, индолизинил, изоксазолил, нафтиридинил, оксадиазолил, 2-оксазепинил, оксазолил, оксиранил, 1-оксидопиридинил, 1-оксидопиримидинил, 1-оксидопиразинил, 1-оксидопиридазинил, 1-фенил-1*H*-пирролил, феназинил, фенотиазинил, феноксазинил, фталазинил, птеридинил, пуринил, пирролил, пиразолил, пиридинил, пиразинил, пиримидинил, пиридазинил, хиназолинил, хиноксалинил, хинолинил, хинуклидинил, изохинолинил, тетрагидрохинолинил, тиазолил, тиадиазолил, триазолил, тетразолил, триазинил и тиофенил (т.е. тиенил). Если в описании конкретно не указано иное, гетероарильная группа может быть необязательно замещенной.

Термин «замещенный», используемый в настоящем документе, означает любую из вышеуказанных групп (т.е. алкил, алкенил, алкинил, циклоалкил, циклоалкенил, циклоалкинил, гетероцикл, арил, гетероарил, алкокси, арилокси, ацил, алкилкарбамоил, алкилкарбоксамидил, алкоксикарбонил, алкилтио или арилтио), в которой по меньшей мере один атом заменен неводородными атомами, такими как, но не ограничиваясь перечисленными: атом галогена, такой как F, Cl, Br и I; атом кислорода в таких группах как гидроксильные группы, алкоксигруппы и сложноэфирные группы; атом серы в таких группах как тиоловые группы, тиоалкильные группы, сульфоновые группы, сульфонильные группы и сульфоксидные группы; атом азота в таких группах как амины, амиды, алкиламины, диалкиламины, ариламины, алкилариламины, диариламины, N-оксиды, имиды и енамины; атом кремния в таких группах как триалкилсилильные группы, диалкиларилсилильные группы, алкилдиарилсилильные группы и триарилсилильные группы; и другие гетероатомы в различных других группах. «Замещенный» также означает любую из вышеуказанных групп, в которых один или более атомов заменены связью более высокого порядка (например, двойной или тройной связью) с гетероатомом, таким как кислород в оксо, карбонильных, карбоксильных и сложноэфирных группах; и

азот в таких группах как имины, оксимы, гидразоны и нитрилы. Например, «замещенный» включает любую из вышеуказанных групп, в которых один или более атомов заменены  $-NR_gR_h$ ,  $-NR_gC(=O)R_h$ ,  $-NR_gC(=O)NR_gR_h$ ,  $-NR_gC(=O)OR_h$ ,  $-NR_gSO_2R_h$ ,  $-OC(=O)NR_gR_h$ ,  $-OR_g$ ,  $-SR_g$ ,  $-SOR_g$ ,  $-SO_2R_g$ ,  $-OSO_2R_g$ ,  $-SO_2OR_g$ ,  $=NSO_2R_g$  и  $-SO_2NR_gR_h$ . «Замещенный» также означает любую из вышеуказанных групп, в которых один или более атомов водорода заменены  $-C(=O)R_g$ ,  $-C(=O)OR_g$ ,  $-C(=O)NR_gR_h$ ,  $-CH_2SO_2R_g$ ,  $-CH_2SO_2NR_gR_h$ . В вышеизложенном  $R_g$  и  $R_h$  являются одинаковыми или разными и независимо представляют собой водород, алкил, алкенил, алкинил, алкокси, алкиламино, тиаалкил, арил, аралкил, циклоалкил, циклоалкенил, циклоалкинил, циклоалкилалкил, галогеналкил, галогеналкенил, галогеналкинил, гетероциклил, *N*-гетероциклил, гетероциклилалкил, гетероарил, *N*-гетероарил и/или гетероарилалкил. «Замещенный» дополнительно означает любую из вышеуказанных групп, в которых один или более атомов заменены амино, циано, гидроксид, имино, нитро, оксо, тиоксо, галоген, алкил, алкенил, алкинил, алкокси, алкиламино, тиаалкил, арил, аралкил, циклоалкил, циклоалкенил, циклоалкинил, циклоалкилалкил, галогеналкил, галогеналкенил, галогеналкинил, гетероциклил, *N*-гетероциклил, гетероциклилалкил, гетероарил, *N*-гетероарил и/или гетероарилалкильной группой. «Замещенный» также может означать аминокислоту, в которой один или более атомов на боковой цепи заменены алкилом, алкенилом, алкинилом, ацилом, алкилкарбоксамидилом, алкоксикарбонилом, карбоциклилом, гетероциклилом, арилом или гетероарилом. Кроме того, каждый из вышеуказанных заместителей также может быть необязательно замещен одним или более из вышеуказанных заместителей.

В настоящем документе символ «  » (далее может называться «связь в положении присоединения») обозначает связь, которая представляет собой положение присоединения между двумя химическими частицами, одна из которых изображена как присоединенная к связи в положении присоединения, а другая не изображена как присоединенная к связи в положении присоединения. Например, «  $XY \text{---} \text{---}$  » указывает на то, что химическая частица «XY» связана с другой химической частицей посредством связи в положении присоединения. Кроме того, конкретное положение присоединения к не изображенной химической частице может быть определено посредством логического вывода. Например, соединение  $CH_3-R^3$ , где  $R^3$  представляет собой H или «  $XY \text{---} \text{---}$  », логически означает, что если  $R^3$  представляет собой «XY», то связь в положении присоединения представляет собой ту же связь, что и связь, посредством которой  $R^3$ , как изображено, связан с  $CH_3$ .

В настоящем документе «субъект» представляет собой индивидуума. Таким образом, «субъект» может включать домашних животных (например, кошек, собак и т. д.),

сельскохозяйственных животных (например, крупный рогатый скот, лошадей, свиней, овец, коз и т. д.), лабораторных животных (например, мышь, кролика, крысу, морскую свинку и т. д.) и птиц. «Субъект» также может включать млекопитающее, такое как примат или человек. Таким образом, субъект может представлять собой пациента-человека или ветеринарного пациента. Термин «пациент» относится к субъекту, проходящему лечение у медицинского работника, например, врача.

Термины «ингибировать», «осуществлять ингибирование» или «ингибирование» относятся к уменьшению активности, экспрессии, функции или другого биологического параметра и могут включать, но не требуют, полного устранения активности, экспрессии, функции или другого биологического параметра. Ингибирование может включать, например, по меньшей мере примерно 10% снижение активности, ответа, состояния или заболевания по сравнению с контролем. Согласно вариантам реализации экспрессия, активность или функция гена или белка уменьшается на статистически значимую величину. Согласно вариантам реализации активность или функция уменьшается на по меньшей мере примерно 10%, примерно 20%, примерно 30%, примерно 40%, примерно 50% и до примерно 60%, примерно 70%, примерно 80%, примерно 90% или примерно 100%.

Под «снижать» или другими формами слова, такими как «осуществлять снижение» или «снижение», подразумевается уменьшение события или характеристики (например, роста опухоли). Понятно, что это, как правило, относится к некоторому стандартному или ожидаемому значению, другими словами, оно является относительным, но не всегда необходимо ссылаться на стандартное или относительное значение. Например, «снижает рост опухоли» означает снижение скорости роста опухоли по сравнению со стандартом или контролем (например, необработанной опухолью).

В настоящем документе «лечить», «осуществлять лечение», «лечение» и их варианты относятся к любому введению раскрытых соединений, которые частично или полностью ослабляют, уменьшают, облегчают, ингибируют, отсрочивают начало, снижают тяжесть и/или снижают частоту возникновения одного или более симптомов или признаков заболевания, описанных в настоящем документе. В отношении пациента термин «лечение» относится к медицинскому ведению пациента с целью излечения, уменьшения, стабилизации или предотвращения заболевания, патологического состояния или нарушения. Данный термин включает активное лечение, то есть лечение, направленное конкретно на улучшение заболевания, патологического состояния или нарушения, а также включает причинно-следственное лечение, то есть лечение, направленное на устранение причины ассоциированного заболевания, патологического состояния или нарушения. Кроме того, этот термин включает паллиативное лечение, то есть лечение, направленное на облегчение симптомов, а не на излечение заболевания, патологического состояния или нарушения; профилактическое лечение, то есть лечение, направленное на минимизацию или частичное или полное ингибирование развития ассоциированного заболевания, патологического состояния или нарушения; и

поддерживающее лечение, то есть лечение, применяемое в дополнение к другой специфической терапии, направленной на улучшение ассоциированного заболевания, патологического состояния или нарушения.

Термин «терапевтически эффективный» относится к применяемому количеству раскрытого соединения и/или композиции, достаточному для уменьшения одной или более причин или симптомов заболевания или нарушения. Такое уменьшение требует только снижения или изменения, но не обязательно устранения.

Термин «фармацевтически приемлемый» относится к соединениям, материалам, композициям и/или лекарственным формам, которые в рамках здравого медицинского суждения подходят для применения в контакте с тканями человека и/или животных без избыточной токсичности, раздражения, аллергического ответа или других проблем или осложнений, в соответствии с разумным соотношением польза/риск.

Термин «носитель» означает соединение, композицию, вещество или структуру, которые в комбинации с соединением или композицией согласно настоящему изобретению способствуют или облегчают получение, хранение, введение, доставку, эффективность, селективность или любой другой признак соединения или композиции для его(ее) предполагаемого применения или цели, или их комбинации. Например, носитель может быть выбран для минимизации любого разрушения активного ингредиента и для минимизации любых нежелательных побочных эффектов у субъекта.

В настоящем документе термин «фармацевтически приемлемый носитель» относится к носителю, подходящему для введения пациенту. Фармацевтический носитель может представлять собой вещество, которое способствует или облегчает получение, хранение, введение, доставку, эффективность, селективность или любой другой признак соединения или композиции согласно настоящему изобретению для его(ее) предполагаемого применения или цели, или их комбинации. Например, носитель может быть выбран для снижения разрушения соединения или для снижения нежелательных побочных эффектов у пациента. Согласно вариантам реализации фармацевтически приемлемый носитель может представлять собой стерильные водные или неводные растворы, дисперсии, суспензии или эмульсии, а также стерильные порошки для восстановления в стерильные инъекционные растворы или дисперсии непосредственно перед применением. Примеры подходящих водных и неводных носителей, разбавителей, растворителей или сред включают воду, этанол, полиолы (такие как глицерин, пропиленгликоль, полиэтиленгликоль и т. п.), карбоксиметилцеллюлозу и их подходящие смеси, растительные масла (такие как оливковое масло) и инъекционные органические сложные эфиры, такие как этилолеат. Подходящую текучесть можно поддерживать, например, за счет использования материалов покрытия, таких как лецитин, путем поддержания необходимого размера частиц в случае дисперсий и за счет использования поверхностно-активных веществ. Такие композиции также могут содержать адъюванты, такие как консерванты, смачивающие агенты, эмульгирующие агенты и диспергирующие агенты. Предотвращение действия микроорганизмов может быть обеспечено путем

включения различных антибактериальных и противогрибковых агентов, таких как парабен, хлорбутанол, фенол, сорбиновая кислота и т. п. Также может быть желательным включать изотонические агенты, такие как сахара, хлорид натрия и т. п. Инъекционные составы можно стерилизовать, например, путем фильтрации через фильтр, удерживающий бактерии, или путем включения стерилизующих агентов в виде стерильных твердых композиций, которые можно растворять или диспергировать в стерильной воде или другой стерильной среде для инъекций непосредственно перед применением. Подходящие инертные носители могут включать сахара, такие как лактоза.

Термин «фармацевтически приемлемые соли» включает соли, полученные путем взаимодействия активного соединения, функционирующего в качестве основания, с неорганической или органической кислотой с образованием соли, например, солей хлористоводородной кислоты, серной кислоты, фосфорной кислоты, метансульфоновой кислоты, камфорсульфоновой кислоты, щавелевой кислоты, малеиновой кислоты, янтарной кислоты, лимонной кислоты, муравьиной кислоты, бромистоводородной кислоты, бензойной кислоты, винной кислоты, фумаровой кислоты, салициловой кислоты, миндальной кислоты, карбоновой кислоты и т. д. Специалистам в данной области техники также будет понятно, что соли присоединения кислоты могут быть получены путем взаимодействия соединений с соответствующей неорганической или органической кислотой с помощью любого из ряда известных способов. Термин «фармацевтически приемлемые соли» также включает соли, полученные путем взаимодействия активного соединения, функционирующего как кислота, с неорганическим или органическим основанием с образованием соли, например, солей этилендиамина, N-метилглюкамина, лизина, аргинина, орнитина, холина, N,N'-дибензилэтилендиамина, хлорпрокаина, диэтанолamina, прокаина, N-бензилфенэтиламина, диэтиламина, пиперазина, трис-(гидроксиметил)-аминометана, гидроксида тетраметиламмония, триэтиламина, дибензиламина, эфенамина, дегидроабетиламина, N-этилпиперидина, бензиламина, тетраметиламмония, тетраэтиламмония, метиламина, диметиламина, триметиламина, этиламина, основных аминокислот и т. п. Неограничивающие примеры неорганических солей или солей металлов включают соли лития, натрия, кальция, калия, магния и т. п.

В настоящем документе термин «парентеральное введение» относится к введению посредством инъекции или инфузии. Парентеральное введение включает, но не ограничивается перечисленным, подкожное введение, внутривенное введение или внутримышечное введение.

В настоящем документе термин «подкожное введение» относится к введению непосредственно под кожу. «Внутривенное введение» означает введение в вену.

В настоящем документе термин «доза» относится к определенному количеству фармацевтического агента, которое обеспечивается за одно введение. Согласно вариантам реализации дозу можно вводить в виде двух или более болюсов, таблеток или инъекций. Согласно вариантам реализации, в которых желательным является подкожное введение, желаемая доза требует объема, который нелегко вместить за одну инъекцию. В таких вариантах

реализации две или более инъекций можно применять для обеспечения желаемой дозы. Согласно вариантам реализации дозу можно вводить в виде двух или более инъекций для снижения реакции в месте инъекции у пациента.

В настоящем документе термин «единичная лекарственная форма» относится к форме, в которой обеспечивается фармацевтический агент. Согласно вариантам реализации единичная лекарственная форма представляет собой флакон, который включает лиофилизированные соединения или композиции, описанные в настоящем документе. Согласно вариантам реализации единичная лекарственная форма представляет собой флакон, который включает восстановленные соединения или композиции, описанные в настоящем документе.

Термин «терапевтический фрагмент» (ТМ) относится к соединению, которое можно применять для лечения по меньшей мере одного симптома заболевания или нарушения, и может включать, но не ограничивается перечисленными, терапевтические полипептиды, олигонуклеотиды, малые молекулы и другие агенты, которые можно применять для лечения по меньшей мере одного симптома заболевания или нарушения. Согласно вариантам реализации ТМ модулирует активность, экспрессию и/или уровни целевого транскрипта. Согласно вариантам реализации ТМ уменьшает уровни целевого транскрипта посредством механизмов распада. Согласно вариантам реализации активность представляет собой способность целевого транскрипта связываться (например, секвестрировать) с одним или более белками. Согласно вариантам реализации ТМ модулирует активность целевого транскрипта путем снижения аффинности между целевым транскриптом и одним или более белками, которые связываются с целевым транскриптом. В результате снижения аффинности между целевым транскриптом и одним или более белками активность одного или более белков может модулироваться. Например, если один или более белков не связаны с целевым транскриптом, они доступны для выполнения своих функций, таких как, например, способствование сплайсингу, альтернативному сплайсингу и/или пропуску экзона других транскриптов. В результате функции ТМ активность, экспрессия и/или уровни нисходящих генов, которые регулируются одним или более белками, взаимодействие которых с целевым транскриптом нарушается ТМ, могут модулироваться.

Термины «модулировать», «осуществлять модуляцию» и «модуляция» относятся к нарушению экспрессии, функции или активности по сравнению с уровнем экспрессии, функции или активности до модуляции. Модуляция может включать повышение (стимуляцию или индукцию) или уменьшение (ингибирование или снижение) экспрессии, функции или активности. Согласно вариантам реализации модулируется активность целевого транскрипта. Согласно вариантам реализации модуляция активности целевого транскрипта включает уменьшение способности целевого транскрипта связываться с одним или более белками. Согласно вариантам реализации уменьшение аффинности между целевым транскриптом и одним или более белками приводит к модуляции активности одного или более белков, которые взаимодействуют с целевым транскриптом.



Например, если один или более белков не связаны с целевым транскриптом, они доступны для выполнения своих функций, таких как, например, способствование сплайсингу, альтернативному сплайсингу и/или пропуску экзона других транскриптов (например, нисходящих транскриптов). Таким образом, модуляция активности целевого транскрипта может привести к модуляции активности, экспрессии и/или уровней нисходящих генов, которые регулируются одним или более белками, взаимодействие которых с целевым транскриптом может быть нарушено.

«Аминокислота» относится к органическому соединению, которое включает

аминогруппу и карбоксильную группу и имеет общую формулу 
$$\begin{array}{c} \text{R} \\ | \\ \text{H}_2\text{N} - \text{C}_\alpha - \text{COOH} \\ | \\ \text{H} \end{array}$$
, где R может представлять собой любую органическую группу. Аминокислота может представлять собой встречающуюся в природе аминокислоту и не встречающуюся в природе аминокислоту. Аминокислота может представлять собой протеогенную аминокислоту или непротеогенную аминокислоту. Аминокислота может представлять собой L-аминокислоту или D-аминокислоту. Термин «боковая цепь аминокислоты» или «боковая цепь» относится к характеристическому заместителю («R»), связанному с  $\alpha$ -углеродом природной или не природной  $\alpha$ -аминокислоты. Аминокислота может быть включена в полипептид за счет пептидной связи.

В настоящем документе «незаряженная» аминокислота представляет собой аминокислоту, боковая цепь которой имеет суммарный нейтральный заряд при pH от 7,35 до 7,45. Примеры незаряженных аминокислот включают, но не ограничиваются перечисленными, глицин и цитруллин.

В настоящем документе «заряженная» аминокислота представляет собой аминокислоту, боковая цепь которой имеет суммарный заряд при pH от 7,35 до 7,45. Примером заряженной аминокислоты является аргинин.

В настоящем документе термин «идентичность последовательности» относится к проценту нуклеиновых кислот или аминокислот между двумя олигонуклеотидными или полипептидными последовательностями, соответственно, которые являются одинаковыми и находятся в одном и том же относительном положении. Таким образом, одна последовательность имеет определенный процент идентичности последовательности по сравнению с другой последовательностью. Для сравнения последовательностей, как правило, одна последовательность выступает в качестве референсной последовательности, с которой сравниваются тестовые последовательности. Специалистам в данной области техники будет понятно, что две последовательности обычно считаются «по существу идентичными», если они содержат идентичные остатки в соответствующих положениях. Согласно вариантам реализации идентичность последовательности между последовательностями можно определить с применением алгоритма Нидлмана-Вунша (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453), реализованного в программе

Needle пакета EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., Trends Genet.(2000), 16: 276-277), в версии, которая существует на дату подачи. Используемые параметры представляют собой штраф за открытие пропуска 10, штраф за продление пропуска 0,5 и матрицу замены EBLOSUM62 (версия EMBOSS BLOSUM62). Выходной показатель Needle, называемый «наибольшей длиной идентичного фрагмента» (полученный с применением опции «-nobrief») используют в качестве процента идентичности и вычисляют следующим образом:  $(\text{Идентичные остатки} \times 100) / (\text{Длина выравнивания} - \text{Общее число пропусков в выравнивании})$ .

В других вариантах реализации идентичность последовательности может быть определена с использованием алгоритма Смита-Уотермана в версии, которая существует на дату подачи.

В настоящем документе «гомология последовательностей» относится к проценту аминокислот между двумя полипептидными последовательностями, которые являются гомологичными и находятся в одном и том же относительном положении. Таким образом, одна полипептидная последовательность имеет определенный процент гомологии последовательности по сравнению с другой полипептидной последовательностью. Как будет понятно специалистам в данной области техники, две последовательности обычно считаются «по существу гомологичными», если они содержат гомологичные остатки в соответствующих положениях. Гомологичные остатки могут быть идентичными остатками. В качестве альтернативы, гомологичные остатки могут представлять собой неидентичные остатки с соответствующим образом сходными структурными и/или функциональными характеристиками. Например, как хорошо известно обычным специалистам в данной области техники, определенные аминокислоты, как правило, классифицируют как «гидрофобные» или «гидрофильные» аминокислоты и/или как имеющие «полярные» или «неполярные» боковые цепи, и замена одной аминокислоты другой такого же типа часто может считаться «гомологичной» заменой.

Как хорошо известно в данной области техники, аминокислотные последовательности можно сравнивать с применением любого из множества алгоритмов, в том числе доступных в коммерческих компьютерных программах, таких как BLASTP, gapped BLAST и PSI-BLAST, существующих на дату подачи. Такие программы описаны в Altschul, et al., J. Mol. Biol., (1990), 215(3): 403-410; Altschul, et al., Nucleic Acids Res. (1997), 25:3389-3402; Baxevanis et al., Bioinformatics A Practical Guide to the Analysis of Genes and Proteins, Wiley, 1998; и Misener, et al., (eds.), Bioinformatics Methods and Protocols (Methods in Molecular Biology, Vol. 132), Humana Press, 1999. В дополнение к идентификации гомологичных последовательностей программы, упомянутые выше, как правило, обеспечивают показатель степени гомологии.

В настоящем документе термин «нацеливающий на клетки фрагмент» относится к молекуле или макромолекуле, которая специфично связывается с молекулой, такой как рецептор, на поверхности клетки-мишени. Согласно вариантам реализации молекула клеточной поверхности экспрессируется только на поверхности клетки-мишени. Согласно

вариантам реализации молекула клеточной поверхности также присутствует на поверхности одной или более нецелевых клеток, но величина экспрессии молекулы клеточной поверхности выше на поверхности клеток-мишеней. Примеры нацеливающего на клетку фрагмента включают, но не ограничиваются перечисленными, антитело, пептид, белок, аптамер или малую молекулу.

В настоящем документе термины «антисмысловое соединение» и «АС» используются взаимозаменяемо для обозначения полимерной структуры нуклеиновой кислоты, которая по меньшей мере частично комплементарна молекуле целевой нуклеиновой кислоты, с которой она (АС) гибридизуется. АС может представлять собой короткий (в вариантах реализации менее 50 оснований) полинуклеотид или гомолог полинуклеотида, который включает по меньшей мере часть последовательности, комплементарную целевой последовательности. Согласно вариантам реализации АС представляет собой полинуклеотид или гомолог полинуклеотида, который включает часть, которая имеет последовательность, комплементарную целевой последовательности в цепи целевой пре-мРНК. АС может быть образован природными нуклеиновыми кислотами, синтетическими нуклеиновыми кислотами, гомологами нуклеиновых кислот или любой их комбинацией. Согласно вариантам реализации АС включает олигонуклеозиды. Согласно вариантам реализации АС включает антисмысловые олигонуклеотиды. Согласно вариантам реализации АС включает конъюгатные группы. Неограничивающие примеры АС включают, но не ограничиваются перечисленными, праймеры, зонды, антисмысловые олигонуклеотиды, олигонуклеотиды внешней направляющей последовательности (EGS), кРНК, олигонуклеотиды, олигонуклеозиды, аналоги олигонуклеотидов, миметики олигонуклеотидов и химерные комбинации указанных выше. Таким образом, эти соединения могут быть введены в одноцепочечной форме, двухцепочечной форме, кольцевой, разветвленной или шпилечной форме и могут содержать структурные элементы, такие как внутренние или концевые выпуклости или петли. Олигомерные двухцепочечные соединения могут представлять собой две цепи, гибридизованные с образованием двухцепочечных соединений, или одну цепь с достаточной самокомплементарностью для обеспечения гибридизации и образования полностью или частично двухцепочечного соединения. Согласно вариантам реализации АС модулирует (повышает, уменьшает или изменяет) экспрессию, уровни и/или активность целевого транскрипта (например, целевой нуклеиновой кислоты). Согласно вариантам реализации АС уменьшает уровень целевого транскрипта посредством индукции механизмов распада. Согласно вариантам реализации АС модулирует активность целевого транскрипта. Согласно вариантам реализации АС модулирует активность целевого транскрипта путем уменьшения способности целевого транскрипта связывать один или более белков. Согласно вариантам реализации уменьшение аффинности между целевым транскриптом и одним или более белками может привести к модуляции активности одного или более белков. Например, если один или более белков не связаны с целевым транскриптом, они доступны для выполнения своих функций, таких

как, например, способствование сплайсингу, альтернативному сплайсингу и/или пропуску экзона других транскриптов (нисходящих транскриптов). Таким образом, АС-опосредуемая модуляция активности целевого транскрипта может привести к модуляции активности, экспрессии и/или уровней нисходящих генов, которые регулируются одним или более белками, взаимодействие которых с целевым транскриптом может быть нарушено.

В настоящем документе термины «нацеливающий» или «нацеленный на» относятся к ассоциации терапевтического фрагмента, например, антисмыслового соединения, с молекулой целевой нуклеиновой кислоты или областью молекулы целевой нуклеиновой кислоты. Согласно вариантам реализации терапевтический фрагмент включает антисмысловое соединение, способное гибридизоваться с целевой нуклеиновой кислотой в физиологических условиях. Согласно вариантам реализации антисмысловое соединение нацелено на конкретную часть или сайт в целевой нуклеиновой кислоте, например, часть целевой нуклеиновой кислоты, обладающую по меньшей мере одной идентифицируемой структурой, функцией или характеристикой, такую как конкретный экзон или интрон, или выбранные нуклеиновые основания или мотивы в экзоне или интроне.

В настоящем документе термины «целевая последовательность нуклеиновой кислоты», «целевая нуклеотидная последовательность» и «целевая последовательность» относятся к последовательности нуклеиновой кислоты или нуклеотидной последовательности, с которой связывается или гибридизуется терапевтический фрагмент, такой как антисмысловое соединение. Целевые нуклеиновые кислоты включают, но не ограничиваются перечисленными, часть целевого транскрипта, целевую РНК (включая, но не ограничиваясь перечисленными, пре-мРНК и мРНК или их части), часть целевой кДНК, полученной из такой РНК, а также часть целевой нетранслируемой РНК, такой как миРНК. Например, в вариантах реализации целевая нуклеиновая кислота может представлять собой часть целевого клеточного гена (или мРНК, транскрибированной из такого гена), экспрессия или транскрипция которой ассоциирована с конкретным нарушением или патологическим состоянием. Термин «часть» относится к определенному числу смежных (т. е. связанных) нуклеотидов нуклеиновой кислоты.

В настоящем документе термин «транскрипт» или «транскрипт гена» относится к молекуле РНК, транскрибированной из ДНК, и включает, но не ограничивается этим, мРНК, пре-мРНК и частично процессированную РНК.

Термины «целевой транскрипт» и «целевая РНК» относятся к пре-мРНК или транскрипту мРНК, который связан терапевтическим фрагментом. Целевой транскрипт может включать целевую нуклеотидную последовательность. Согласно вариантам реализации целевой транскрипт включает целевую нуклеотидную последовательность, которая включает экспансированный тринуклеотидный повтор CUG.

Термины «целевой ген» и «представляющий интерес ген» относятся к гену, модуляция экспрессии и/или активности которого является желаемой или предусмотренной.

Целевой ген может быть транскрибирован в целевой транскрипт, который включает целевую нуклеотидную последовательность. Целевой транскрипт может быть транслирован в представляющий интерес белок.

Термин «целевой белок» относится к полипептиду или белку, кодируемому целевым транскриптом (например, целевой мРНК).

В настоящем документе термин «мРНК» относится к молекуле РНК, которая кодирует белок и включает пре-мРНК и зрелую мРНК. «Пре-мРНК» относится к недавно синтезированной молекуле эукариотической мРНК непосредственно после транскрипции ДНК. Согласно вариантам реализации пре-мРНК кэпирована 5'-кэпом, модифицирована 3'-поли-А-хвостом и/или сплайсирована с получением зрелой последовательности мРНК. Согласно вариантам реализации пре-мРНК включает один или более интронов. Согласно вариантам реализации пре-мРНК подвергается процессу, известному как сплайсинг, для удаления интронов и соединения экзонов. Согласно вариантам реализации пре-мРНК включает один или более элементов сплайсинга или регуляторных элементов сплайсинга. Согласно вариантам реализации пре-мРНК включает сайт полиаденилирования.

В настоящем документе термин «экспрессия», «генная экспрессия», «экспрессия гена» или т. п. относится ко всем функциям и стадиям, посредством которых информация, кодируемая в гене, превращается в функциональный генный продукт, такой как полипептид или некодирующая РНК, в клетке. Примеры некодирующей РНК включают транспортную РНК (тРНК) и рибосомную РНК. Экспрессия гена полипептида включает транскрипцию гена с образованием пре-мРНК, процессинг пре-мРНК с образованием зрелой мРНК, транслокацию зрелой мРНК из ядра в цитоплазму, трансляцию зрелой мРНК в полипептид и сборку кодируемого полипептида. Экспрессия включает частичную экспрессию. Например, экспрессия гена может называться образованием транскрипта гена. Трансляция зрелой мРНК может называться экспрессией зрелой мРНК.

В настоящем документе термин «модуляция экспрессии гена» или т. п. относится к модуляции одного или более процессов, ассоциированных с экспрессией гена. Например, модификация экспрессии гена может включать модификацию одного или более из транскрипции гена, процессинга РНК, транслокации РНК из ядра в цитоплазму и трансляции мРНК в белок.

В настоящем документе термин «ген» относится к последовательности нуклеиновой кислоты, которая охватывает 5'-промоторную область, ассоциированную с экспрессией генного продукта, и любые интронные и экзонные области и 3'-нетранслируемые области («НТО»), ассоциированные с экспрессией генного продукта.

Термин «иммунная клетка» относится к клетке гемопозитического происхождения, которая играет роль в иммунном ответе. Иммунные клетки включают, но не ограничиваются перечисленными, лимфоциты (например, В-клетки и Т-клетки), естественные клетки-киллеры (НК) и миелоидные клетки. Термин «миелоидные клетки» включает моноциты, макрофаги и гранулоциты (например, базофилы, нейтрофилы, эозинофилы и тучные клетки). Моноциты представляют собой лимфоциты, которые

циркулируют в крови в течение 1-3 дней, после чего они либо мигрируют в ткани и дифференцируются в макрофаги или воспалительные дендритные клетки, либо погибают. Термин «макрофаг» в контексте настоящего документа включает макрофаги фетального происхождения (которые также могут называться резидентными тканевыми макрофагами) и макрофаги, происходящие из моноцитов, которые мигрировали из кровотока в ткань в организме (которые могут называться макрофагами моноцитарного происхождения). В зависимости от того, в какой ткани расположен макрофаг, его называют клеткой Купфера (печень), внутриклубочковой мезангиальной клеткой (почка), альвеолярным макрофагом (легкие), синусовым гистиоцитом (лимфатические узлы), клеткой Гофбауэра (плацента), микроглией (головной и спинной мозг) или клеткой Лангерганса (кожа), помимо прочего.

В настоящем документе термин «олигонуклеотид» относится к олигомерному соединению, содержащему множество связанных нуклеотидов или нуклеозидов. Один или более нуклеотидов олигонуклеотида могут быть модифицированы. Олигонуклеотид может содержать рибонуклеиновую кислоту (РНК) или дезоксирибонуклеиновую кислоту (ДНК). Олигонуклеотиды могут состоять из природных и/или модифицированных нуклеиновых оснований, сахаров и ковалентных межнуклеозидных связей и могут дополнительно включать конъюгаты, отличные от нуклеиновой кислоты.

В настоящем документе термин «нуклеозид» относится к гликозиламину, который включает нуклеиновое основание и сахар. Нуклеозиды включают, но не ограничиваются перечисленными, природные нуклеозиды, лишенные азотистого основания нуклеозиды, модифицированные нуклеозиды и нуклеозиды, имеющие миметические основания и/или сахарные группы. «Природный нуклеозид» или «немодифицированный нуклеозид» представляет собой нуклеозид, который включает природное нуклеиновое основание и природный сахар. Природные нуклеозиды включают нуклеозиды РНК и ДНК.

В настоящем документе термин «природный сахар» относится к сахару нуклеозида, который не является модифицированным относительно его встречающейся в природе формы в РНК (2'-ОН) или ДНК (2'-Н).

В настоящем документе термин «нуклеотид» относится к нуклеозиду, имеющему фосфатную группу, ковалентно связанную с сахаром. Нуклеотиды можно модифицировать с помощью любого из множества заместителей.

В настоящем документе термин «нуклеиновое основание» относится к части основания нуклеозида или нуклеотида. Нуклеиновое основание может включать любой атом или группу атомов, способных к образованию водородных связей с основанием другой нуклеиновой кислоты. Природное нуклеиновое основание представляет собой нуклеиновое основание, которое является немодифицированным по сравнению с его встречающейся в природе формой в РНК или ДНК.

В настоящем документе термин «фрагмент гетероциклического основания» относится к нуклеиновому основанию, которое включает гетероцикл.

В настоящем документе термин «межнуклеозидная связь» относится к ковалентной связи между соседними нуклеозидами.

В настоящем документе термин «природная межнуклеозидная связь» относится к фосфодиэфирной связи от 3' к 5'.

В настоящем документе термин «модифицированная межнуклеозидная связь» относится к любой связи между нуклеозидами или нуклеотидами, отличной от встречающейся в природе межнуклеозидной связи.

В настоящем документе термин «олигонуклеозид» относится к олигонуклеотиду, в котором межнуклеозидные связи не содержат атома фосфора.

В настоящем документе термин «химерное антисмысловое соединение» относится к антисмысловому соединению, имеющему по меньшей мере один сахар, нуклеиновое основание и/или межнуклеозидную связь, которые являются дифференциально модифицированными по сравнению с другими сахарами, нуклеиновыми основаниями и межнуклеозидными связями в том же олигомерном соединении. Остальная часть сахаров, нуклеиновых оснований и межнуклеозидных связей может быть независимо модифицированной или немодифицированной. В целом, химерное олигомерное соединение будет содержать модифицированные нуклеозиды, которые могут находиться в отдельных положениях или могут группироваться вместе в областях, которые будут определять конкретный мотив. Любая комбинация модификаций и/или миметических групп может включать химерное олигомерное соединение, описанное в настоящем документе.

В настоящем документе термин «антисмысловой олигонуклеотид со смешанным остовом» относится к антисмысловому олигонуклеотиду, в котором по меньшей мере одна межнуклеозидная связь антисмыслового олигонуклеотида отличается от по меньшей мере одной другой межнуклеозидной связи антисмыслового олигонуклеотида.

В настоящем документе термин «комплементарность нуклеинового основания» относится к нуклеиновому основанию, которое способно к спариванию оснований с другим нуклеиновым основанием. Например, в ДНК аденин (A) комплементарен тимину (T). Например, в РНК аденин (A) комплементарен урацилу (U). Согласно вариантам реализации комплементарное нуклеиновое основание относится к нуклеиновому основанию антисмыслового соединения, которое способно к спариванию оснований с нуклеиновым основанием своей целевой нуклеиновой кислоты. Например, если нуклеиновое основание в определенном положении антисмыслового соединения способно к образованию водородной связи с нуклеиновым основанием в определенном положении целевой нуклеиновой кислоты, то положение водородной связи между олигонуклеотидом и целевой нуклеиновой кислотой считается комплементарным в этой паре нуклеиновых оснований.

В настоящем документе термин «некомплементарное нуклеиновое основание» относится к паре нуклеиновых оснований, которые не образуют водородных связей друг с другом или иным образом не поддерживают гибридизацию.

В настоящем документе термин «комплементарный» относится к способности олигомерного соединения гибридизоваться с другим олигомерным соединением или

нуклеиновой кислотой посредством комплементарности нуклеиновых оснований. Согласно вариантам реализации антисмысловое соединение и его мишень комплементарны друг другу, когда достаточное число соответствующих положений в каждой молекуле занято нуклеиновыми основаниями, которые могут связываться друг с другом для обеспечения стабильной ассоциации между антисмысловым соединением и мишенью. Специалисту в данной области техники понятно, что включение несовпадений возможно без устранения способности олигомерных соединений оставаться в ассоциации. Таким образом, в настоящем документе описаны антисмысловые соединения, которые могут включать до примерно 20% нуклеотидов, которые не совпадают (т. е. не являются нуклеиновыми основаниями, комплементарными соответствующим нуклеотидам мишени). Согласно вариантам реализации антисмысловые соединения содержат не более примерно 15%, например, не более примерно 10%, например, не более 5% несоответствия или несоответствия отсутствуют. Остальные нуклеотиды представляют собой нуклеиновые основания, комплементарные или иным образом не нарушающие гибридизацию (например, универсальные основания). Обычный специалист в данной области техники поймет, что соединения, предложенные в настоящем документе, составляют по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или 100% нуклеиновых оснований, комплементарных целевой нуклеиновой кислоте.

В настоящем документе термин «гибридизация» означает спаривание комплементарных олигомерных соединений (например, антисмыслового соединения и его целевой нуклеиновой кислоты). Хотя это не ограничивается конкретным механизмом, наиболее распространенный механизм спаривания включает образование водородной связи, которая может быть водородной связью по Уотсону-Крику, по Хугстину или обратной водородной связью по Хугстину, между комплементарными нуклеозидными или нуклеотидными основаниями (нуклеиновыми основаниями). Например, природное основание аденин является нуклеиновым основанием, комплементарным природным нуклеиновым основаниям тимину и урацилу, которые спариваются посредством образования водородных связей. Природное основание гуанин представляет собой нуклеиновое основание, комплементарное природным основаниям цитозину и 5-метилцитозину. Гибридизация может происходить при различных условиях.

В настоящем документе термин «специфично гибридизуется» относится к способности олигомерного соединения гибридизоваться с одним сайтом нуклеиновой кислоты с большей аффинностью, чем оно гибридизуется с другим сайтом нуклеиновой кислоты. Согласно вариантам реализации антисмысловой олигонуклеотид специфично гибридизуется с более чем одним целевым сайтом. Согласно вариантам реализации олигомерное соединение специфично гибридизуется с его мишенью в строгих условиях гибридизации.



«Строгие условия гибридизации» и «строгие условия промывки при гибридизации» в контексте гибридизации нуклеиновых кислот зависят от последовательности и отличаются при различных параметрах окружающей среды. Подробное руководство по гибридизации нуклеиновых кислот можно найти в Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes* часть I глава 2 «Overview of principles of hybridization and the strategy of nucleic acid probe assays» Elsevier, New York (1993). Обычно очень строгие условия гибридизации и промывки выбирают так, чтобы они были примерно на 5°C ниже температуры плавления ( $T_m$ ) для конкретной последовательности при определенной ионной силе и pH.  $T_m$  представляет собой температуру (при определенной ионной силе и pH), при которой 50% целевой последовательности гибридизуется с идеально соответствующим зондом. Очень строгие условия выбирают таким образом, чтобы они были равны  $T_m$  для конкретного зонда. Примером строгих условий гибридизации для гибридизации комплементарных нуклеотидных последовательностей, которые содержат более 100 комплементарных остатков, на фильтре при Саузерн-блоттинге или Нозерн-блоттинге является 50% формамид с 1 мг гепарина при 42°C, при этом гибридизацию проводят в течение ночи. Примером очень строгих условий промывки являются 0,15 М NaCl при 72°C в течение примерно 15 минут. Примером строгих условий промывки является промывка 0,2x SSC при 65°C в течение 15 минут (см. Sambrook and Russel, *Molecular Cloning: A laboratory Manual*, 3<sup>rd</sup> ed., Cold Spring Harbor Laboratory Press, 2001, где приведено описание буфера SSC). Часто промывке в строгих условиях предшествует промывка в условиях низкой строгости для удаления фонового сигнала зонда. Примером промывки в условиях средней строгости для дуплекса, состоящего, например, более чем из 100 нуклеотидов, является 1x SSC при 45°C в течение 15 минут. Примером промывки с низкой строгостью для дуплекса, состоящего, например, более чем из 100 нуклеотидов, является 4-6x SSC при 40°C в течение 15 минут. Для коротких зондов (например, примерно от 10 до 50 нуклеотидов) строгие условия, как правило, включают концентрации солей менее примерно 1,0 М ионов Na, как правило, от примерно 0,01 до 1,0 М концентрации ионов Na (или других солей) при pH от 7,0 до 8,3, и температура, как правило, составляет по меньшей мере примерно 30°C. Строгие условия также могут быть достигнуты путем добавления дестабилизирующих агентов, таких как формамид.

В настоящем документе термин «2'-модифицированный» или «2'-замещенный» означает сахар, который включает заместитель в 2'-положении, отличный от H или OH. 2'-модифицированные мономеры, включают, но не ограничиваются перечисленными, BNA и мономеры (например, нуклеозиды и нуклеотиды) с 2'-заместителями, такими как аллил, амин, азид, тио, O-аллил, O-C1-C10 алкил, -OCF<sub>3</sub>, O-(CH<sub>2</sub>)<sub>2</sub>-O-CH<sub>3</sub>, 2'-O(CH<sub>2</sub>)<sub>2</sub>SCH<sub>3</sub>, O-(CH<sub>2</sub>)<sub>2</sub>-O-N(R<sub>m</sub>)(R<sub>n</sub>) или O-CH<sub>2</sub>-C(=O)-N(R<sub>m</sub>)(R<sub>n</sub>), где каждый R<sub>m</sub> и R<sub>n</sub> независимо представляет собой H или замещенный или незамещенный C<sub>1</sub>-C<sub>10</sub> алкил.

В настоящем документе термин «МОЕ» относится к 2'-O-метоксиэтильному заместителю.

В настоящем документе термин «модифицированный нуклеотид с высокой аффинностью» относится к нуклеотиду, содержащему по меньшей мере одно модифицированное нуклеиновое основание, межнуклеозидную связь или фрагмент сахара так, что модификация повышает аффинность антисмыслового соединения, которое включает модифицированный нуклеотид, в отношении целевой нуклеиновой кислоты. Высокоаффинные модификации включают, но не ограничиваются перечисленными, BNA, LNA и 2'-МОЕ.

В настоящем документе термин «миметик» относится к группам, которые замещают сахар, нуклеиновое основание и/или межнуклеозидную связь в АС. Обычно миметик используется вместо сахара или комбинации сахар-межнуклеозидная связь, а нуклеиновое основание сохраняется для гибридизации с выбранной мишенью. Типичные примеры миметика сахара включают, но не ограничиваются перечисленными, циклогексенил или морфолино. Типичные примеры миметика для комбинации сахар-межнуклеозидная связь включают, но не ограничиваются перечисленными, пептидные нуклеиновые кислоты (PNA) и морфолиновые группы, связанные незаряженными ахиральными связями. В некоторых случаях вместо нуклеинового основания используют миметик. Типичные миметики нуклеиновых оснований хорошо известны в данной области техники и включают, но не ограничиваются перечисленными, аналоги на основе трициклического феноксазина и универсальные основания (Berger et al., Nuc Acid Res. 2000, 28:2911-14, включена в настоящий документ посредством ссылки). Способы синтеза миметиков сахаров, нуклеозидов и нуклеиновых оснований хорошо известны специалистам в данной области техники.

В настоящем документе термин «бициклический нуклеозид» или «BNA» относится к нуклеозиду, в котором фуранозная часть нуклеозида включает мостик, соединяющий два атома на фуранозном кольце с образованием бициклической кольцевой системы. BNA включают, но не ограничиваются перечисленными,  $\alpha$ -L-LNA,  $\beta$ -D-LNA, ENA, оксиамино BNA (2'-O-N(CH<sub>3</sub>)-CH<sub>2</sub>-4') и аминокси BNA (2'-N(CH<sub>3</sub>)-O-CH<sub>2</sub>-4').

В настоящем документе термин «4'-2' бициклический нуклеозид» относится к BNA, в которой мостик, соединяющий два атома фуранозного кольца, соединяет 4'-атом углерода и 2'-атом углерода фуранозного кольца с образованием бициклической кольцевой системы.

В настоящем документе «блокированная нуклеиновая кислота» или «LNA» относится к нуклеотиду, модифицированному таким образом, что 2'-гидроксильная группа рибозильного сахарного кольца связана с 4'-атомом углерода сахарного кольца за счет метиленовой группы с образованием 2'-C,4'-C-оксиметиленовой связи. LNA включают, но не ограничиваются перечисленными,  $\alpha$ -L-LNA и  $\beta$ -D-LNA.

В настоящем документе термин «кэп-структура» или «концевой кэп-фрагмент» относится к химическим модификациям, которые были включены на любом конце АС. Термин «терапевтический полипептид» относится к встречающейся в природе или полученной рекомбинантным способом макромолекуле, которая включает две или более

аминокислот и обладает терапевтической, профилактической или другой биологической активностью.

Термин «малая молекула» относится к органическому соединению с фармакологической активностью и молекулярной массой менее примерно 2000 дальтон, или менее примерно 1000 дальтон, или менее примерно 500 дальтон. Низкомолекулярные терапевтические средства, как правило, изготавливают с помощью химического синтеза.

«Целевой белок дикого типа» относится к изомеру нативного функционального белка, продуцируемому целевым геном дикого типа, нормальным или немутированным вариантом целевого гена. Целевой белок дикого типа также относится к белку, полученному из целевой пре-мРНК, которая была повторно сплайсирована.

«Повторно сплайсированный целевой белок» в контексте настоящего документа относится к белку, кодируемому мРНК, полученной в результате сплайсинга целевой пре-мРНК, с которой гибридизуется АС. Повторно сплайсированный целевой белок может быть идентичен целевому белку дикого типа, может быть гомологичен целевому белку дикого типа, может представлять собой функциональный вариант целевого белка дикого типа, может представлять собой изоформу целевого белка дикого типа или может быть активным фрагментом целевого белка дикого типа.

В настоящем документе «экспансированный тринуклеотидный повтор», такой как «экспансированный» CUG и/или «экспансированный» повтор CTG, означает ген, содержащий или кодирующий тринуклеотидный повтор, содержащий ряд повторяющихся последовательных тринуклеотидов, который больше, чем присутствует в гене дикого типа. Экспансированные нуклеотидные повторы могут быть записаны как XXX·NNN или (XXX·NNN), где XXX относится к повтору ДНК, и NNN относится к повтору РНК, который транскрибируется из повтора ДНК. Например, повтор CTG·CUG относится к гену, содержащему повтор ДНК CTG, из которого транскрибируется РНК, содержащая повтор CUG. Согласно вариантам реализации число повторов в экспансированном тринуклеотидном повторе составляет 5 или более, 10 или более, 15 или более, или 20 или более по сравнению с геном дикого типа. Согласно вариантам реализации экспансированный тринуклеотидный повтор включает 2х, 3х, 4х, 5х, 10х, 20х, 50х или более тринуклеотидных повторов, чем ген дикого типа. Экспансированный тринуклеотидный повтор может привести к заболеванию у субъекта, имеющего ген, который содержит экспансированный тринуклеотидный повтор. Например, субъект, имеющий экспансированный повтор CTG в гене, может страдать DM1 или FECO. При DM1 ген DPMK содержит экспансированный повтор CTG. Субъекты, которые страдают DM1, могут иметь 50 или более повторов CTG в 3'-нетранслируемой области (НТО) гена DPMK, тогда как непораженные субъекты, как правило, имеют 5-34 повтора CTG в 3'-НТО гена DPMK. При FECO ген TCF4 содержит экспансированный повтор CTG. Субъекты, которые страдают FECO, могут иметь 40 или более повторов CTG в локусе CTG18.1 гена TCF4, в то время как непораженные субъекты, как правило, имеют 30 или менее повторов CTG в локусе CTG18.1 гена TCF4. мРНК, транскрибируемая из гена,

содержащего экспансированный повтор CTG, будет содержать экспансированный повтор CUG.

Термин «нисходящий» в настоящем раскрытии, поскольку он связан с геном, мРНК или белком, относится к гену, мРНК или белку, на которые влияет связывание АС с нуклеотидом-мишенью (например, целевым транскриптом), но не является геном, мРНК или белком, соответствующим целевому нуклеотиду. Связывание АС с целевым нуклеотидом может снижать агрегацию или секвестрацию РНК-связывающего белка, такого как MBNL1 или CUGBP1, на накопленной мРНК, содержащей повторы CUG, что может сделать такие РНК-связывающие белки доступными для надлежащей транскрипции, процессинга РНК и/или экспрессии продуктов нисходящих генов.

В настоящем документе «функциональный фрагмент» или «активный фрагмент» относится к части эукариотического целевого белка дикого типа, которая проявляет активность, такую как один или более видов активности полноразмерного целевого белка дикого типа, или которая обладает другой активностью. Согласно вариантам реализации повторно сплайсированный целевой белок, который обладает по меньшей мере одним видом биологической активности целевого белка дикого типа, считается активным фрагментом целевого белка дикого типа. Активность может представлять собой любой процент активности (т.е. более или менее) полноразмерного целевого белка дикого типа, включая, но не ограничиваясь перечисленными, примерно 1% от активности, примерно 2%, примерно 3%, примерно 4%, примерно 5%, примерно 10%, примерно 20%, примерно 30%, примерно 40%, примерно 50%, примерно 60%, примерно 70%, примерно 80%, примерно 90%, примерно 95%, примерно 96%, примерно 97%, примерно 98%, примерно 99%, примерно 100%, примерно 200%, примерно 300%, примерно 400%, примерно 500% или более (включая все значения и диапазоны между этими значениями) активности по сравнению с целевым белком дикого типа. Таким образом, в вариантах реализации активный фрагмент может сохранять по меньшей мере часть одного или более видов биологической активности целевого белка дикого типа. Согласно вариантам реализации активный фрагмент может усиливать один или более видов биологической активности целевого белка дикого типа.

«Целевой белок дикого типа» относится к изомеру нативного функционального белка, продуцируемому целевым геном дикого типа, нормальным или немутированным вариантом целевого гена. Целевой белок дикого типа также относится к белку, полученному из целевой пре-мРНК, которая была надлежащим образом сплайсирована.

В настоящем документе термины «сплайсинг» и «процессинг» относятся к модификации пре-мРНК после транскрипции, в которой интроны удаляются и экзоны соединяются. Сплайсинг осуществляется в ряде реакций, которые катализируются крупным комплексом РНК-белок, состоящим из пяти небольших ядерных рибонуклеопротеинов (мяРНП), называемым сплайсосомой. В интроне для сплайсинга требуются 3'-сайт сплайсинга, 5'-сайт сплайсинга и сайт ветвления. РНК-компоненты мяРНП взаимодействуют с интроном и могут участвовать в катализе.

В настоящем документе альтернативный сплайсинг относится к сплайсингу различных комбинаций экзонов, присутствующих в гене, что приводит к образованию различных транскриптов мРНК из одного гена.

«Повторно сплайсированный целевой белок» в контексте настоящего документа относится к белку, кодируемому мРНК, полученной в результате сплайсинга целевой пре-мРНК, с которой гибридизуется АС. Повторно сплайсированный целевой белок может быть идентичен целевому белку дикого типа, может быть гомологичен целевому белку дикого типа, может представлять собой функциональный вариант целевого белка дикого типа или может представлять собой активный фрагмент целевого белка дикого типа.

Все публикации, патенты и заявки на патент, упомянутые в данном описании, указывают на уровень квалификации специалистов в области техники, к которой относится данное изобретение. Все публикации, патенты и заявки на патенты включены в настоящую заявку посредством ссылки в той же степени, как если бы каждая отдельная публикация или заявка на патент была конкретно и по отдельности указана для включения посредством ссылки.

## **ПРИМЕРЫ**

### **Пример 1. Оценка влияния соединений РМО и РМО-EEV на образование очагов РНК и восстановление сплайсинга в клеточных линиях, связанных с DM1**

Влияние соединения РМО с повтором (CUG)<sub>7</sub> и РМО-EEV с повтором (CUG)<sub>7</sub> (А и D в Таблице 12) на образование очагов РНК и восстановление сплайсинга оценивали в клеточных линиях, связанных с DM1. Также включали соединение РМО-EEV, имеющее несовпадение последовательности РМО, и РМО, имеющий случайную последовательность (В и С в Таблице 12).

#### **Экспериментальная часть**

*Клетки.* РМО и РМО-EEV оценивали в модели DM1 на клетках HeLa (HeLa-480), которая представляет собой стабильную клеточную линию с высокой нагрузкой повтором CUG и последующими дефектами сплайсинга, и в миобластах DM1, полученных от пациента, страдающего DM1 с примерно 2600 повторами CTG и последующими дефектами сплайсинга. HeLa-480 воспроизводит патогенные отличительные признаки DM1, включая рибонуклеарные очаги CUG и ошибочный сплайсинг пре-мРНК-мишеней факторов альтернативного сплайсинга muscleblind (MBNL). Было отмечено, что миобласты DM1 достаточно медленно растут (время удвоения составляет примерно 7 дней) и плохо трансфицируются. Контрольные клетки HeLa и HeLa-480 обрабатывали ENDOPORTER без дополнительных соединений. Клетки HeLa-480 представляют патологическое состояние, и клетки HeLa представляют непатологическое состояние.

*Анализ очагов РНК.* Клетки HeLa-480 или миобласты DM1 обрабатывали 1 мкМ, 3 мкМ или 10 мкМ соединений А-D (Таблица 12). Все соединения трансфицировали без трансфекционного реагента или с использованием трансфекционного агента ENDOPORTER (доступного от GENETOOLS LLC в Филомете, Орегон), разработанного для доставки естественно заряженных РМО в клетки. Клетки инкубировали в течение 24

часов, а затем фиксировали для качественного анализа очагов РНК с помощью микроскопии. При качественном визуальном осмотре обработанных клеток HeLa-480 соединение А (РМО-ЕЕV) показало наименьшее количество очагов РНК. Не делали заключения в отношении миобластов DM1.

*Анализ сплайсинга и ОТ-ПЦР.* Клетки HeLa-480, обработанные, как описано выше, собирали через 24 или 48 часов инкубации и экстрагировали общую РНК. Выполняли ОТ-ПЦР для измерения и/или количественного определения профилей сплайсинга экзонов, пораженных при DM, в целевых РНК, таких как MBNL1 и CLASP1. Оценивали процент представляющего интерес включения экзона.

**Таблица 12:** РМО и РМО-ЕЕV, протестированные в Примере 1.

Идентификатор	Последовательность (все РМО)	Чистота	Растворитель
A	CAGCAGCAGCAGCAGCAGCAG-клик-К-PEG12-К(FfΦRrRrQ)-PKKKRKV-Ас (SEQ ID NO: 154)-клик-К-PEG12-(SEQ ID NO:78)-(SEQ ID NO: 42)-Ас	91	Солевой раствор (0,9% NaCl)
B	GТААСТGТАТТТGGТАСТТСС-С3-NH2-PEG4-COT-PEG12-PKKKRKV-(FfΦRrRrQ) (SEQ ID NO: 317)-С3-NH2-PEG4-COT-PEG12-(SEQ ID NO: 78)-(SEQ ID NO:42)-Ас	99	PBSX1
C	AGCCAGAGCACCGCAACCGGACGAG (SEQ ID NO: 318)	81	Солевой раствор (0,9% NaCl)
D	CAGCAGCAGCAGCAGCAGCAG (SEQ ID NO: 154)	75	Солевой раствор (0,9% NaCl)

**Результаты.** На **ФИГ. 6А-6D** показаны результаты ОТ-ПЦР событий альтернативного сплайсинга MBNL1 (**6А** и **6С**, включение экзона 5) и CLASP1 (**6В** и **6D**, включение экзона 19) через 24 часа (**6А** и **6В**) и 48 часов (**6С** и **6D**) после обработки клеток HeLa-480 соединениями А-D и трансфекционным агентом ENDOPORTER. Снижение включения экзона 5 в MBNL1 наблюдали для клеток, обработанных каждым из соединений А-D, как в момент времени 24 ч, так и в момент времени 48 ч (**ФИГ. 6А** и **6С**). Клетки, обработанные соединением А, показали наибольшее восстановление (уменьшение включения экзона 5). Увеличение включения экзона 19 в CLASP1 наблюдали для клеток, обработанных каждым из соединений А-D, как в момент времени 24 ч, так и в момент времени 48 ч (**ФИГ. 6В** и **6D**). Клетки, обработанные соединением А, показали наибольшее восстановление (увеличение включения экзона 19). Напротив, когда клетки

HeLa-480 обрабатывали A-D без трансфекционного реагента ENDOPORTER, не наблюдали изменений в событиях сплайсинга MBNL1 (ФИГ. 6Е) или CLASP1 (ФИГ. 6F), за исключением соединения А.

На ФИГ. 7А-7В показаны результаты ОТ-ПЦР событий альтернативного сплайсинга MBNL1 и CLASP1 после обработки миобластов DM1 соединением A-D, отрицательным контролем DM-04 или положительным контролем DM-05. Обработка всеми соединениями привела к восстановлению событий сплайсинга MBNL1 (ФИГ. 7А) и CLASP1 (ФИГ. 7В).

**Пример 2. Оценка влияния EEV-PMO 221-1106 на восстановление сплайсинга в модели DM1 на мышах**

Влияние PMO по отдельности и PMO-EEV (Таблица 13) на восстановление сплайсинга оценивали *in vivo* с использованием модели DM1 на мышцах HSA-LR.

**Экспериментальная часть.**

*Модель на мышах.* HSA-LR представляет собой трансгенную мышиную модель, характеризующуюся экспансированными длинными повторами CUG (LR) в 3'-НТО трансгена скелетного актина человека (HSA) и экспрессирующую РНК CUGexp (например, РНК с экспансированным CUG) на высоких уровнях в скелетной мышце (Mankodi et al., Science 2000, 289(5485):1769-1773). Мыши HSA-LR характеризуются миотоническим фенотипом с дефектами сплайсинга. В качестве контроля и для получения трансгенных мышей HSA-LR использовали модель на мышцах Friend Virus В NIH Jackson (FVB/NJ).

*Экспериментальный дизайн.* Соединения А-D вводили мышам путем ретроорбитальной инъекции или внутривенной (в/в) инъекции в виде однократной дозы 100 мкл раствора соединения на 20 г массы тела. Шкала была пропорциональна массе тела каждой мыши (например, 150 мкл на 30 г массы тела).

**Таблица 13:** PMO и PMO-EEV, протестированные в Примере 2.

Идентификатор	Соединение	Последовательность (все PMO)	Конц. (мг/мл)	Растворитель
(А)	Солевой раствор	Контрольный носитель	Н/П	Солевой раствор (0,9% NaCl)
(В)	221	5'-CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154)	6	Солевой раствор (0,9% NaCl)

(C)	221-1106 при 20 мг/кг	5'-CAG CAG CAG CAG CAG CAG CAG-3'-PEG4COT -клик-К- миниPEG2-Lys(FfFGRGRE)- миниPEG2-VKRKKKP-Ас (SEQ ID NO: 154)-PEG4COT- клик-К-миниPEG2-К(SEQ ID NO: 76)-миниPEG2-(SEQ ID NO: 42)- Ас	4	Солевой раствор (0,9% NaCl)
(D)	221-1106 при 40 мг/кг	5'-CAG CAG CAG CAG CAG CAG CAG-3'-PEG4COT-клик-К- миниPEG2-Lys(FfFGRGRE)- миниPEG2-VKRKKKP-АС (SEQ ID NO: 154)-PEG4COT- клик-К-миниPEG2-К(SEQ ID NO: 76)-миниPEG2-(SEQ ID NO: 42)- Ас	8	Солевой раствор (0,9% NaCl)

Животных соотносили по возрасту и распределяли в шесть групп обработки. Использовали две контрольные группы; Группа 1: мыши FVB/NJ (FVB/NJ; контроль без заболевания) и Группа 2: мыши HSA-LR (контрольные пораженные животные), каждой мышши вводили путем инъекции солевой раствор. Использовали четыре группы обработки (Группы А-D); мышам HSA-LR вводили путем инъекции соединения А, В, С или D. Мышам FVB/NJ было 5 недель и 4 дня на момент инъекции, в то время как мышам HSA-LR было 6 недель и 1 или 2 дня на момент инъекции. Для этого эксперимента использовали по четыре мышши (два самца и две самки) на группу. Мышей умерщвляли через 1 неделю после обработки. Ткани (икроножная мышца, четырехглавая мышца, передняя большеберцовая мышца (ТА)) собирали и мгновенно замораживали в жидком азоте и хранили при -80°C для дальнейшей оценки в анализе восстановления сплайсинга.

Общую РНК экстрагировали из образцов тканей и анализировали с помощью ОТ-ПЦР для оценки АС-индуцированного восстановления событий альтернативного сплайсинга РНК в (i) экзоне 22 *Atp2a1*, (ii) экзоне 7 *Nfix*, (iii) экзоне 7a *C1cn1* и (iv) экзоне 5 *Mbnl1*. Оценивали процент представляющего интерес включения экзона.

**Результаты.** Перед обработкой все мышши HSA-LR имели миотонию. После инъекции соединений все мышши были дезориентированы. Мыши, которым инъецировали соединения А и В (группы А и В), восстановились в течение 15 минут, тогда как мышсам, которым инъецировали соединения С и D (группы С и D), требовалось несколько часов для восстановления. Все обработанные мышши полностью восстановились к следующему



дню. Во время умерщвления группы А и В имели миотонию; однако группы С и D явно не имели миотонии.

На **ФИГ. 8А-10D** показаны измерения сплайсинга РНК для *Atp2a1* (для включения экзона 22; **ФИГ. 8А, 9А и 10А**), *Nfix* (для включения экзона 7; **ФИГ. 8В, 9В и 10В**), *Cln1* (для включения экзона 7а; **ФИГ. 8С, 9С и 10С**) и *Mbn1* (для включения экзона 5; **ФИГ. 8D, 9D и 10D**) в ткани икроножной мышцы (**ФИГ. 8А-8D**), четырехглавой мышцы (**ФИГ. 9А-9D**) и передней большеберцовой мышцы (**ФИГ. 10А-10D**) обработанных мышей. У мышей, обработанных соединениями С и D (РМО-EEV), наблюдали восстановление событий сплайсинга *Atp2a1* и *Nfix* в ткани икроножной, четырехглавой и передней большеберцовой мышцы, в то время как в группах, получавших РМО и солевой раствор, не наблюдали восстановления событий сплайсинга (**ФИГ. 8А, 8В, 9А, 9В, 10А и 10В**). Аналогичным образом, в ткани икроножной и четырехглавой мышц мышей, обработанных соединениями С и D, наблюдали восстановление событий сплайсинга *Cln1* (**ФИГ. 8С и 9С**) и *Mbn1* (**ФИГ. 8D и 9D**), в то время как в группах, получавших РМО и солевой раствор, не наблюдали восстановления событий сплайсинга. В отношении сплайсинга *Cln1* и *Mbn1* в ткани большеберцовой мышцы (**ФИГ. 10С-10D**) никаких дефектов альтернативного сплайсинга не детектировали в контрольной линии мышей (FVB/NJ) и модели DM1 на мышах (HSA-LR). Таким образом, обработка соединениями А-Д не приводила к восстановлению сплайсинга в генах *Cln1* и *Mbn1* в ткани большеберцовой мышцы.

Эти результаты демонстрируют положительное влияние обработки РМО-EEV на восстановление сплайсинга в исследовании *in vivo* с использованием модели DM1 на мышах, а также возможность применения соединений РМО-EEV для лечения миотонической дистрофии (DM).

### **Пример 3. Оценка различных соединений РМО-EEV для коррекции событий ошибочного сплайсинга в иммортализованных миобластах от пациентов, страдающих DM1**

Влияние двух DMPK CUG-нацеленных РМО-EEV (197-777 и 221-1106; **Таблица 14**) на восстановление сплайсинга генов, связанных с DM1, оценивали *in vitro* с применением мышечных миобластов и мышечных трубочек, полученных от пациентов, страдающих DM1.

#### **Экспериментальная часть.**

*Клеточная культура.* Получали иммортализованные миобласты от пациентов, страдающих DM1 (ASA308DM1), и непораженных индивидуумов (KM1421; AB1190). Миобласты пациентов, страдающих DM1, несут 2600 повторов СТГ в 3'-НТО DMPK. Миобласты культивировали в ростовой среде для выращивания клеток скелетных мышц (доступной от PromoCell в Гейдельберге, Германия) с 2% лошадиной сыворотки (доступной от Gibco в Бристоле, Род-Айленд), 1% куриного эмбрионального экстракта (доступного от USB Corp в Кливленде, Огайо) и 0,5 мг/мл пенициллина/стрептомицина (Gibco). Для миогенной дифференцировки конфлюэнтные культуры переключали на

среду для дифференцировки DMEM, дополненную 2% лошадиной сыворотки, и культивировали в течение четырех дней.

**Таблица 14:** Соединения, протестированные в Примере 3

Идентифика- тор соединения	Последовательность EEV (от N к C)	Последователь- ность РМО (5'-3')	Модификации РМО	Химия конъюгации
Только РМО	Н/П	CAGCAGCAG CAGCAGCAG CAG (SEQ ID NO: 154)	5'-ОН	Н/П
197-777 (DM1-1)	Ac-PKKKRKV- Lys(цикло[Ff-Nal- RrRrQ])-PEG <sub>12</sub> -K(N3)- NH <sub>2</sub> Ac-(SEQ ID NO: 42)-- Lys(SEQ ID NO: 78)- PEG <sub>12</sub> -K(N3)-NH <sub>2</sub>	CAGCAGCAG CAGCAGCAG CAG (SEQ ID NO: 154)	5'- саркозинамид ; 3'-C4- циклооктин	Клик
221-1106 (DM1-2)	Ac-PKKKRKV- миниPEG2- Lys(цикло[FfFGRGRQ]) -миниPEG2-K(N3)-NH <sub>2</sub> Ac-(SEQ ID NO: 42)- миниPEG2-Lys(SEQ ID NO: 76)-миниPEG2- K(N3)NH <sub>2</sub>	CAGCAGCAG CAGCAGCAG CAG (SEQ ID NO: 154)	5'-ОН; 3'- вторичный амин морфолино	?Клик с линкером PEG4-COT

*Обработка.* Мышечные клетки пациентов, страдающих DM1, обрабатывали 10 мкМ, 3 мкМ, 1 мкМ или 0,3 мкМ соединений с применением двух различных условий обработки. В первом условии миобласты высевали при 75-80% конфлюентности, соединения серийно разводили в ростовой среде и клетки промывали в течение 24 часов для обеспечения свободного захвата соединения. Среду, содержащую соединение, удаляли, миобласты промывали 1X ДФСБ (Gibco) и дифференцировали в течение четырех дней перед сбором. Для второго условия, выполняемого параллельно, миобласты дифференцировали за три дня до обработки, соединения серийно разводили в среде для дифференцировки и мышечные трубочки собирали для анализа через 24 часа.

*Выделение РНК и ПЦР.* Общую РНК выделяли с помощью мининабора RNEASY (доступного от Qiagen в Германтауне, Мэриленд) в соответствии с инструкциями производителя. Для включения экзона 100 нг РНК подвергали обратной транскрипции и использовали для ПЦР (набор для ОТ-ПЦР OneStep, Qiagen). Образцы анализировали с помощью LabChip (доступного от PerkinElmer в Уолтеме, Мэриленд) с помощью набора для высокопроизводительного и высокочувствительного анализа ДНК.

**Результаты.** DMPK CUG-нацеленный РМО (не конъюгированный с EEV) обеспечивал улучшение ошибочного сплайсинга экзона 5 MBNL1 (данные не показаны). На **ФИГ. 11А-11F** показано смешанное восстановление дефекта сплайсинга MBNL1 (**ФИГ. 11А**) и его мишеней (SOS1, IR, DMD, BIN1, LDB3; **ФИГ. 11В-11F**) в мышечных клетках, полученных от пациентов, страдающих DM1, обработанных различными концентрациями DMPK CUG-нацеленного EEV-РМО (CUG<sup>exp</sup> 197-777 и CUG<sup>exp</sup> 221-1106). EEV-РМО 197-777 вызывал умеренную коррекцию событий ошибочного сплайсинга в мышечных клетках пациентов, страдающих DM1. MBNL1 и SOS1 показали лучший ответ в отношении коррекции ошибочного сплайсинга. EEV-РМО 197-777 выбрали в качестве инструментального соединения для последующих экспериментов, описанных ниже.

Миобласты и мышечные трубочки, полученные от пациентов, страдающих DM1, обрабатывали 10 мкМ, 3 мкМ и 1 мкМ DMPK CUG-нацеленного EEV-РМО 197-777 с использованием методов, аналогичных описанным выше. Оценивали восстановление событий альтернативного сплайсинга РНК MNBL1 и мишеней MNBL1. После обработки миобластов и мышечных трубочек EEV-РМО наблюдали коррекцию сплайсинга различной степени для MNBL1 (исключение экзона 5; **ФИГ. 12А**), SOS1 (включение экзона 25; **ФИГ. 12В**), INSR (включение экзона 11; **ФИГ. 12С**), DMD (включение экзона 78, **ФИГ. 12D**), BIN1 (включение экзона 11; **ФИГ. 12Е**) и LDB3 (исключение экзона 11; **ФИГ. 12F**).

На **ФИГ. 44А-44D** показана обратимость фенотипов миотонии у мышей HSA-LR, обработанных 20 мг/кг 221-1106, которую количественно определяли с помощью анализа мышечной релаксации. На **ФИГ. 44А** и **44С** показаны графики времени релаксации до 80% от пиковой изометрической силы и на **ФИГ. 44В** показаны исходные данные для кривых усилия. На **ФИГ. 44D** показано обращение фенотипов миотонии у мышей HSA-LR, обработанных 20 мг/кг 221-1106, количественно определенное с помощью типичных электромиографических (EMG) кривых.

#### **Пример 4. Оценка РМО-EEV 221-1120 в модели DM1 на мышцах**

Модель DM1 на мышцах создавали для исследования влияния EEV-РМО 221-1120 (также называемого EEV-РМО-DM1-3 или DM1-3; РМО 221=5'-CAG CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154; все мономеры РМО); EEV 1120=Ac-РKKKРKV-AEEA-Lys(цикло[FGFGRGRQ]-PEG12-OH (Ac-(SEQ ID NO: 42)-AEEA-Lys(SEQ ID NO: 82)-PEG12-OH)) на сплайсинг и уровни мРНК нисходящих генов в той же модели на

трансгенных мышцах HSA-LR, что и описанная в Примере 2. РМО и EEV конъюгировали с применением химии амидов.

**Экспериментальная часть.** Были две общие группы обработки: 1) мыши дикого типа и 2) мыши HSA-LR (модель заболевания DM1). В группе обработки HSA-LR было две подгруппы обработки: 1) HSA-LR, получавшие солевой раствор (контроль); и 2) HSA-LR+EEV-РМО 221-1120. Мышам вводили 15 мг/кг, 30 мг/кг, 60 мг/кг или 90 мг/кг (на основе РМО) РМО-EEV или солевой раствор путем внутривенной инъекции в хвост. Через семь дней после обработки мышей умерщвляли и ткани собирали для анализа.

*Анализ методом ОТ-ПЦР (корректировка сплайсинга).* Ткани гомогенизировали с помощью гомогенизатора OMNI BEAD MILL, и РНК экстрагировали с помощью QIACUBEQ. Анализ методом ОТ-ПЦР выполняли с использованием набора для одностадийной ОТ-ПЦР (Qiagen) в соответствии с протоколами производителя с 35 циклами ПЦР при 94°C в течение 30 секунд; 60°C в течение 30 секунд и 72°C в течение 30 секунд. Последовательность ген-специфичных праймеров приведена далее: включение экзона 7a Clcn1, прямой праймер=5'-TTCACATCGCCAGCATCTGTGC-3' (SEQ ID NO: 319), обратный праймер=5'-CACGGAACACAAAGGCACTGAATGT-3' (SEQ ID NO: 320); включение экзона 5 Mbn11, прямой праймер:=5'-GCTGCCCAATACCAGGTCAAC-3' (SEQ ID NO: 321), обратный праймер:=5'-TGGTGGGAGAAATGCTGTATGC-3' (SEQ ID NO: 322);

Включение экзона 22 Atp2a1, прямой праймер=5'-GCTCATGGTCCSTCAAGATCTCAC-3' (SEQ ID NO: 323), обратный праймер: 5'-GGGTCAGTGCSTCAGCTTTG-3' (SEQ ID NO: 324);

Включение экзона 7 Nfix, прямой праймер=5'-TCGACGACAGTGAGATGGAG-3' (SEQ ID NO: 325), обратный праймер 5' CAAACTCCTTCAGCGAGTCC-3' (SEQ ID NO: 326). Праймеры для Clcn1, Mbn11 и Atp2a1 были получены из Klein et al., The Journal of Clinical Investigation. 2019, 129 (11), pg. 4739; и праймеры для Nfix были получены из Chen et al., Scientific Reports. 2016, 6(1), pg. 1. Продукты кДНК разделяли на 2% агарозном Е-геле с использованием красителя SYBR SAFE. Процент включения экзона рассчитывали по соотношению непропущенная полоса/(непропущенная полоса+пропущенная полоса).

*Расчет индекса сплайсинга при DM1 мышши (mDSI).* mDSI рассчитывали в соответствии с протоколом, приведенном в литературном источнике Tanner et. al. (Nucleic acids research. 2021, 49 (4), pg. 2240-54). Для каждого образца  $i$  нормированные значения сплайсинга для каждого события сплайсинга  $j$  рассчитывали как  $(PSI_{i,j} - PSI_{\text{дикий тип},j}) / (PSI_{\text{HSA LR},j} - PSI_{\text{дикий тип},j})$ , где  $PSI_{\text{дикий тип},j}$  представляет собой среднее значение PSI для события  $j$  среди мышей дикого типа, и  $PSI_{\text{HSA LR},j}$  представляет собой среднее значение PSI для события  $j$  среди мышей HSA LR. Затем mDSI рассчитывают как среднее значение всех нормированных значений сплайсинга, которыми в исследованиях являются Atp2a1, Nfix, Mbn11 и Clcn1.

*Анализ методом кОТ-ПЦР (нокдаун мРНК HSA).* Обратную транскрипцию выполняли с использованием высокопроизводительного набора для обратной

транскрипции кДНК от Life Technologies Corporation в соответствии с протоколами производителя. Количественную ПЦР в режиме реального времени выполняли с использованием SyBr Green Supermix от Bio-Rad и аппарата QuantStudio3 для кПЦР с ген-специфичными праймерами: прямой праймер для мРНК HSA=5'-TTCCATCGTCCACCGCAAAT-3' (SEQ ID NO: 327), обратный праймер=5'-AGTTTACGATGGCAGCAACG-3' (SEQ ID NO; 328), оба праймера из Klein et al., The Journal of Clinical Investigation. 2019, 129 (11), pg. 4739; и прямой праймер для GAPDH мыши=5'-AGGTCGGTGTGAACGGATTTG-3' (SEQ ID NO: 329), обратный праймер=5'-TGTAGACCATGTAGTTGAGGTCA-3' (SEQ ID NO: 330).

*Секвенирование РНК.* Для профилирования транскриптома проводили полиА-РНК-секвенирование с применением секвенирования нового поколения. Z-оценку для каждого гена рассчитывали как (значение образца - среднее значение)/(стандартное отклонение). Анализ дифференциального сплайсинга проводили на данных RNAseq для расчета процента сплайсинга (PSI) отдельных экзонов для каждого гена. PSI представляет собой соотношение нормированного количества считываний, указывающего на включение элемента транскрипта, к общему нормированному количеству считываний для данного события (считывания включения и исключения). Например, если экзон включен в считывания 100% времени, PSI равен 1. Кроме того, если экзон исключен из считываний 100% времени, PSI равен 0.

Анализировали двадцать два представляющих интерес гена, которые, как известно, являются прогнозирующими для развития DM1. Кроме того, анализировали гены, исследованные в Wagner et al. (PLOS Gen 2016 (47)) и Tanner et al. (NAR 2021 (48), 4, 2240-2254). Экзоны мыши картировали по местоположению у человека. В некоторых случаях граница экзонов у мышей и/или в геноме человека была не полностью известна. Таким образом, данные анализировали с использованием различных границ. Верные границы подтверждали с использованием данных РНК-секвенирования.

*Анализ очагов РНК CUG.* Срезы передней большеберцовой мышцы окрашивали на очаги CUG (FISH, красный) и ядра (Hoechst, синий). Срезы передней большеберцовой мышцы визуализировали и количественно определяли число ядер, имеющих очаги РНК CUG.

## Результаты

*Нокдаун мРНК HSA.* Диафрагма экспрессировала только 5-10% мРНК HSA по сравнению с четырехглавой, передней большеберцовой и трехглавой мышцами (**ФИГ. 13А**). Обработка EEV-РМО, по-видимому, не изменяла уровни мРНК HSA в диафрагме (**ФИГ. 13В**). Уровень экспрессии повторов HSA 220 CUG может быть недостаточным для фенотипа ошибочного сплайсинга при DM1 в диафрагме.

EEV-РМО приводил к нокдауну мРНК HSA дозозависимым образом, что подтверждает взаимодействие с мишенью в ткани четырехглавой мышцы (**ФИГ. 14А**), икроножной мышцы (**ФИГ. 14В**), трехглавой мышцы (**ФИГ. 14С**) и передней большеберцовой мышцы (**ФИГ. 14D**). Кроме того, значение Ct (пороговое значение

цикла) мРНК HSA близко к уровню GAPDH мыши (~15), что позволяет предположить высокую экспрессию трансгена HSA у мышцей HSA-LR в четырехглавой мышце.

*mDSI (коррекция сплайсинга).* Индекс сплайсинга при DM1 у мышцей (mDSI) для четырехглавой, икроножной, трехглавой и передней большеберцовой мышцей показан на **ФИГ. 15А-15D**. Обработка EEV-PMO приводила к коррекции дефектов сплайсинга, имеющих значение при DM1 (Atp2a1 экзон 22, Nfix экзон 7, Clcn1 экзон 7а, Mbnl1 экзон 5) через 1 неделю после инъекции в четырехглавой мышце (**ФИГ. 15А**), икроножной мышце (**ФИГ. 15В**), трехглавой мышце (**ФИГ. 15С**) и передней большеберцовой мышце (**ФИГ. 15D**) дозозависимым образом, при этом более высокие дозы обеспечивали показатели, приближающиеся к дикому типу или эквивалентные ему (полная коррекция). Приблизительно 50%-60% нокдаун РНК скелетного актина человека у мышцей HSA-LR был достигнут при концентрациях лекарственного средства, которые обеспечивали почти полную коррекцию сплайсинга.

На **ФИГ. 16А-В** показаны изображения ткани передней большеберцовой мышцей мышцей HSA-LR (**ФИГ. 16А**) и мышцей HSA-LR, обработанных EEV-PMO (**ФИГ. 16В**), окрашенной на очаги CUG (красный) и ядра (синий). Качественная и количественная оценка (**ФИГ. 16С**) показала, что обработка EEV-PMO снижала количество ядер, имеющих очаги CUG.

*Воздействие лекарственного средства.* Воздействие лекарственного средства исследовали методом ЖХ-МС. На **ФИГ. 17А-17D** показан дозозависимый ответ для воздействия PMO-EEV в четырехглавой мышце (**ФИГ. 17А**), трехглавой мышце (**ФИГ. 17В**), сердце (**ФИГ. 17С**), икроножной мышце (**ФИГ. 17D**), передней большеберцовой мышце (ТА; **ФИГ. 17F**), печени (**ФИГ. 17I**) и почке (**ФИГ. 17J**). В диафрагме дозозависимый ответ не наблюдали (**ФИГ. 17G**). EEV-PMO не детектировался в головном мозге, за исключением уровней дозировки 60 мг/кг и 90 мг/кг. На **ФИГ. 17К** показано воздействие лекарственного средства в различных тканях при уровне дозировки 60 мг/кг.

*Ответ при миотонии:* Наблюдали дозозависимое снижение миотонии у мышцей HSA-LR через 7 дней после обработки EEV-PMO-DM1-3 при 15, 30, 60 и 90 мг/кг (**ФИГ. 18А**). Миотония, вероятно, уменьшается через неделю после обработки EEV-PMO-DM1-3. У мышцей HSA-LR, получивших однократную дозу 90 мг/кг EEV-PMO-DM1-3, не наблюдали очевидных признаков миотонии задних конечностей после индукции.

*Анализ данных РНК-секвенирования.* На **ФИГ. 19А-19D** показаны результаты анализа основных компонентов. Анализ основных компонентов может быть использован для выявления сходства между образцами на основе матрицы расстояния. Данный тип графика полезен для визуализации общего эффекта экспериментальных ковариат и пакетных эффектов. Ось x представляет собой направление, которое объясняет наибольшую дисперсию, а ось y - второе по значимости. Процент общей дисперсии на направление показан как PCA. Мыши дикого типа и HSA-LR находятся в отдельных группах. Экспрессия генов в икроножной мышце мышцей HSA-LR, обработанных PMO-EEV, была смещена к экспрессии генов мышцей дикого типа.

На **ФИГ. 20А** представлена тепловая карта, показывающая дифференциально экспрессируемые гены (по Z-оценке) из трех групп обработки: 1) мыши ДТ; 2) мыши HSA-LR и 3) мыши HSA-LR+EEV-PMO 221-1120 (60 мг/кг). В общей сложности 956 ( $p < 0,5$ ) генов дифференциально экспрессировались в группах обработки, что указывает на разницу между мышами дикого типа (ДТ) и мышами в модели заболевания (HSA-LR). Обработка EEV-PMO (HSA-LR (+,+)) приводила к общей коррекции экспрессии генов, сдвигаясь от профиля заболевания (красный цвет, HSA-LR (-,-)) и к профилю у мышей дикого типа (ДТ выше).

**ФИГ. 20В** представляет собой тепловую карту, используемую для визуализации профиля экспрессии 40 из 43 генов, для которых было обнаружено, что они имеют более 7 повторов СТГ на основании анализа BLAST. Три гена с повторами СТГ, обнаруженные в анализе Blast (Crb2, Hsd3b6 и Inhbe), не включали из-за низкого числа считываний. Этот анализ можно применять для выявления совместно регулируемых генов в условиях обработки.

На **ФИГ. 21** показан вулканический график общего изменения транскрипции в группе, обработанной EEV-PMO, и в группе HSA-LA. Каждая точка данных на графике рассеяния представляет ген. Кратное изменение каждого гена представлено по оси x, и  $\log_{10}$  его скорректированного p-значения представлен по оси y. Гены со скорректированным p-значением менее 0,05 и кратным изменением более 2 обозначены красными точками. Они представляют собой гены с положительной регуляцией. Гены со скорректированным p-значением менее 0,05 и кратным изменением менее -2 обозначены синими точками. Они представляют собой гены с отрицательной регуляцией. Была обнаружена значительная отрицательная регуляция трех генов (Txlnb, Scube2 и Greb1) и значительная положительная регуляция одного гена (Txlnb). Большинство транскриптов, содержащих по меньшей мере (CUG)<sub>7</sub>, не подвергались значительному влиянию. РСА-анализ этих генов показал, что для Scube2 (**ФИГ. 22А**), Greb1 (**ФИГ. 22В**), Ttc7 (**ФИГ. 22С**), Txlnb(CUG)<sub>9</sub> и Ndr3 (**ФИГ. 22Е**) наблюдали коррекцию при обработке EEV-PMO. Txlnb избыточно корректируются при обработке (**ФИГ. 22D**).

На **ФИГ. 23А-23D** показаны данные транскриптома для различных генов и различных групп обработки. В группе обработки HSA-LR+ EEV-PMO 221-1120 наблюдали коррекцию включения экзона 22 Atp2a1 (**ФИГ. 23А**), исключения экзона 7 Clcn1 (**ФИГ. 23В**), исключения экзона 7 Nfix (**ФИГ. 23С**) и исключения экзона 7 Mbnl1 (**ФИГ. 23D**).

На **ФИГ. 24** показан процент сплайсинга (PSI) отдельных экзонов для различных генов. Гены представляют собой биомаркеры сплайсинга, восприимчивого к MBNL-1 (например, нисходящие гены). Зависимые от MBNL-1 биомаркеры отбирали на основании динамического диапазона между группами дикого типа и заболевания, как описано в литературе. В группе обработки HSA-LR+EEV-PMO 221-1120 наблюдали коррекцию включения/исключения экзонов для всех 20 представляющих интерес генов, включая

Mbn1, Nfix, Atp2a1, Ldb3, Camk2g, Trim55, Fbox31, Slc8a3, Map3k4, Dctn4, Casna1s, Ryr1, Slain2, Phka1, Ppp3cc, Ttn, Neb, Irrfp2, Rapgef1 и Vsp39.

**Пример 5. Оценка РМО-ЕЕV 221-1120 в модели DM1 на мышах, второй модели DM1 на мышах**

РМО-ЕЕV DM1-3 (221-1120; см. последовательность в Примере 4) оценивали во второй модели DM1 на мышах с применением метода, аналогичного тому, что описан в Примере 4.

**Экспериментальная часть.** Мышам HSA-LR в возрасте семи недель внутривенно вводили однократную дозу 80 мг/кг ЕЕV-РМО-DM1-3 или дозу 20 мг/кг ЕЕV-РМО-DM1-3 один раз в две недели в течение шести недель (всего 80 мг/кг за 4 дозы), и ткани собирали через 1-12 недель после введения однократной дозы или через две недели после введения последней дозы. ОТ-ПЦР применяли для определения альтернативного сплайсинга для конкретных генов (Atp2a1, Clcn1, Nfix, MBNL1). К-ПЦР применяли для определения снижения уровня мРНК актина-HSA после обработки. ЖХ-МС использовали для определения уровня лекарственного средства в четырехглавой мышце, икроножной мышце, передней большеберцовой мышце, трехглавой мышце, диафрагме, сердце, почке, печени, головном мозге и плазме крови. Уменьшение миотонии регистрировали через 7 дней после обработки соединением ЕЕV-РМО-DM1-3.

**Результаты.** Тенденции, подобные таковым в Примере 4, наблюдали для восстановления сплайсинга для Atp2a1, Clcn1, Nfix и MBNL1 в ткани передней большеберцовой мышцы, икроножной мышцы, трехглавой мышцы и четырехглавой мышцы (данные не показаны). Кроме того, тенденции, подобные таковым в Примере 4, наблюдали для нокдауна мРНК HSA в ткани передней большеберцовой мышцы, икроножной мышцы, трехглавой мышцы и четырехглавой мышцы как через 1 неделю, так и через 4 недели после обработки (данные не показаны).

**ФИГ. 25A-25D** представляют собой графики, показывающие уменьшение уровня лекарственного средства при 80 мг/кг ЕЕV-РМО-DM1-3 через 1-8 недель в ткани передней большеберцовой мышцы (**ФИГ. 25A**), икроножной мышцы (**ФИГ. 25B**), трехглавой мышцы (**ФИГ. 25C**) и четырехглавой мышцы (**ФИГ. 25D**). ЕЕV-РМО-DM1-3 (60 мг/кг олигонуклеотида, 80 мг/кг полного лекарственного средства) полностью корректирует ошибочный сплайсинг в икроножной мышце, трехглавой мышце, передней большеберцовой мышце и четырехглавой мышце через 1 неделю после обработки. **ФИГ. 26A-26B** представляют собой графики, показывающие уменьшение уровней лекарственного средства, наблюдаемое при введении однократной дозы 80 мг/кг ЕЕV-РМО-DM1-3 через от 1 до 4 недель, 8 недель и 12 недель в печени. Относительно более высокое количество ЕЕV-РМО-DM1-3 в печени наблюдали через 2 недели после введения последней дозы в 6-недельном режиме дозирования, чем через 4 недели после введения однократной дозы.

На **ФИГ. 26C-26D** показано уменьшение уровней лекарственного средства, наблюдаемое при введении однократной дозы 80 мг/кг ЕЕV-РМО-DM1-3 через от 1 до 4



недель, 8 недель и 12 недель в почке. Относительно низкое количество EEV-PMO-DM1-3 в почке также наблюдали через 2 недели после введения последней дозы в 6-недельном режиме дозирования, чем через 4 недели после введения однократной дозы. Через 12 недель после введения однократной дозы EEV-PMO-DM1-3 лекарственное средство все еще присутствовало в почке, но не в печени.

Субъективные наблюдения за миотонией были сделаны и показаны в **Таблице 15**. При многодозовом режиме (Q2W) не наблюдали признаков восстановления миотонии через две недели после обработки. Наблюдался смешанный эффект в отношении миотонии у мышей, обработанных однократной дозой 80 мг/кг, который исчезал через 12 недель после обработки.

**Таблица 15.** Наблюдения за миотонией

Группа	До введения дозы		Через 1 неделю после введения дозы		Через 4 недели после введения дозы		Через 8 недель после введения дозы		Через 12 недель после введения дозы	
			F	M	F	M	F	M	F	M
Пол	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 мг/кг	++	++	+	+	+	+	0	+	++	++
221-1120 20 мг/кг 1 раз/2 недели	++	++	Н/П	Н/П	Н/П	Н/П	++	++	Н/П	Н/П

F=самка; M=самец; 0=ни у одной из мышей не наблюдали миотонии; +=смешанная миотония, у некоторых мышей наблюдали уменьшение миотонии; ++=у всех мышей наблюдали миотонию

Сходный эксперимент выполняли для оценки внутривенного введения EEV-PMO-DM1-3 в течение более длительного периода времени и в более высокой дозе. Мышам HSA-LR в возрасте восьми недель вводили 40 мг/кг, 60 мг/кг, 80 мг/кг или 120 мг/кг EEV-PMO-DM1-3 путем внутривенной инъекции, и ткани собирали через 4-12 недель. ОТ-ПЦР применяли для определения альтернативного сплайсинга для конкретных генов (*Atp2a1*, *Clcn1*, *Nfix*, *MBNL1*). Уменьшение миотонии регистрировали через 7 дней после обработки EEV-PMO-DM1-3. Тенденции, подобные таковым в Примере 4, наблюдали для восстановления сплайсинга для *Atp2a1*, *Clcn1*, *Nfix* и *MBNL1* в ткани передней большеберцовой мышцы и икроножной мышцы (данные не показаны).

Субъективные наблюдения за миотонией были сделаны и показаны в **Таблице 16**. У самок миотония была более выраженной, чем у самцов. Как у самцов, так и у самок мышей, которым вводили дозу 120 мг/кг, через 8 недель после обработки отсутствуют признаки миотонии.

Таблица 16. Наблюдения за миотонией

Группа	До введения дозы		Через 1 неделю после введения дозы		Через 4 недели после введения дозы		Через 8 недель после введения дозы		Через 12 недель после введения дозы		
			F	M	F	M	F	M	F	M	F
FVB	0	0	0	0	0	0	0	0	0	0	0
HSA-LR	++	++	++	++	++	++	++	++	++	++	++
221-1120 40 мг/кг	++	++	+	++	++	++	Н/П	Н/П	Н/П	Н/П	Н/П
221-1120 60 мг/кг	++	++	+	0	+	+	Н/П	Н/П	Н/П	Н/П	Н/П
221-1120 80 мг/кг	++	++	0	0	+	+	+	0	+	0	0
221-1120 120 мг/кг	++	++	0	0	0	0	0	0	0	0	0

F=самка; M=самец; 0=ни у одной из мышей не наблюдали миотонии; +=смешанная миотония, у некоторых мышей наблюдали уменьшение миотонии; ++=у всех мышей наблюдали миотонию

#### Пример 6. Обработка клеток, полученных от пациентов, страдающих DM1, EEV-PMO-DM1-3

PMO-EEV DM1-3 (221-1120; см. последовательность в Примере 4) оценивали в миобластах, полученных от пациентов, страдающих DM1.

**Экспериментальная часть.** Миобласты пациентов обрабатывали 30 мкМ DM1-3 в течение четырех дней дифференцировки. Коррекцию сплайсинга оценивали с помощью одностадийной ОТ-ПЦР и Labrchip (на графике представлено среднее значение  $\pm$  CO; n=4). Для детектирования очагов РНК CUG использовали HCR-FISH и анализы для детектирования секвестрированного белка MBNL1. Результаты: EEV-PMO-DM1-3 стимулировал значительную коррекцию сплайсинга биомаркеров и уменьшение ядерных очагов в мышечных клетках, полученных от пациентов, страдающих DM1.

**Результаты.** ФИГ. 27А-27С представляют собой графики, на которых показано, что EEV-PMO-DM1-3 стимулирует значительную коррекцию сплайсинга биомаркеров (MBLN1, SOS1 и NFIX) в мышечных клетках, полученных от пациентов, страдающих DM1. Кроме того, обработка DM1-3 приводила к уменьшению ядерных очагов в мышечных клетках, полученных от пациентов, страдающих DM1 (ФИГ. 28А-28С).

#### Пример 7. Скрининг цитотоксичности EEV-PMO-DM1-3 в клетках почки

PMO-EEV DM1-3 (221-1120; см. последовательность в Примере 4) оценивали в клетках почки человека.

**Экспериментальная часть.** Первичные эпителиальные клетки проксимальных канальцев почки человека (RPTEC) подвергали воздействию различных концентраций

(серийное разведение 1:2 в солевом растворе с конечным коэффициентом разведения 4х от примерно 6 мкМ до примерно 800 мкМ) РМО-DM1 и EEV-РМО-DM1-3 в течение 24 часов и проверяли на жизнеспособность с использованием люминесцентного анализа жизнеспособности CELLTITER-GLO. Мелиттин использовали в качестве положительного контроля при 16,6 мкМ.

**Результаты.** На **ФИГ. 29А-29В** показано, что РМО-DM1 или его конъюгированный вариант EEV-РМО-DM1-3 не проявляли какой-либо токсичности даже при наивысшей концентрации 817 мкМ или 797 мкМ, соответственно.

**Пример 8. Оценка способности РМО-EEV 221-1113 корректировать события ошибочного сплайсинга и последующего сплайсинга в иммортализованных клетках пациентов, страдающих DM1, и клетках HeLa-480**

Иммортализованные мышечные клетки, полученные от пациентов, страдающих DM1 (2600 повторов CUG), и HeLa-480 (линия клеток модели DM1, см. Пример 1), обрабатывали конструкцией EEV-РМО 221-1113 и анализировали в отношении коррекции aberrантного сплайсинга и количественного определения очагов. EEV 1113 представляет собой Ас-РКККРКV-миниPEG-K(цикло(Ff-Nal-GrGrQ))-PEG12-ОН (Ас-(SEQ ID NO: 42)-миниPEG-K(цикло(SEQ ID NO: 80)-PEG12-ОН). EEV-РМО 221-1113 представляет собой EEV 1113, конъюгированный с последовательностью 221 РМО (5'-CAG CAG CAG CAG CAG CAG CAG-3' (SEQ ID NO: 154; все мономеры РМО) посредством химии амидных связей.

**Экспериментальная часть.** Использовали методы, подобные описанным в Примере 1 и Примере 3.

**Результаты.**

*Анализ очагов РНК CUG.* Клетки окрашивали на ядра (Hoeschet, синий) и на очаги РНК с повторами CUG (зеленый) и визуализировали. Уменьшение очагов РНК CUG наблюдали между необработанными клетками от пациента, страдающего DM1, и обработанными EEV-РМО клетками DM1 (**ФИГ. 30А-30С**). Аналогичным образом, уменьшение очагов РНК CUG наблюдали между необработанными клетками HeLa-480 и клетками HeLa-480, обработанными EEV-РМО (**ФИГ. 31А-31В**).

*Коррекция последующего сплайсинга.* Обработанные РМО-EEV клетки, полученные от пациентов, страдающих DM1, и клетки HeLa-480 анализировали в отношении процента включения экзона 5 для MBLN1, процента включения экзона 25 для SOS1 и процента включения экзона 7 для NFIX. Обработка EEV-РМО приводила к восстановлению событий сплайсинга для Mbnl1 (**ФИГ. 32А**), Sos1 (**ФИГ. 32В**) и NFIX (**ФИГ. 32С**).

Кроме того, в клетках HeLa-480, обработанных EEV-РМО, наблюдали дозозависимую коррекцию сплайсинга MBNL1 (**ФИГ. 33А**) и последующего ошибочного сплайсинга SOS1 (**ФИГ. 33В**), CLASP1 (**ФИГ. 33С**), NFIX (**ФИГ. 33Д**) и INSR (**ФИГ. 33Е**) дозозависимым образом.

**Пример 9. Оценка EEV-РМО 221-1106 во второй модели DM1 на мышах**

Модель DM1 на мышах создавали для исследования влияния EEV-PMO 221-1106 на сплайсинг и уровни мРНК нисходящих генов.

**Экспериментальная часть.** В качестве модели заболевания DM1 использовали трансгенных мышей с длинным повтором в гене скелетного актина человека (HSA-LR). Использовали методы, аналогичные описанным в Примере 5.

*Результаты.* На **ФИГ. 34А-34D** показана дозозависимая коррекция включения экзона 22 в Atp2a1 (**ФИГ. 34А**), экзона 7 в Nfix (**ФИГ. 34В**), экзона 7А в Clcn1 (**ФИГ. 34С**) и Mbnl1 (**ФИГ. 34D**) в икроножной мышце мышей, обработанных различными концентрациями EEV-PMO 221-1106. Обработка только PMO 221 не приводила к коррекции сплайсинга.

#### **Пример 10. Модель DM1 на мышах для исследования влияния повторов CUG разной длины в PMO**

Модель DM1 на мышах создавали для исследования влияния PMO-EEV 221-1121 (PMO имеет 7 повторов CAG, 21-мер) и PMO-EEV 0325-1121 (PMO имеет 8 повторов CAG, 24-мер) на сплайсинг и уровни мРНК нисходящих генов. PMO-EEV 221-1121 представляет собой PMO 221 (5'-CAG CAG CAG CAG CAG CAG CAG-3'; SEQ ID NO: 154; все мономеры PMO), конъюгированный с EEV 1121 (Ac-PKKKRKV-миниPEG2-Lys(цикло[GfFGGrGrQ])-PEG12-OH; Ac-(SEQ ID NO: 42)-миниPEG2-Lys(SEQ ID NO:74)-PEG12-OH) посредством химии амидов. PMO-EEV 0325-1121 представляет собой PMO 0325 (5'-CAG CAG CAG CAG CAG CAG CAG-CAG-3'; SEQ ID NO: 155; все мономеры PMO), конъюгированный с EEV 1121 посредством химии амидов.

**Экспериментальная часть.** В качестве модели заболевания DM1 использовали трансгенных мышей с длинным повтором в гене скелетного актина человека (HSA-LR). В общих чертах, мышам HSA-LR вводили 20 мг/кг, 40 мг/кг или 60 мг/кг либо 0221-1121, либо 0325-1121 путем внутривенной инъекции в хвостовую вену. Через неделю после инъекции мышей умерщвляли и собирали ткань. Использовали другие экспериментальные методы, аналогичные описанным в Примере 5.

**Результаты.** 0221-1121 (21-мер) был более эффективным в коррекции сплайсинга экзонов в Mbnl1 (**ФИГ. 35А**), Nfix (**ФИГ. 35В**) и Atp2a1 (**ФИГ. 35С**), чем 0325-1121 (24-мер), в ткани передней большеберцовой мышцы. Этот результат был неожиданным. Ожидалось, что 24-мер будет более эффективным, поскольку он будет иметь более высокую эффективность гибридизации и более высокую температуру плавления. В ткани икроножной мышцы различия были менее выраженными, как показано на **ФИГ. 36А-36С**. Субъективные наблюдения за миотонией осуществляли у самцов мышей (**Таблица 17**). Смешанную миотонию наблюдали через 1 неделю после обработки для 21-мер при 40 мг/кг, что близко результатам в Таблице 11.

**Таблица 17:** Наблюдения за миотонией

Группа	До введения дозы	Через 1 неделю после введения дозы	Через 4 недели после введения дозы

Пол	М	М	М
FVB	0	0	Н/П
HSA-LR	++	++	Н/П
221-112 20 мг/кг	++	++	Н/П
221-1120 40 мг/кг	++	+	
325-1120 20 мг/кг	++	++	Н/П
325-1120 40 мг/кг	++	++	

М=самец; 0=ни у одной из мышей не наблюдали миотонии; +=смешанная миотония, у некоторых мышей наблюдали уменьшение миотонии; ++=у всех мышей наблюдали миотонию

#### **Пример 11. Фармакокинетические исследования EEV-PMO 221-1120 на мышях CD1**

Модель на мышях CD1 использовали для исследования воздействия лекарственного средства (AUC) в плазме, почке и передней большеберцовой мышце для конструкции EEV-PMO 221-1120 (последовательность см. в Примере 4), также измеряли PMO-0221a, основной метаболит 221-1120 (см. **ФИГ. 37**).

**Экспериментальная часть.** Мышам CD1 в возрасте от пяти до семи недель вводили 80 мг/кг конструкции EEV-PMO 221-1120 путем внутривенной инъекции. У мышей брали образцы крови, и/или умерщвляли в различные моменты времени.

**Результаты.** В **Таблице 18**, **Таблице 19** и **Таблице 20** показаны фармакокинетические свойства, наблюдаемые в плазме, почке и передней большеберцовой мышце, соответственно. Для таблиц:  $AUC_{last}$ =площадь под кривой от нуля до последней поддающейся количественному определению концентрации; Д=доза;  $C_{max}$ =максимальная концентрация в сыворотке или плазме крови;  $T_{max}$ =время до достижения  $C_{max}$ ;  $C_L$ =общий клиренс из плазмы, сыворотки или крови;  $t_{1/2}$ =период полувыведения;  $V_{ss}$ =кажущийся объем распределения в равновесном состоянии;  $Q_h$ =печеночный кровоток (мл/мин/кг).

Значения AUC для метаболита в ~1000 раз ниже в передней большеберцовой мышце по сравнению с почкой. Средние значения времени удержания метаболита (MRT) в плазме могут быть непосредственно связаны со значениями MRT в тканях в результате перемещения из тканей в плазму до выведения с мочой.

**Таблица 18:** Фармакокинетические свойства в плазме

	221-1120	PMO-0221a
$AUC_{last}(HM^*ч)$	9290	953

AUC <sub>last</sub> /Д	1121	115
C <sub>max</sub> (нМ)	16217	12
C <sub>max</sub> /Д	1956	1,4
T <sub>max</sub> (ч)	0,1	24
CL (мл/мин/кг)	15	-
Q <sub>h</sub> (%)	16	-
t <sub>1/2</sub> (ч)	19	68
MRT <sub>last</sub> (ч)	1,2	60,0
V <sub>ss</sub> (мл/кг)	1325	-

**Таблица 19: Фармакокинетические свойства в почках**

	221-1120	РМО-0221a
AUC <sub>last</sub> (пмоль/г*ч)	51609	10113709
AUC <sub>last</sub> /Д	6225	1219865
C <sub>max</sub> (пмоль/г)	8754	105053
C <sub>max</sub> /Д	1056	12671
T <sub>max</sub> (ч)	4	24
t <sub>1/2</sub> (ч)	5	87
MRT <sub>last</sub> (ч)	8	68

**Таблица 20: Фармакокинетические свойства в передней большеберцовой мышце**

	221-1120	РМО-0221a
AUC <sub>last</sub> (пмоль/г*ч)	-	9140
AUC <sub>last</sub> /Д	-	1102
C <sub>max</sub> (пмоль/г)	34	162
C <sub>max</sub> /Д	4	20
T <sub>max</sub> (ч)	4	24
t <sub>1/2</sub> (ч)	-	48
MRT <sub>last</sub> (ч)	-	50

**Пример 12. Оценка РМО-ЕЕV 221-1120 в третьей модели DM1 на мышах**

Исследование в модели DM1 на мышах, сходное с Примерами 5 и 9, проводили для оценки влияния различных доз РМО-ЕЕV 221-220 (последовательность представлена в Примере 4) у мышей HSA-LR.

**Экспериментальная часть.** Мышам HSA-LR в возрасте восьми недель вводили 40, 60, 80 или 120 мг/кг РМО-ЕЕV 221-1120 внутривенно, и ткани собирали через 4-12 недель. ОТ-ПЦР применяли для определения альтернативного сплайсинга для конкретных генов (Atp2a1, Clcn1, Nfix, MBNL1). ЖХ-МС использовали для определения уровня

лекарственного средства в четырехглавой мышце, икроножной мышце, передней большеберцовой мышце, трехглавой мышце, диафрагме, сердце, почке, печени, головном мозге, плазме. РНК-секвенирование использовали для определения изменения уровня транскрипции между моделью заболевания с лечением, моделью заболевания без лечения и диким типом. К-ПЦР применяли для определения снижения уровня мРНК актина-HSA после обработки.

**Результаты.** Флуоресцентную визуализацию использовали для определения уменьшения очагов РНК после обработки соединением EEV-олиго (данные не показаны). Уменьшение миотонии регистрировали через 7 дней после обработки соединением EEV-олиго (данные не показаны). Результаты этих экспериментов показывают тенденции, сходные с таковыми у мышей, обработанных РМО-EEV 221-1120, в Примере 5 и Примере 14.

Коррекцию сплайсинга MBNL1, NFIX и ATP2A1 наблюдали в передней большеберцовой мышце (**ФИГ. 38А, 39А, 40А**) и икроножной мышце (**ФИГ. 38В, 39В, 40В**) при различных дозах EEV-РМО. Коррекцию сплайсинга MBNL1, NFIX и ATP2A1 наблюдали как в передней большеберцовой мышце, так и в икроножной мышце через 12 недель после обработки 120 мг/кг EEV-РМО.

#### **Пример 13. Оценка EEV-РМО 221-1120 в клетках HeLa480**

Клетки HeLa480 обрабатывали различными концентрациями EEV-РМО 221-1120 (последовательность представлена в Примере 4) и анализировали на очаги повторов CUG, селективное уменьшение r(CUG) и последующую коррекцию сплайсинга MBNL1 и SOS1.

**Экспериментальная часть.** Клетки HeLa480 конструировали, как описано в предыдущих примерах. ОТ-ПЦР и окрашивание очагов выполняли аналогично другим примерам, описанным в настоящем документе.

**Результаты.** На **ФИГ. 41А-41В** показаны примеры изображений контрольных клеток (необработанных) и клеток, обработанных 5 мкМ, 10 мкМ, 20 мкМ, 50 мкМ и 100 мкМ EEV-РМО 221-1120. Наблюдается уменьшение очагов CUG (зеленый) в обработанной группе HeLa480 по сравнению с необработанными клетками HeLa480. **ФИГ. 41В** представляет собой график количественного определения очагов на ядерную область. На **ФИГ. 41А-41В** показано, что EEV-РМО 221-1120 может уменьшать ядерные очаги РНК CUG. При дозе 5 мкМ наблюдали почти полное уменьшение.

На **ФИГ. 42А-42В** показано, что обработка EEV-РМО 221-1120 может привести к селективному нокдауну транскрипта DMPK, содержащего экспансию повтора, в клеточной линии HeLa480.

На **ФИГ. 42С-42D** показано, что обработка EEV-РМО 221-1120 приводила к коррекции сплайсинга MBNL1 (**ФИГ. 42С**) и SOCS1 (**ФИГ. 42D**) дозозависимым образом.

Описано несколько вариантов реализации настоящего изобретения. Тем не менее, следует понимать, что различные модификации могут быть сделаны без отклонения от

сущности и объема настоящего изобретения. Соответственно, другие варианты реализации включены в объем следующей формулы изобретения.



## ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Соединение, содержащее:

циклический пептид, содержащий от 6 до 12 аминокислот, причем по меньшей мере две аминокислоты указанного циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты указанного циклического пептида представляют собой ароматические гидрофобные аминокислоты и по меньшей мере две аминокислоты указанного циклического пептида представляют собой незаряженные неароматические аминокислоты; и

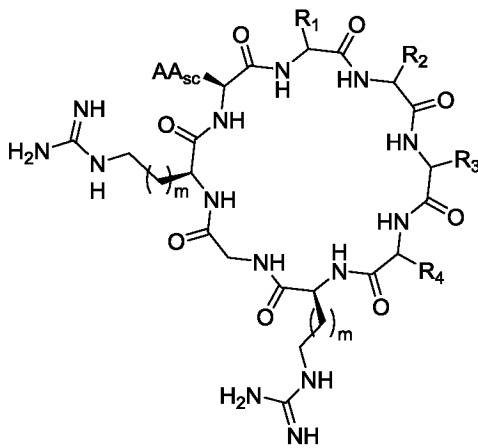
антисмысловое соединение (АС), которое комплементарно по меньшей мере части экспансированного повтора CUG в последовательности целевой мРНК, причем указанный АС содержит фосфордиамидат-морфолиновый (РМО) нуклеотид.

2. Соединение по п. 1, отличающееся тем, что по меньшей мере две заряженные аминокислоты указанного циклического пептида представляют собой аргинин.

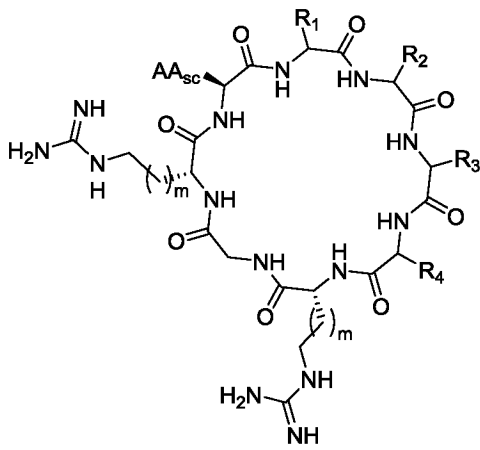
3. Соединение по п. 1 или п. 2, отличающееся тем, что по меньшей мере две ароматические гидрофобные аминокислоты указанного циклического пептида представляют собой фенилаланин, нафтилаланин или их комбинации.

4. Соединение по любому из пп. 1-3, отличающееся тем, что по меньшей мере две незаряженные неароматические аминокислоты представляют собой цитруллин, глицин или их комбинации.

5. Соединение по п. 1, отличающееся тем, что указанный циклический пептид имеет одну из следующих структур:



(I-a);



(I-b); или

ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

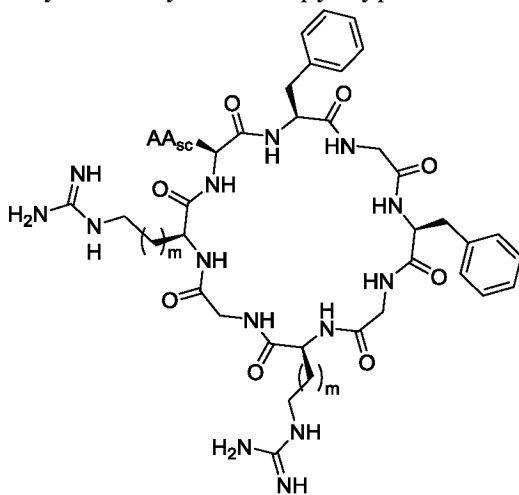
по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  представляет собой H или боковую цепь аминокислоты;

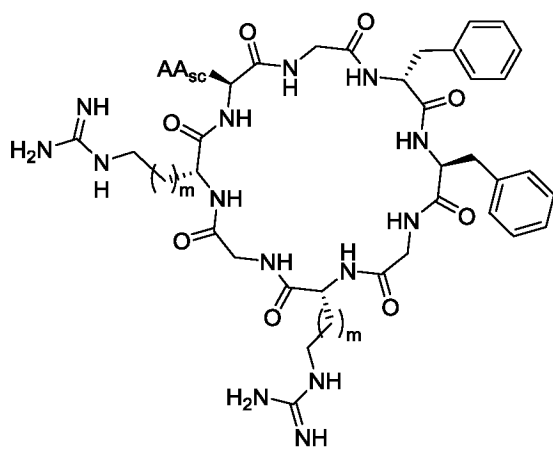
$AA_{sc}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение; и

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3.

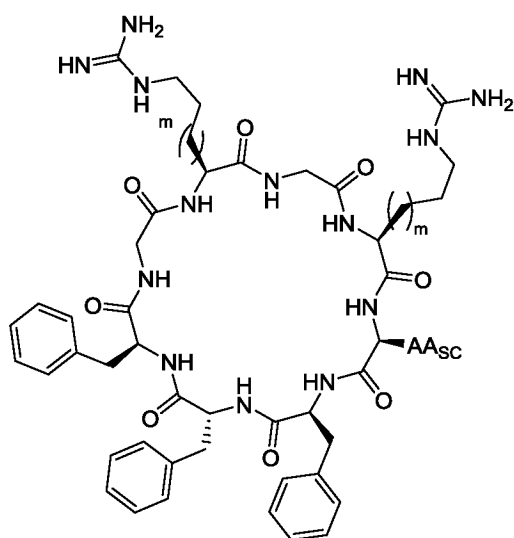
6. Соединение по п. 5, отличающееся тем, что указанный циклический пептид имеет одну из следующих структур:



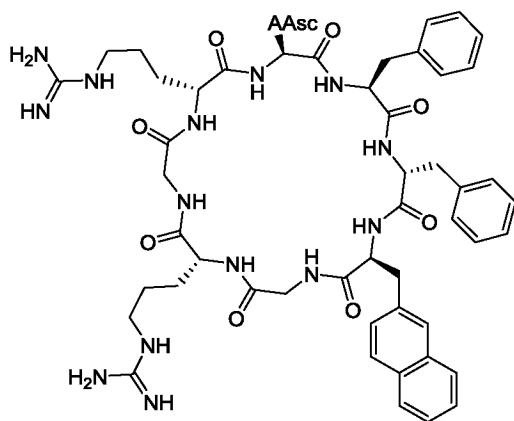
(I-1);



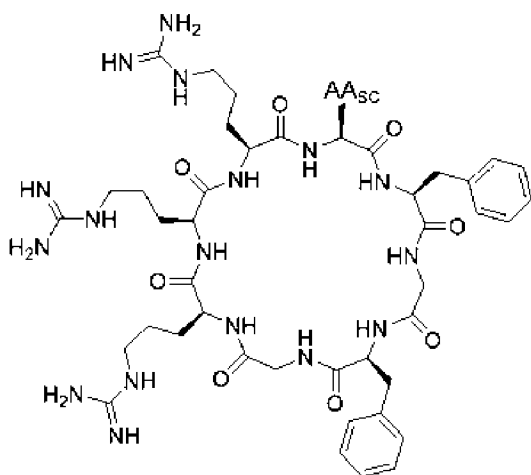
(I-2);



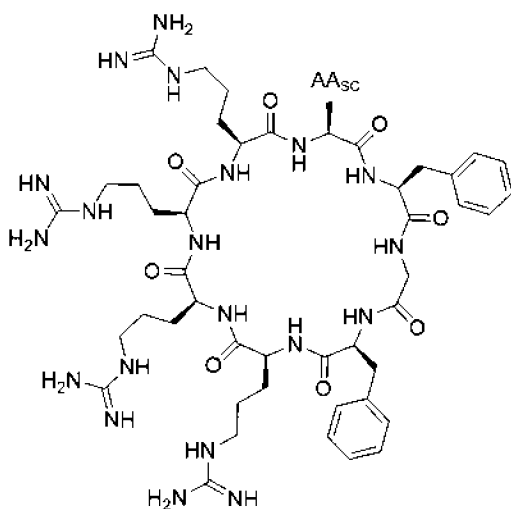
(I-3);



(I-4);



(I-5);



(I-6); или

ее протонированную форму.

7. Соединение по п. 5 или п. 6, отличающееся тем, что AA<sub>SC</sub> представляет собой боковую цепь остатка аспарагина, остатка аспарагиновой кислоты, остатка глутаминовой кислоты, остатка гомоглутаминовой кислоты или остатка гомоглутамата.

8. Соединение по п. 5 или 6, отличающееся тем, что каждый R<sub>1</sub>, R<sub>2</sub> и R<sub>3</sub> независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из R<sub>1</sub>, R<sub>2</sub> и R<sub>3</sub> представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

R<sub>4</sub> представляет собой H или боковую цепь аминокислоты;

m представляет собой 2, и

AA<sub>SC</sub> представляет собой боковую цепь остатка глутаминовой кислоты.

9. Соединение по п. 5 или п. 6, отличающееся тем, что AA<sub>SC</sub> представляет собой:



или

, где t представляет собой целое число от 0 до 5.

10. Соединение по любому из пп. 1-9, дополнительно содержащее линкер, отличающееся тем, что с помощью указанного линкера указанное бессмысленное соединение конъюгировано с указанным AA<sub>SC</sub>.

11. Соединение по п. 10, отличающееся тем, что указанный линкер содержит субъединицу  $-(\text{OCH}_2\text{CH}_2)_{z'}$ , где  $z'$  представляет собой целое число от 1 до 23.

12. Соединение по п. 10, отличающееся тем, что указанный линкер содержит:

(i) субъединицу  $-(\text{OCH}_2\text{CH}_2)_{z'}$ , где  $z'$  представляет собой целое число от 1 до 23;

(ii) один или более аминокислотных остатков, таких как остаток глицина,  $\beta$ -аланина, 4-аминомасляной кислоты, 5-аминопентановой кислоты или 6-аминогексановой кислоты или их комбинации; или

(iii) комбинации (i) и (ii).

13. Соединение по п. 10, отличающееся тем, что указанный линкер содержит:

(i) субъединицу  $-(\text{OCH}_2\text{CH}_2)_z$ , где  $z$  представляет собой целое число от 2 до 20;

(ii) один или более остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты, 5-аминопентановой кислоты, 6-аминогексановой кислоты или их комбинации; или

(iii) комбинации (i) и (ii).

14. Соединение, содержащее:

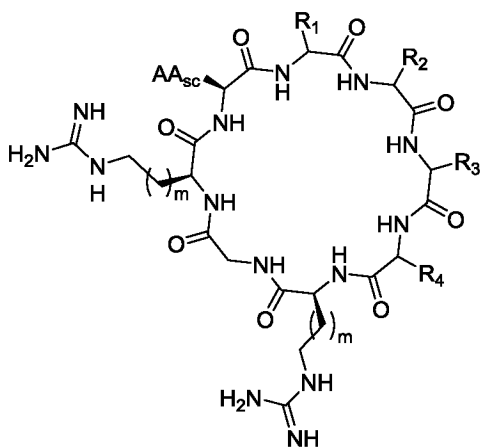
носитель для ускользания из эндосом, содержащий циклический пептид и экзоциклический пептид, отличающееся тем, что указанный циклический пептид содержит от 6 до 12 аминокислот, и указанный экзоциклический пептид содержит от 2 до 10 аминокислот; и

антисмысловое соединение (АС), которое комплементарно по меньшей мере части экспансированного повтора CUG в последовательности целевой мРНК, причем указанный АС содержит фосфордиамидат-морфолиновый (РМО) нуклеотид.

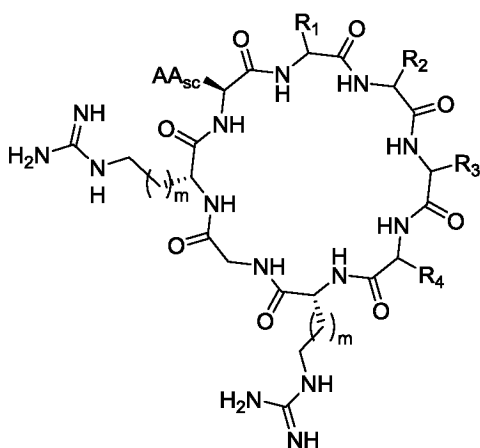
15. Соединение по п. 14, отличающееся тем, что по меньшей мере две аминокислоты указанного циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты указанного циклического пептида представляют собой ароматические гидрофобные аминокислоты и по меньшей мере две аминокислоты указанного циклического пептида представляют собой незаряженные неароматические аминокислоты.

16. Соединение по п. 15, отличающееся тем, что по меньшей мере две заряженные аминокислоты указанных циклических пептидов представляют собой аргинин, по меньшей мере две ароматические гидрофобные аминокислоты указанного циклического пептида представляют собой фенилаланин, нафтилаланин или их комбинации, и по меньшей мере две незаряженные неароматические аминокислоты представляют собой цитруллин, глицин или их комбинации.

17. Соединение по п. 14, отличающееся тем, что указанный циклический пептид имеет одну из следующих структур:



(I-a);



(I-b); или

ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

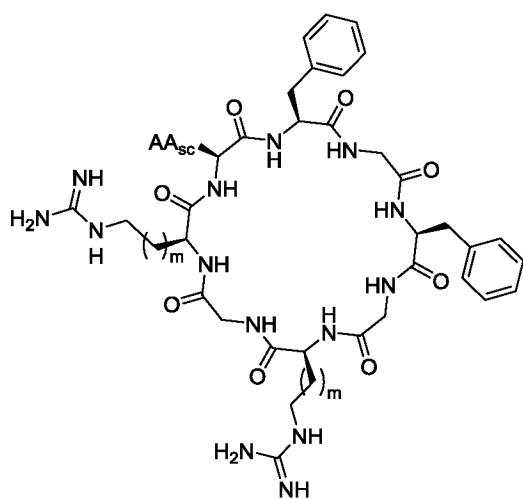
по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  представляет собой H или боковую цепь аминокислоты;

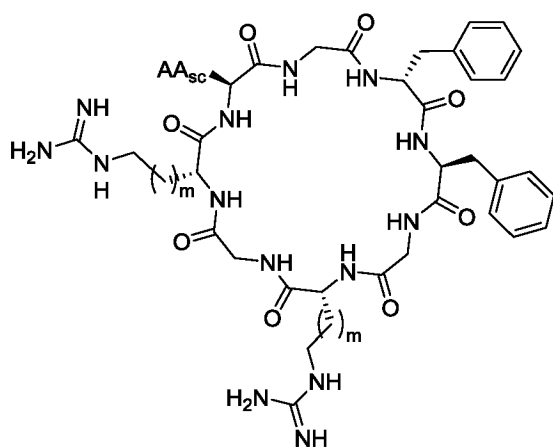
$AA_{SC}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение и указанный экзоциклический пептид; и

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3.

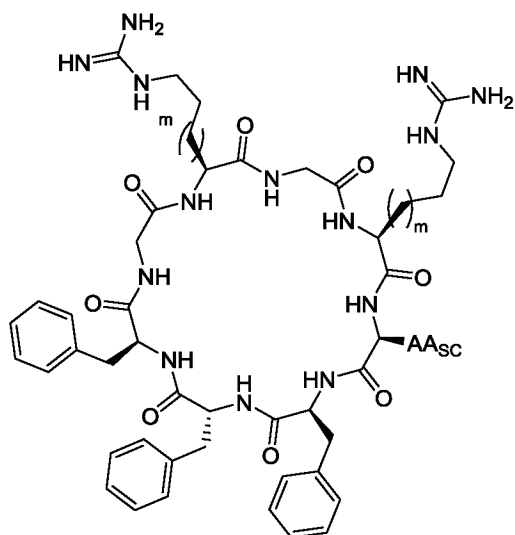
18. Соединение по п. 17, отличающееся тем, что указанный циклический пептид имеет одну из следующих структур:



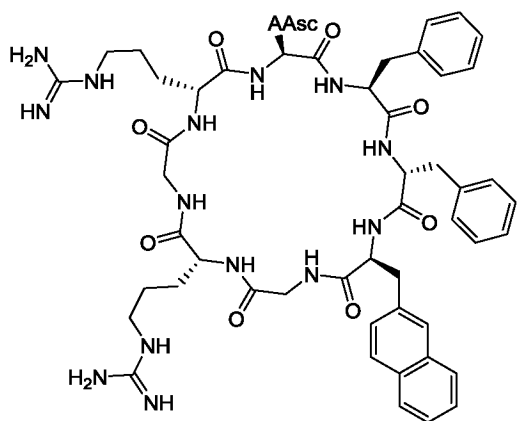
(I-1);



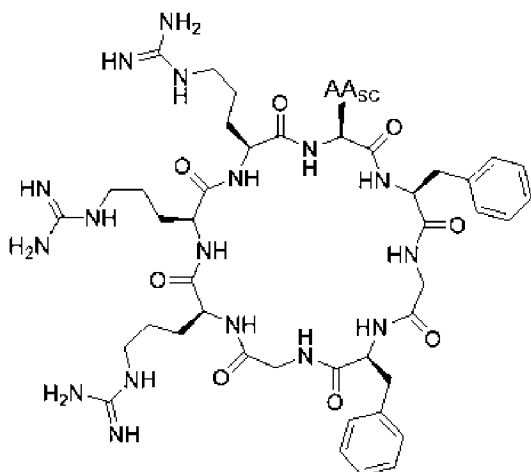
(I-2);



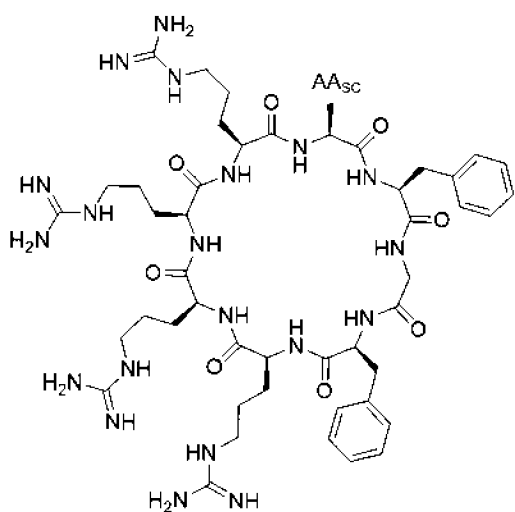
(I-3);



(I-4);



(I-5);



(I-6); или

ее протонированную форму.

19. Соединение по п. 17 или п. 18, отличающееся тем, что AA<sub>SC</sub> представляет собой боковую цепь остатка аспарагина, остатка аспарагиновой кислоты, остатка глутаминовой кислоты, остатка гомоглутаминовой кислоты или остатка гомоглутамата.

20. Соединение по п. 17 или п. 18, отличающееся тем, что AA<sub>SC</sub> представляет собой боковую цепь остатка глутаминовой кислоты.

21. Соединение по п. 17 или 18, отличающееся тем, что каждый R<sub>1</sub>, R<sub>2</sub> и R<sub>3</sub> независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;




по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$  представляет собой H или боковую цепь аминокислоты;

$m$  представляет собой 2, и

$AA_{SC}$  представляет собой боковую цепь остатка глутаминовой кислоты.

22. Соединение по любому из пп. 17-21, отличающееся тем, что  $AA_{SC}$  представляет

собой:  , где  $t$  представляет собой целое число от 0 до 5.

23. Соединение по любому из пп. 14-22, отличающееся тем, что указанный экзоциклический пептид содержит от 4 до 8 аминокислотных остатков.

24. Соединение по любому из пп. 14-23, отличающееся тем, что указанный экзоциклический пептид содержит 1 или 2 аминокислотных остатка, содержащих боковую цепь, содержащую гуанидиновую группу, или ее протонированную форму или соль.

25. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит 2, 3 или 4 остатка лизина.

26. Соединение по п. 25, отличающееся тем, что указанная аминокислотная группа на боковой цепи каждого остатка лизина замещена трифторацетильной ( $-COCF_3$ ), аллилоксикарбонильной (Alloc), 1-(4,4-диметил-2,6-диоксоциклогексиден)этильной (Dde) или (4,4-диметил-2,6-диоксоциклогекс-1-илиден-3)-метилбутильной (ivDde) группой.

27. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит по меньшей мере 2 аминокислотных остатка с гидрофобной боковой цепью.

28. Соединение по п. 27, отличающееся тем, что указанный аминокислотный остаток с гидрофобной боковой цепью выбран из валина, пролина, аланина, лейцина, изолейцина и метионина.

29. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит одну из следующих последовательностей: KK, KR, RR, NN, NK, NR, RN, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKN, KHK, NKK, HRR, HRN, NHR, NBH, NNN, NNNN, KNKK, KKNK, KKKN, KNKN, KNKK, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, NBHNB, NBKBN, RRRRR, KKKKK, KKKRK, RKKKK, KRKKK, KKKKR, KBKKB, RKKKKG, KRKKKG, KKRKKG, KKKKRG, RKKKKB, KRKKKB, KKRKKB, KKKKRB, KKKRKV, RRRRRR, NNNNNN, RHRHRN, HRHRHR, KRKRKR, RKRKRK, RBRBRB, KBKKB, PKKRKV, PGKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV или PKKKRKG, где B представляет собой бета-аланин.

30. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит одну из следующих последовательностей: PKKKRKV, RR, RRR, RHR, RBR, RBRBR, RBHBR или HBRBN, где B представляет собой бета-

аланин.

31. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит одну из следующих последовательностей: KK, KR, RR, KKK, KGK, KBK, KBR, KRK, KRR, RKK, RRR, KKKK, KKRK, KRKK, KRRK, RKKR, RRRR, KGKK, KKGK, KKKKK, KKKRK, KBKVK, KKKRKV, PKKKRKV, PGKKRKV, PKGKRKV, PKKGRKV, PKKKGKV, PKKKRGV или PKKKRKG.

32. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит PKKKRKV.

33. Соединение по любому из пп. 14-24, отличающееся тем, что указанный экзоциклический пептид содержит одну из следующих последовательностей: NLSKRPAAIKKAGQAKKKK, PAAKRVKLD, RQRRNELKRSF, RMRKFKNKGKDTAELRRRRVEVSVELR, KAKKDEQILKRRNV, VSRKRPRP, PPKKARED, PQPKKKPL, SALIKKKKKMAP, DRLRR, PKQKKRK, RKLKKKIKKL, REKKKFLKRR, KRKGDEVDGVDEVAKKKSKK или RKCLQAGMNLEARKTKK.

34. Соединение по любому из пп. 14-33, дополнительно содержащее линкер, с помощью которого указанное антисмысловое соединение и указанный экзоциклический пептид конъюгированы с указанным AA<sub>SC</sub>.

35. Соединение по п. 34, отличающееся тем, что указанный линкер содержит:

(i) субъединицу  $-(OCH_2CH_2)_{z'}$ , где  $z'$  представляет собой целое число от 1 до 23;

(ii) один или более аминокислотных остатков, таких как остаток глицина,  $\beta$ -аланина, 4-аминомасляной кислоты, 5-аминопентановой кислоты или 6-аминогексановой кислоты или их комбинации; или

(iii) комбинации (i) и (ii).

36. Соединение по п. 34, отличающееся тем, что указанный линкер содержит:

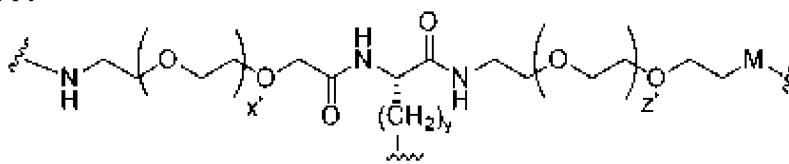
(i) субъединицу  $-(OCH_2CH_2)_z$ , где  $z$  представляет собой целое число от 2 до 20;

(ii) один или более остатков глицина,  $\beta$ -аланина, 4-аминомасляной кислоты, 5-аминопентановой кислоты, 6-аминогексановой кислоты или их комбинации; или

(iii) комбинации (i) и (ii).

37. Соединение по п. 34, отличающееся тем, что указанный линкер содержит двухвалентный или трехвалентный C<sub>1</sub>-C<sub>50</sub> алкилен, в котором 1-25 метиленовых групп необязательно и независимо заменены на -N(H)-, -N(C<sub>1</sub>-C<sub>4</sub> алкил)-, -N(циклоалкил)-, -O-, -C(O)-, -C(O)O-, -S-, -S(O)-, -S(O)<sub>2</sub>-, -S(O)<sub>2</sub>N(C<sub>1</sub>-C<sub>4</sub> алкил)-, -S(O)<sub>2</sub>N(циклоалкил)-, -N(H)C(O)-, -N(C<sub>1</sub>-C<sub>4</sub> алкил)C(O)-, -N(циклоалкил)C(O)-, -C(O)N(H)-, -C(O)N(C<sub>1</sub>-C<sub>4</sub> алкил), -C(O)N(циклоалкил), арил, гетероарил, циклоалкил или циклоалкенил.

38. Соединение по п. 34, отличающееся тем, что указанный линкер имеет структуру:



где:

$x'$  представляет собой целое число от 1 до 23;  $y$  представляет собой целое число от 1 до 5;  $z'$  представляет собой целое число от 1 до 23; \* представляет собой положение присоединения к аминокислотной боковой цепи аминокислотного остатка указанного циклического пептида; и  $M$  представляет собой связывающую группу.

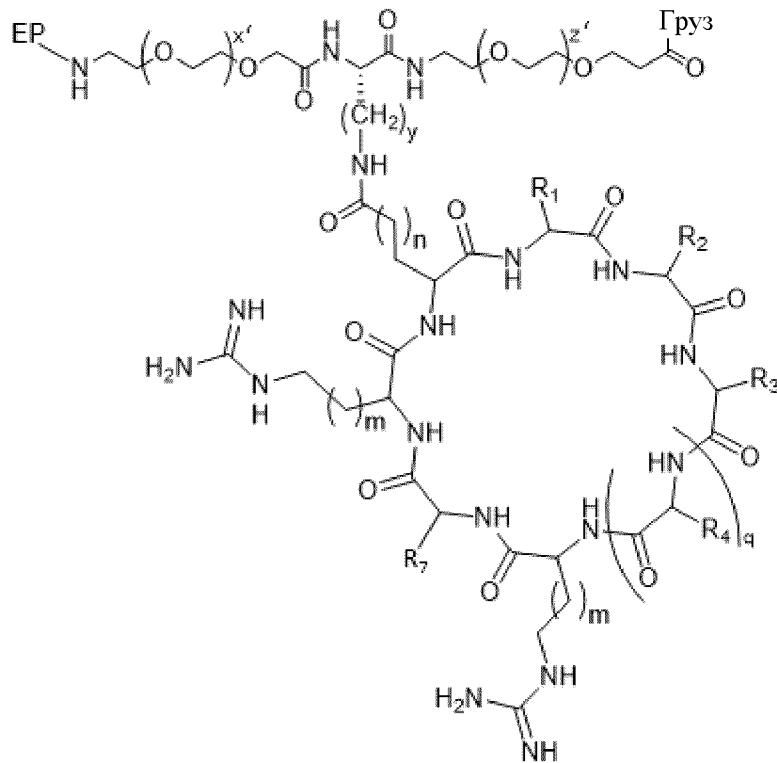
39. Соединение по п. 38, отличающееся тем, что  $z'$  представляет собой 11.

40. Соединение по п. 38 или 39, отличающееся тем, что  $x'$  представляет собой 1.

41. Соединение по любому из пп. 38-40, отличающееся тем, что указанный экзоциклический пептид конъюгирован с указанным линкером на аминоконце указанного линкера.

42. Соединение по любому из пп. 38-41, отличающееся тем, что указанное антисмысловое соединение конъюгировано с  $M$ .

43. Соединение по п. 17, отличающееся тем, что указанное соединение имеет Формулу (С):



(С). или

ее протонированную форму или соль,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или боковую цепь, содержащую арильную или гетероарильную группу, причем по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой боковую цепь, содержащую арильную или гетероарильную группу;

$R_4$  и  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

EP представляет собой указанный экзоциклический пептид;

каждый  $m$  независимо представляет собой целое число от 0 до 3;

$n$  представляет собой целое число от 0 до 2;

$x'$  представляет собой целое число от 1 до 23;

$y$  представляет собой целое число от 1 до 5;

$q$  представляет собой целое число от 1 до 4;

$z'$  представляет собой целое число от 1 до 23, и

Груз представляет собой указанное антисмысловое соединение.

44. Соединение по п. 43, отличающееся тем, что  $R_1$ ,  $R_2$  и  $R_3$  представляют собой  $H$  или боковую цепь, содержащую арильную группу.

45. Соединение по п. 44, отличающееся тем, что указанная боковая цепь, содержащая арильную группу, представляет собой боковую цепь фенилаланина.

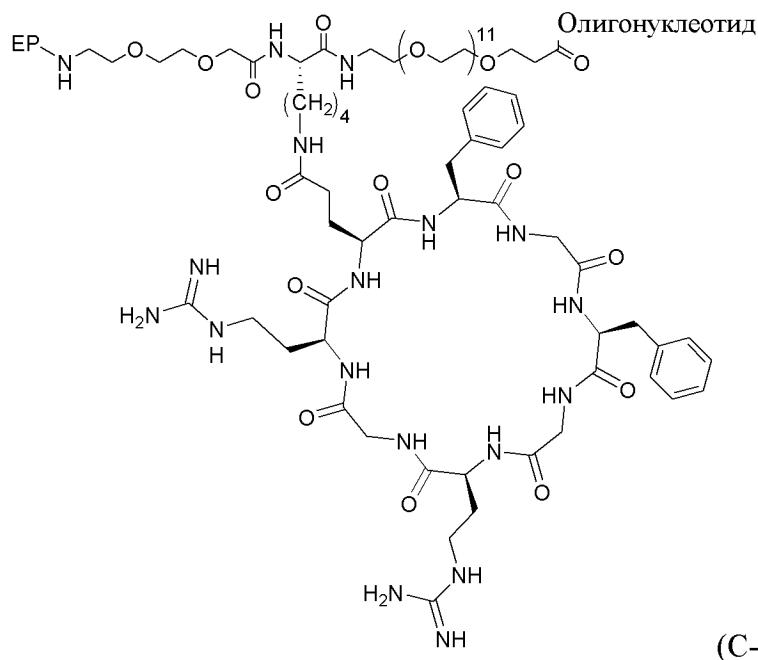
46. Соединение по п. 45, отличающееся тем, что два из  $R_1$ ,  $R_2$  и  $R_3$  представляют собой боковую цепь фенилаланина.

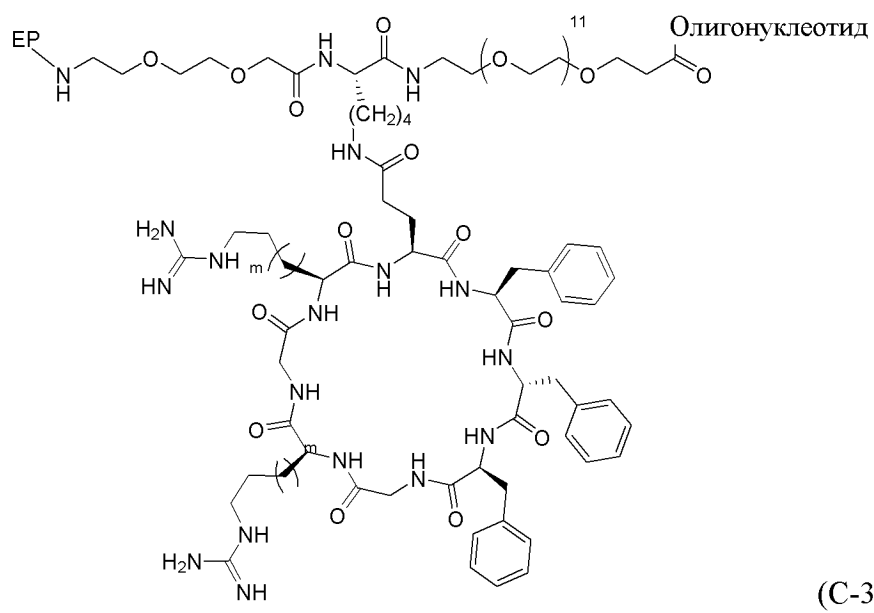
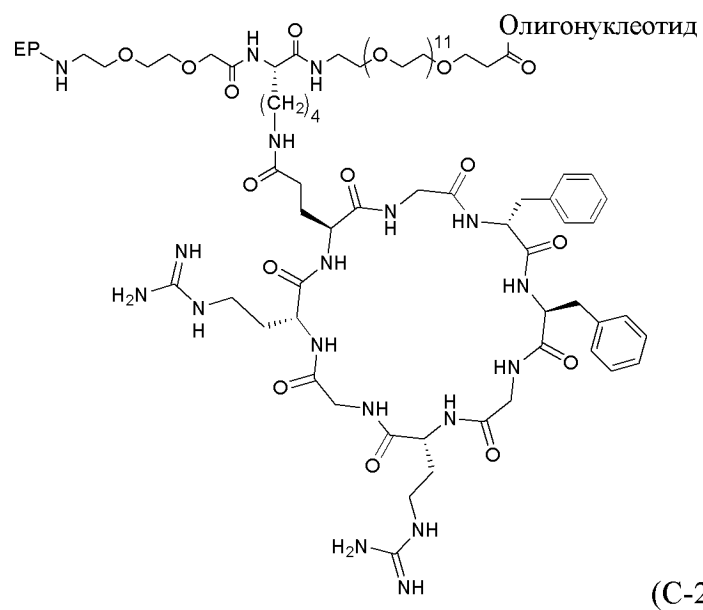
47. Соединение по любому из пп. 44-46, отличающееся тем, что два из  $R_1$ ,  $R_2$ ,  $R_3$  и  $R_4$  представляют собой  $H$ .

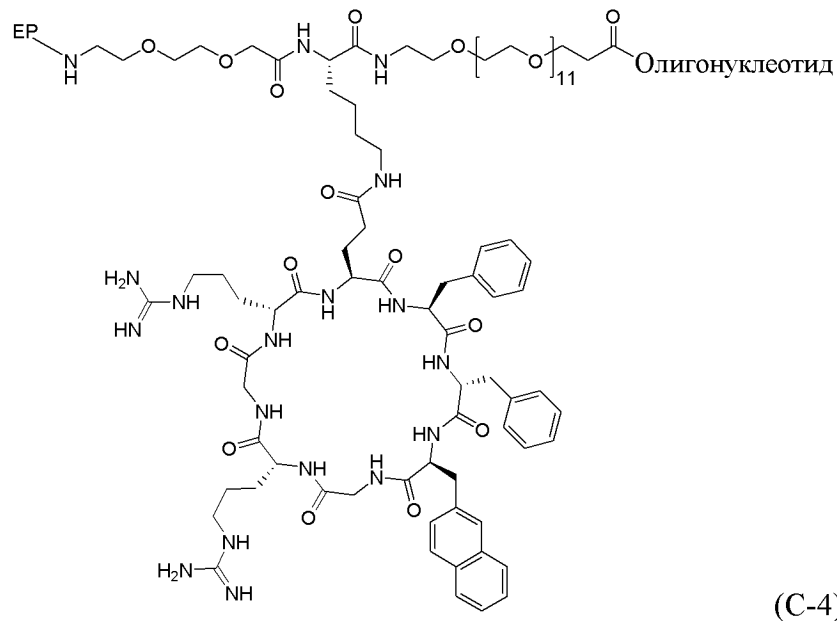
48. Соединение по любому из пп. 44-47, отличающееся тем, что  $z'$  представляет собой 11.

49. Соединение по любому из пп. 44-48, отличающееся тем, что  $x'$  представляет собой 1.

50. Соединение по п. 17, содержащее структуру Формулы (С-1), (С-2), (С-3) или (С-4):







или ее протонированную форму или соль,  
 где EP представляет собой указанный экзоциклический пептид, и  
 олигонуклеотид представляет собой указанное антисмысловое соединение.

51. Соединение по п. 50, отличающееся тем, что указанный олигонуклеотид содержит следующую последовательность: 5'-CAG CAG CAG CAG CAG CAG CAG CAG-3'.

52. Соединение по п. 50 или п. 51, отличающееся тем, что указанный EP содержит следующую последовательность: PKKKRKV.

53. Соединение по любому из предыдущих пунктов, отличающееся тем, что указанный экспансированный тринуклеотидный повтор находится в 3'-НТО последовательности целевой мРНК.

54. Соединение по п. 53, отличающееся тем, что указанная последовательность целевой мРНК представляет собой последовательность мРНК DMPK.

55. Соединение по п. 53, отличающееся тем, что указанная последовательность целевой мРНК представляет собой последовательность мРНК ATXN8OS/ATXN8.

56. Соединение по п. 53, отличающееся тем, что указанная последовательность целевой мРНК представляет собой последовательность мРНК JPH3.

57. Соединение по любому из предыдущих пунктов, отличающееся тем, что указанный AC содержит 5-10 повторов CAG.

58. Соединение по любому из предыдущих пунктов, отличающееся тем, что указанная мРНК представляет собой пре-мРНК или зрелую мРНК.

59. Соединение по любому из предыдущих пунктов, отличающееся тем, что указанная мРНК представляет собой пре-мРНК.

60. Фармацевтическая композиция, содержащая эффективное количество соединения по любому из пп. 1-59.

61. Клетка, содержащая соединение по любому из пп. 1-59.

62. Способ лечения миотонической дистрофии (DM) у субъекта, нуждающегося в этом, включающий введение указанному субъекту соединения по любому из пп. 1-59 или композиции по п. 60.

63. Способ по п. 62, отличающийся тем, что указанное введение приводит к повышению экспрессии белка дикого типа в мышечной ткани.

64. Способ по п. 63, отличающийся тем, что указанное введение приводит к повышению экспрессии белка дикого типа в ткани диафрагмы, ткани четырехглавой мышцы и/или ткани сердца.

65. Способ по п. 63 или п. 64, отличающийся тем, что указанный белок дикого типа представляет собой белок, экспрессируемый из гена, который не имеет экспансированного повтора CUG.

66. Способ по любому из пп. 62-65, отличающийся тем, что указанное введение обеспечивает предотвращение или снижение образования очагов.

67. Способ по любому из пп. 62-66, отличающийся тем, что указанные экспансии нуклеотидных повторов расположены в 3'-НТО мРНК DMPK.

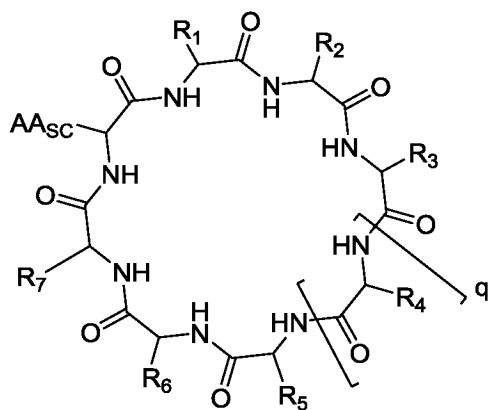
68. Способ лечения заболевания, ассоциированного с мРНК, содержащей экспансированный повтор CUG в 3'-нетранслируемой области (НТО), отличающийся тем, что указанный способ включает введение субъекту, нуждающемуся в этом, соединения, содержащего:

антисмысловое соединение (АС), которое комплементарно по меньшей мере части экспансированного повтора CUG в мРНК; и

циклический пептид, содержащий от 6 до 12 аминокислот, причем по меньшей мере две аминокислоты указанного циклического пептида представляют собой заряженные аминокислоты, по меньшей мере две аминокислоты указанного циклического пептида представляют собой ароматические гидрофобные аминокислоты и по меньшей мере две аминокислоты указанного циклического пептида представляют собой незаряженные неароматические аминокислоты.

69. Способ по п. 68, отличающийся тем, что по меньшей мере две заряженные аминокислоты указанного циклического пептида представляют собой аргинин, по меньшей мере две ароматические гидрофобные аминокислоты указанного циклического пептида представляют собой фенилаланин, нафтилаланин или их комбинации и по меньшей мере две незаряженные неароматические аминокислоты указанного циклического пептида представляют собой цитруллин, глицин или их комбинации.

70. Способ по п. 68, отличающийся тем, что указанный циклический пептид имеет следующую структуру:



(A) или

ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или ароматическую или гетероароматическую боковую цепь аминокислоты;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

$R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  представляет собой H или боковую цепь цитруллина;

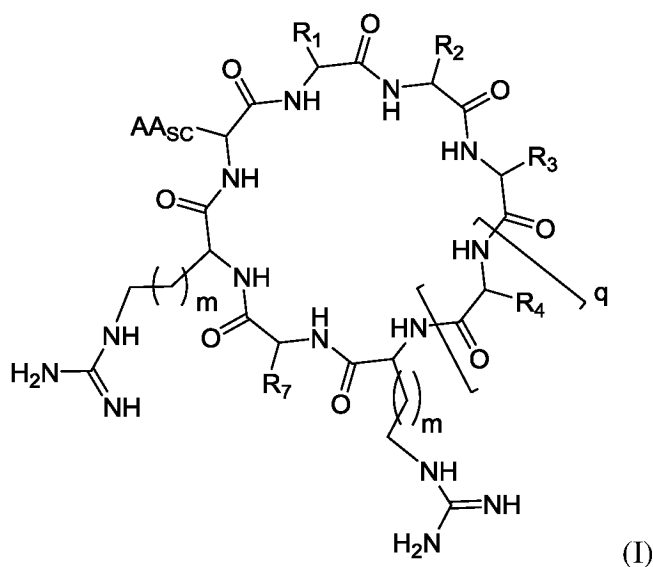
$AA_{sc}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение; и

$q$  представляет собой 1, 2, 3 или 4.

71. Способ по п. 70, отличающийся тем, что по меньшей мере один из  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_7$  представляет собой боковую цепь 3-гуанидин-2-аминопропионовой кислоты, 4-гуанидин-2-аминобутановой кислоты, аргинина, гомоаргинина, N-метиларгинина, N,N-диметиларгинина, 2,3-диаминопропионовой кислоты, 2,4-диаминобутановой кислоты, лизина, N-метиллизина, N, N-диметиллизина, N-этиллизина, N, N, N-триметиллизина, 4-гуанидинофенилаланина, N,N-диметиллизина,  $\beta$ -гомоаргинина или 3-(1-пиперидинил)аланина.

72. Способ по п. 68, отличающийся тем, что указанный циклический пептид имеет следующую структуру:





или ее протонированную форму,

где:

каждый  $R_1$ ,  $R_2$  и  $R_3$  независимо представляет собой H или аминокислотный остаток, имеющий боковую цепь, содержащую ароматическую группу;

по меньшей мере один из  $R_1$ ,  $R_2$  и  $R_3$  представляет собой ароматическую или гетероароматическую боковую цепь аминокислоты;

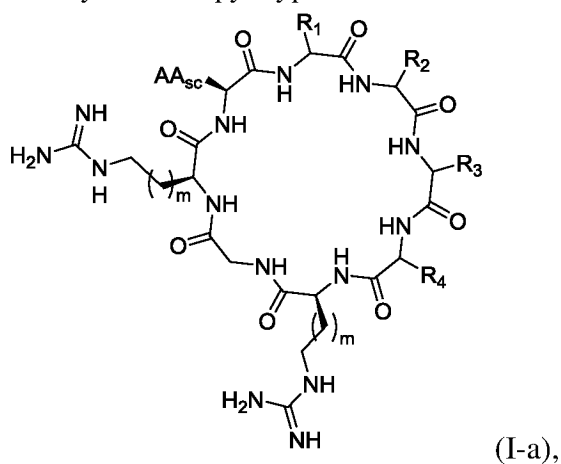
$R_4$  и  $R_7$  независимо представляют собой H или боковую цепь аминокислоты;

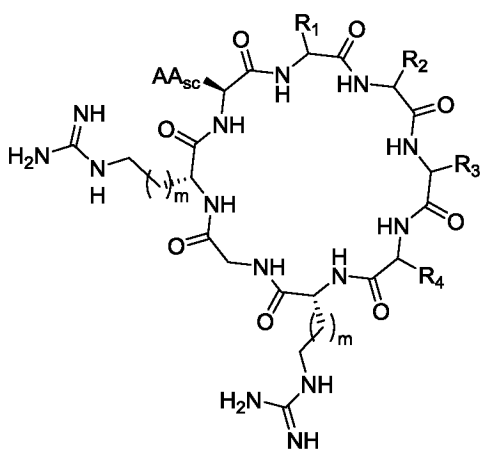
$AA_{sc}$  представляет собой боковую цепь аминокислоты, с которой конъюгировано указанное антисмысловое соединение;

$q$  представляет собой 1, 2, 3 или 4; и

каждый  $m$  независимо представляет собой целое число 0, 1, 2 или 3.

73. Способ по п. 72, отличающийся тем, что указанный циклический пептид имеет одну из следующих структур:

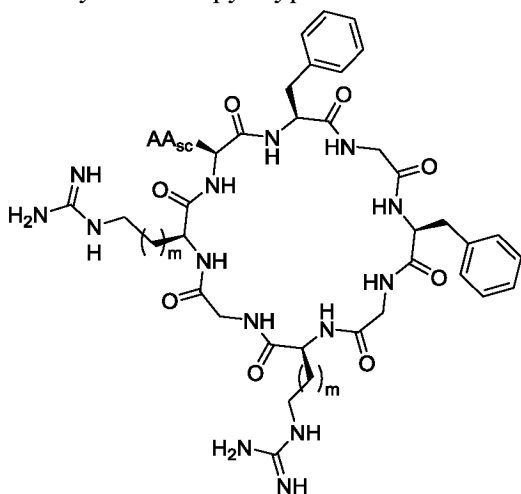




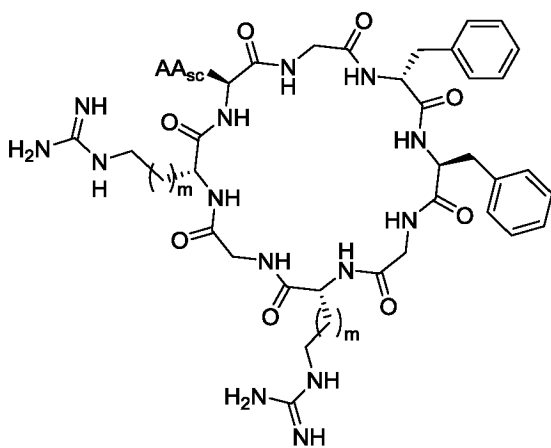
(I-b), или

ее протонированную форму.

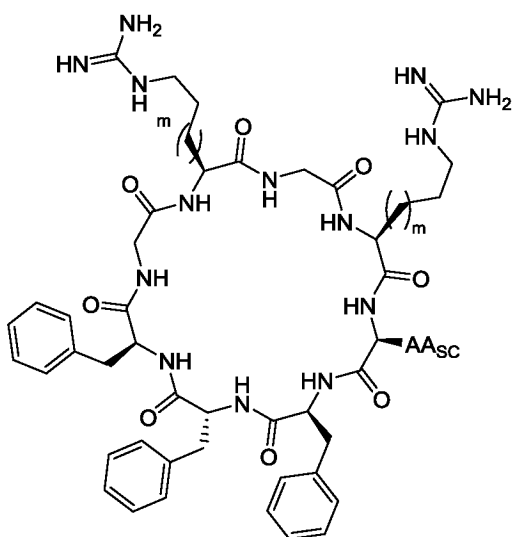
74. Способ по п. 72, отличающийся тем, что указанный циклический пептид имеет одну из следующих структур:



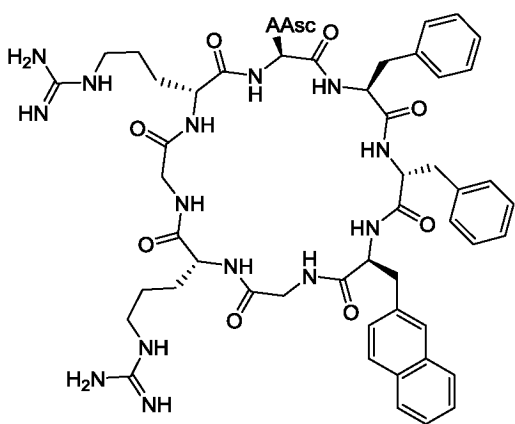
(I-1),



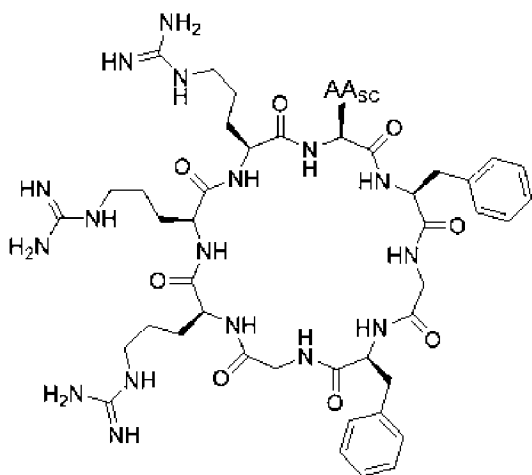
(I-2),



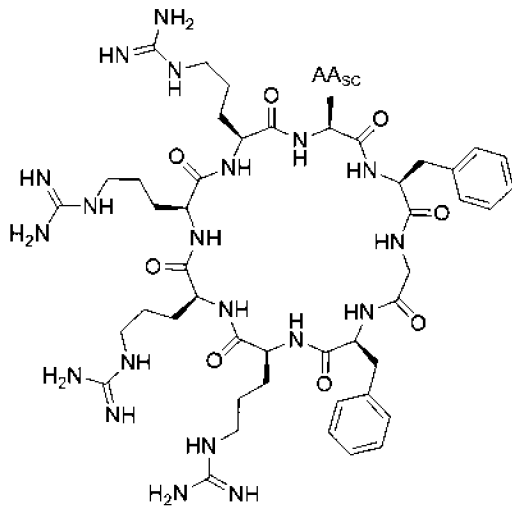
(I-3),



(I-4),



(I-5),



(I-6), или

ее протонированную форму.

75. Способ по любому из пп. 68-74, отличающийся тем, что указанное антисмысловое соединение содержит фосфордиамидат-морфолиновый (РМО) нуклеотид.

76. Способ по любому из пп. 68-75, дополнительно содержащий линкер, с помощью которого указанное антисмысловое соединение конъюгировано с AA<sub>sc</sub>.

77. Способ по любому из пп. 68-76, включающий носитель для ускользания из эндосом, отличающийся тем, что указанный носитель для ускользания из эндосом содержит указанный циклический пептид и экзоциклический пептид.

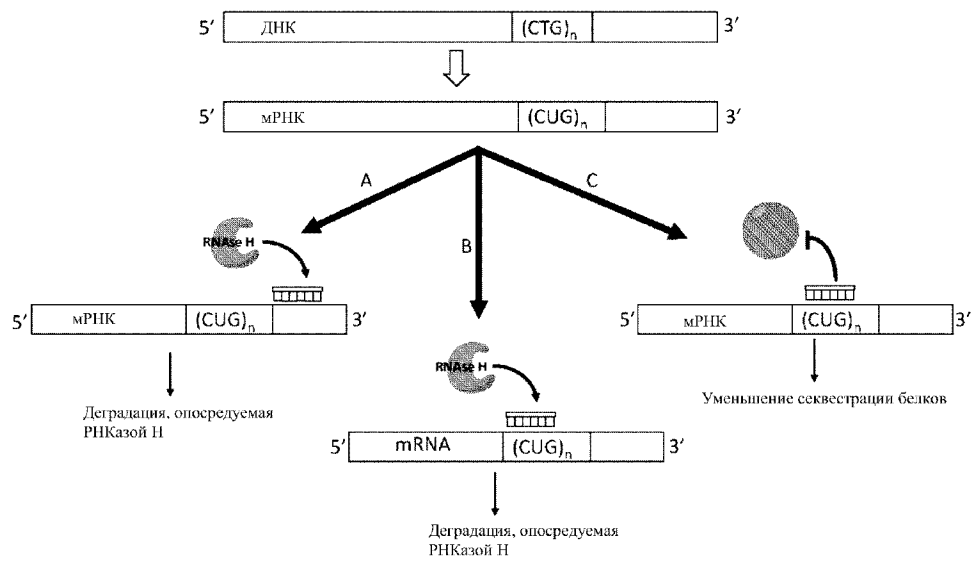
78. Способ по любому из пп. 68-77, отличающийся тем, что указанное заболевание представляет собой миотоническую дистрофию (DM).

79. Способ по любому из пп. 68-77, отличающийся тем, что указанное заболевание представляет собой спиноцеребеллярную атаксию 8 типа (SCA8)

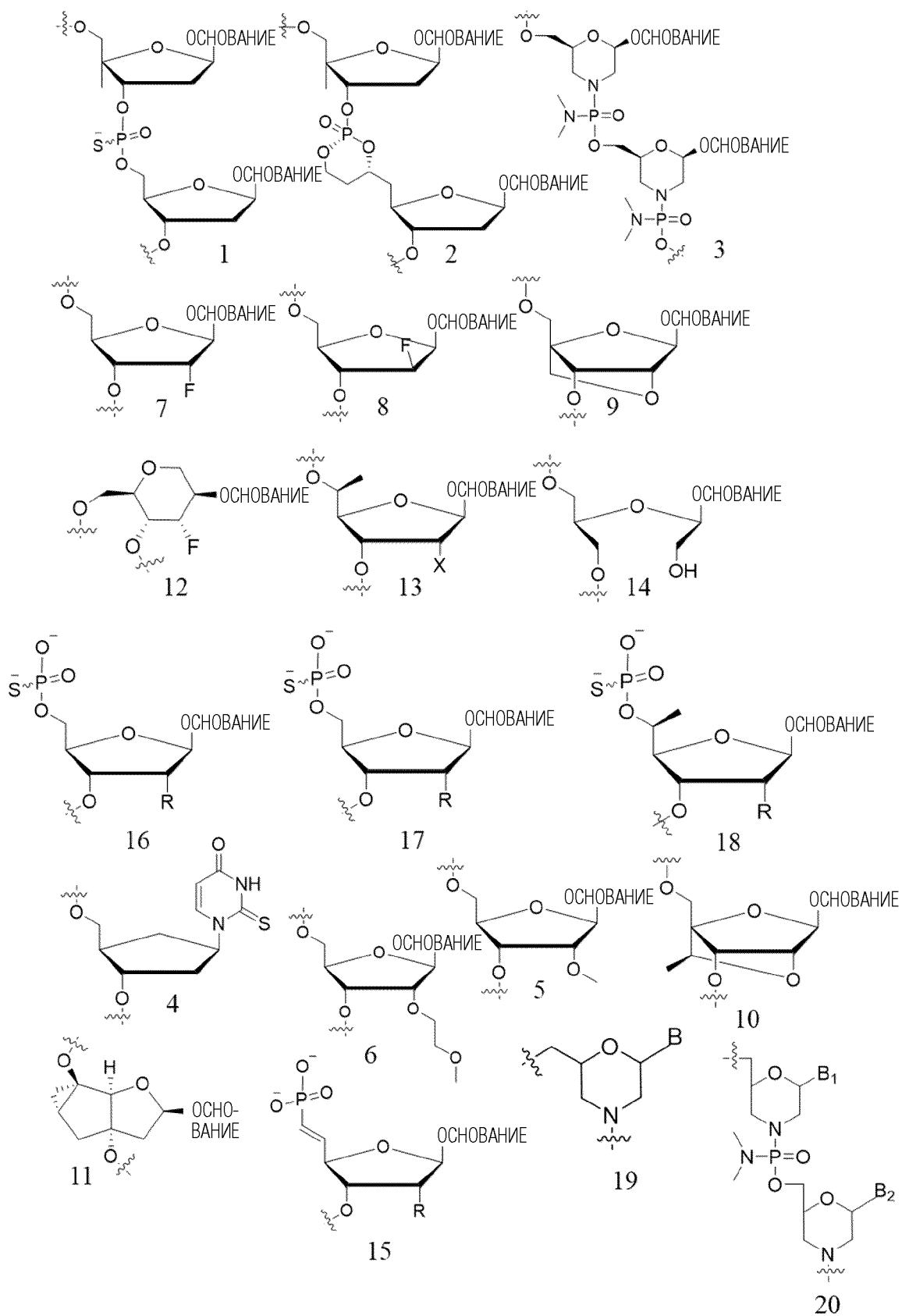
80. Способ по любому из пп. 68-77, отличающийся тем, что указанное заболевание представляет собой болезнь, подобную болезни Хантингтона, 2 типа (HDL2).

81. Соединение, композиция или способ по любому из предыдущих пунктов, отличающиеся тем, что указанное антисмысловое соединение содержит последовательность нуклеотидов, перечисленных в Таблице 2.

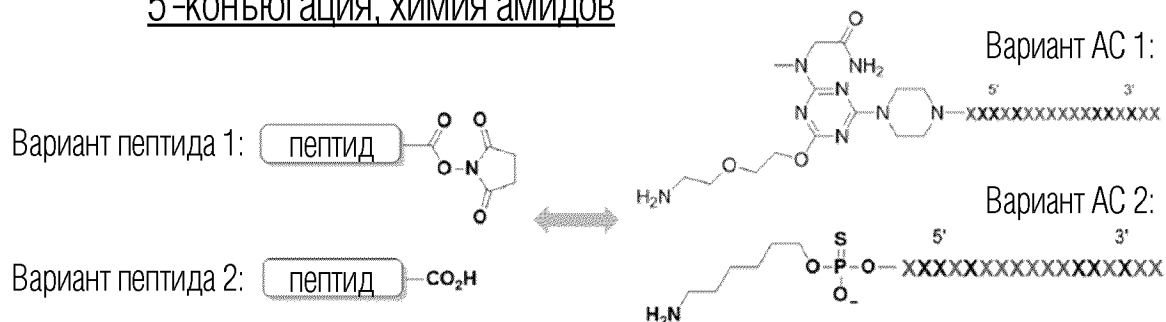
По доверенности



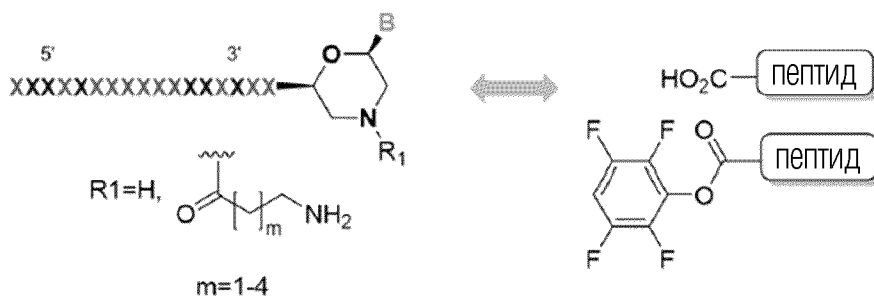
ФИГ. 1



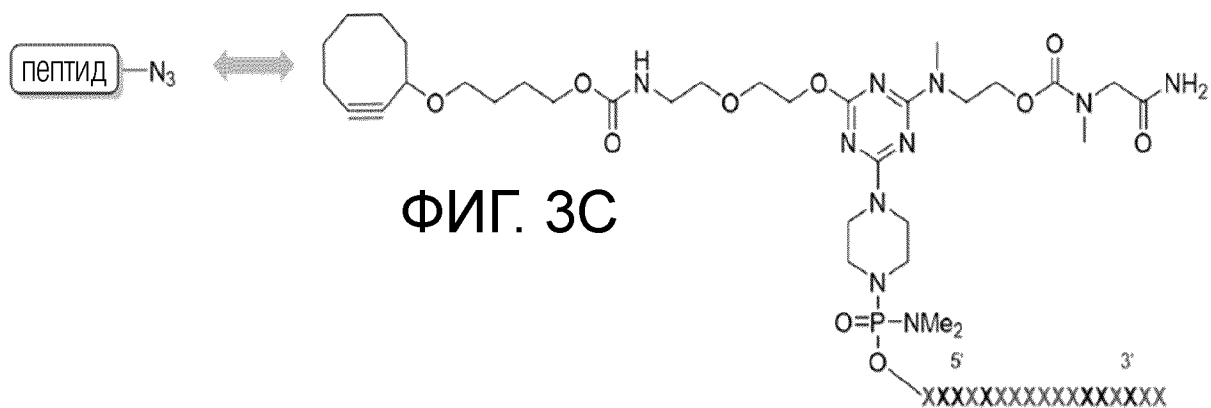
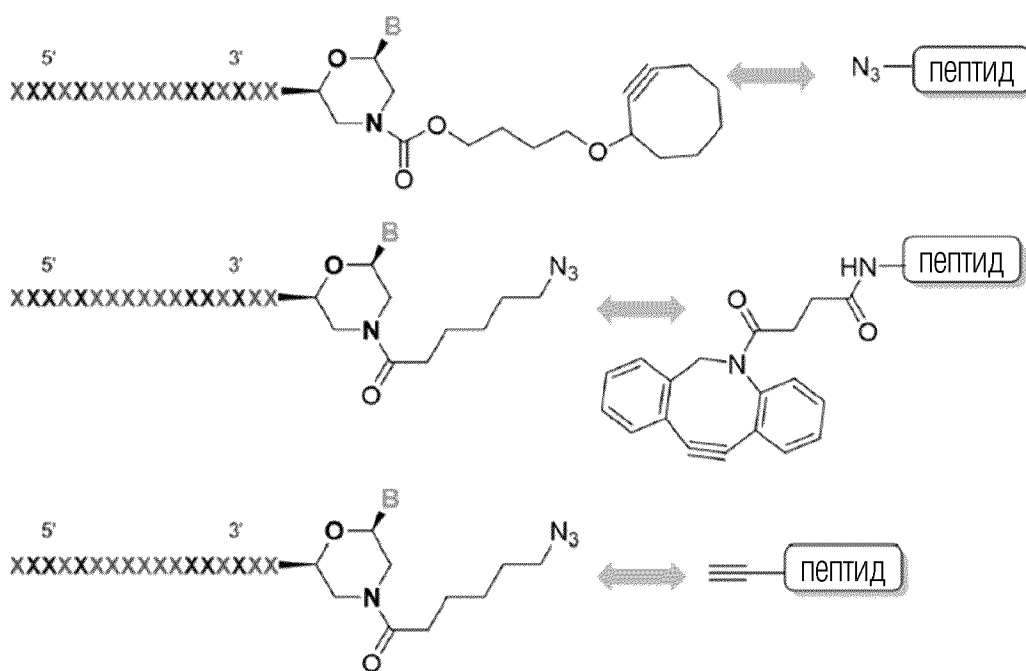
ФИГ. 2

5'-конъюгация, химия амидов

ФИГ. 3А

3'-конъюгация, химия амидов

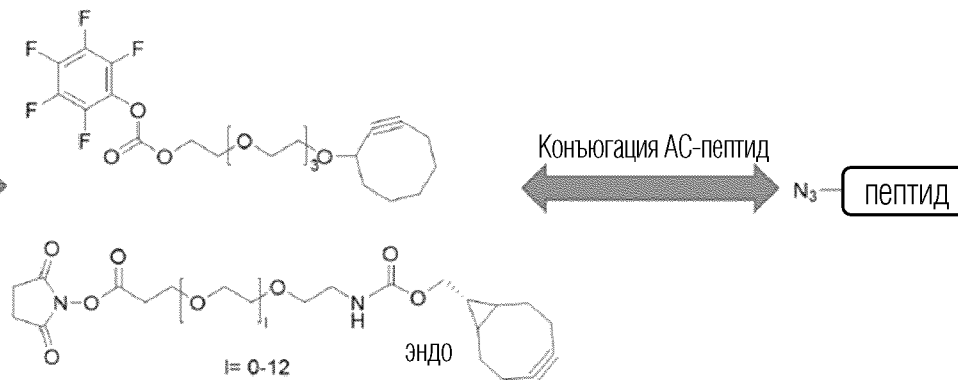
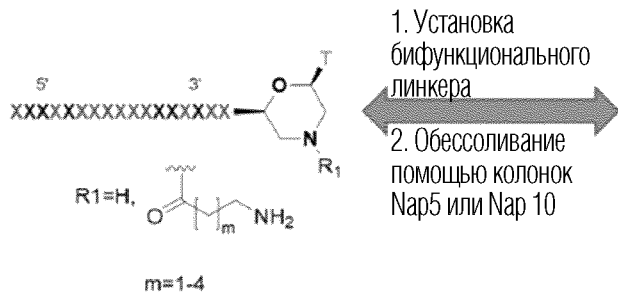
ФИГ. 3В

5'-конъюгация, клик-химия3'-конъюгация, клик-химия

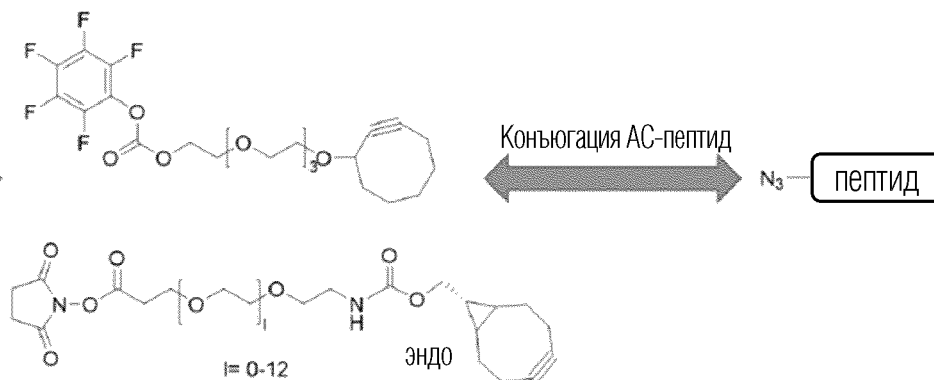
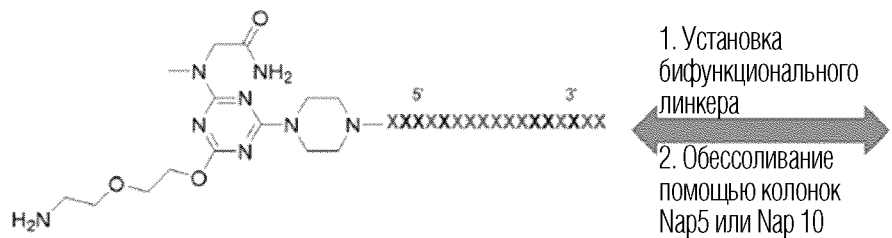
ФИГ. 3D



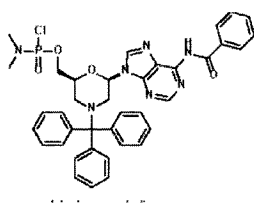
3'-циклооктиновая модификация с применением бифункционального линкера PEG



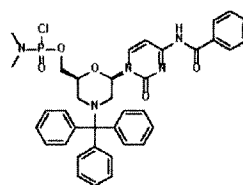
5'-циклооктиновая модификация с применением бифункционального линкера PEG



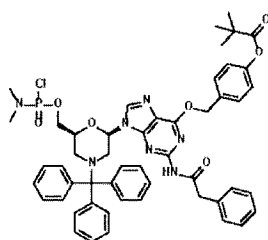
ФИГ. 4



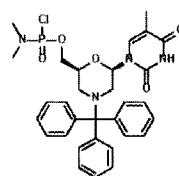
ФИГ. 5А



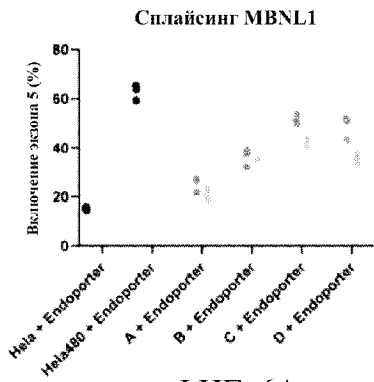
ФИГ. 5В



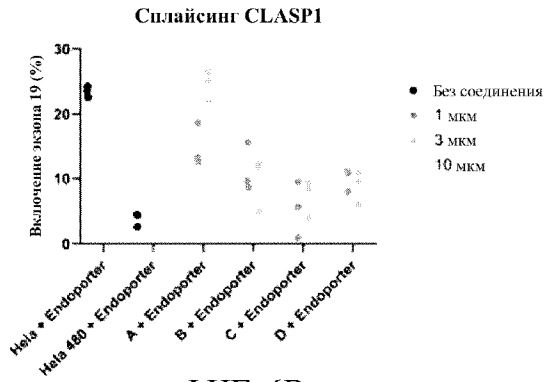
ФИГ. 5С



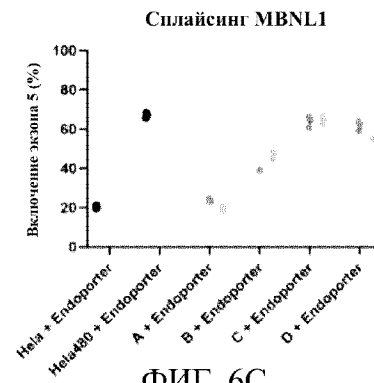
ФИГ. 5D



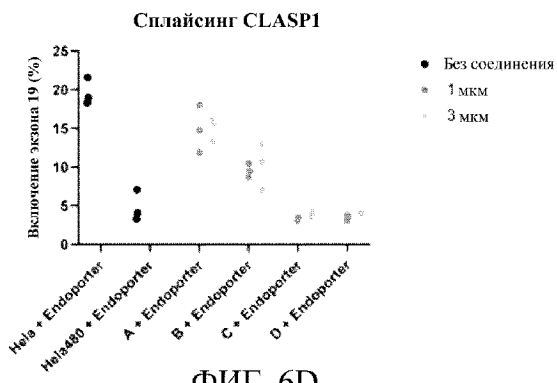
ФИГ. 6А



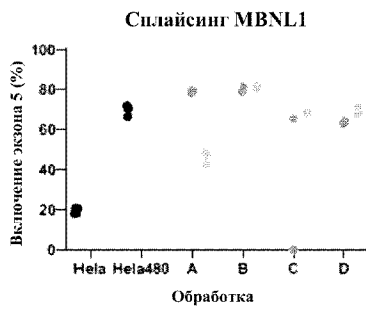
ФИГ. 6В



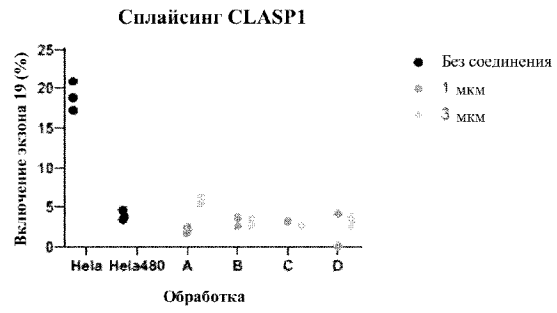
ФИГ. 6С



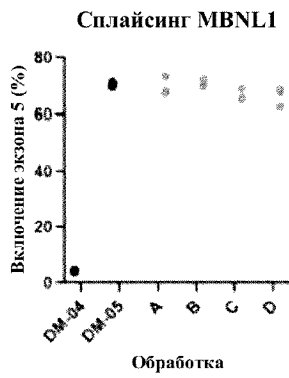
ФИГ. 6D



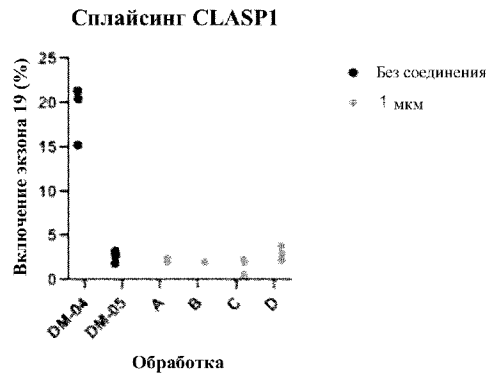
ФИГ. 6Е



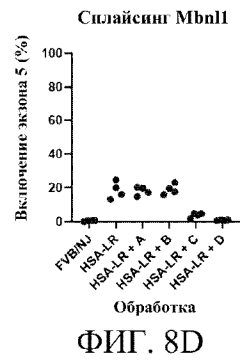
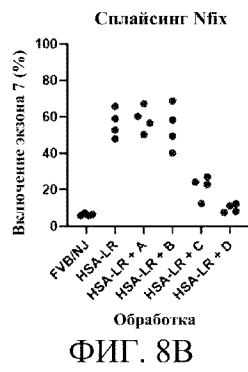
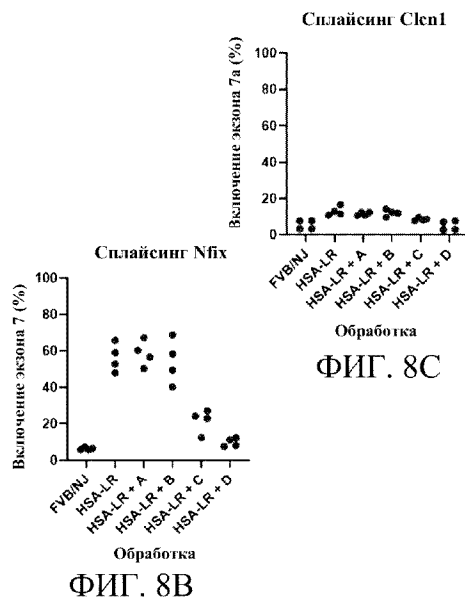
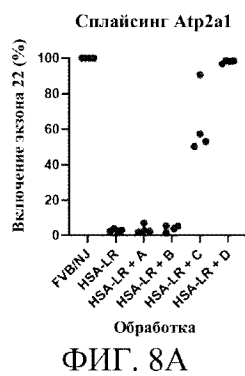
ФИГ. 6F

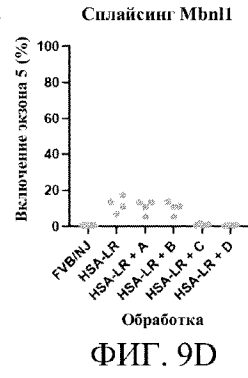
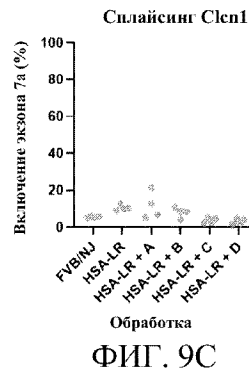
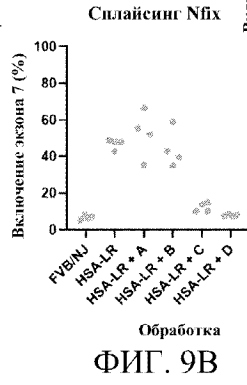
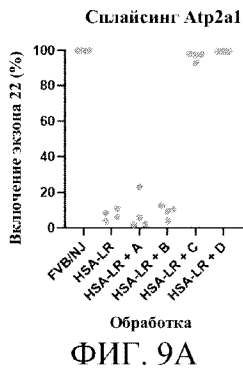


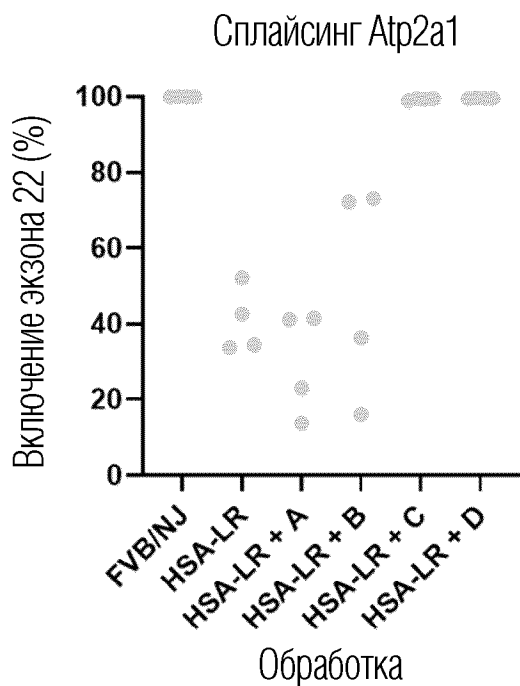
ФИГ. 7А



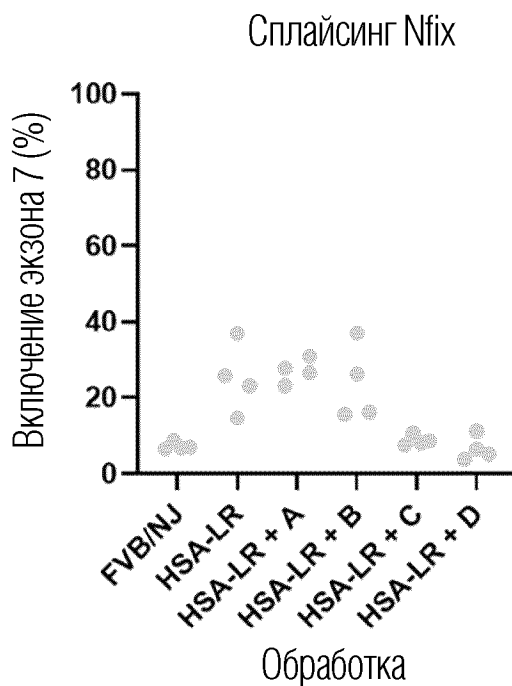
ФИГ. 7В



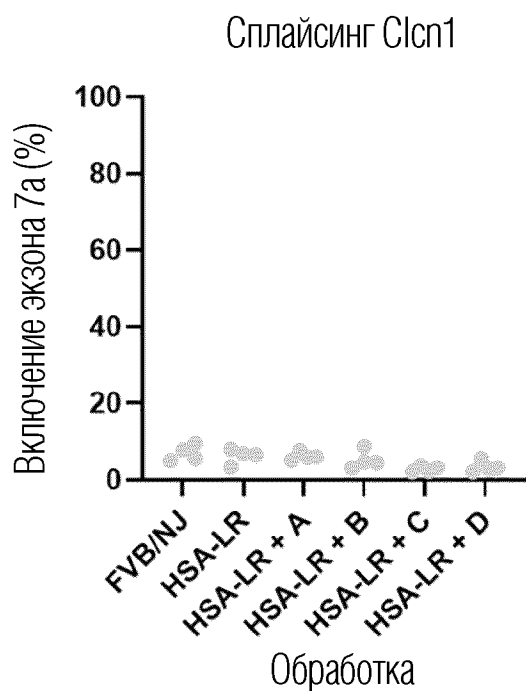




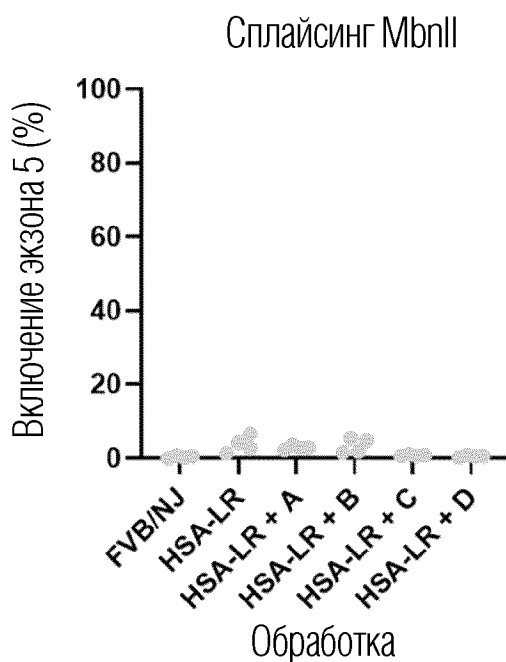
ФИГ. 10А



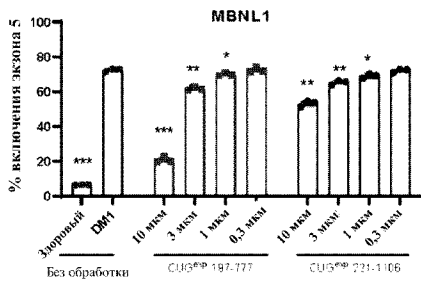
ФИГ. 10В



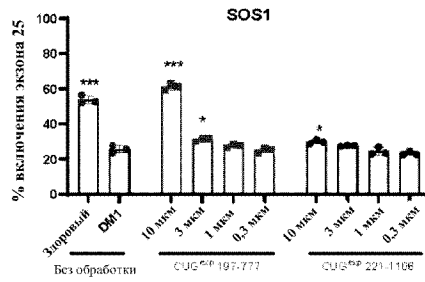
ФИГ. 10С



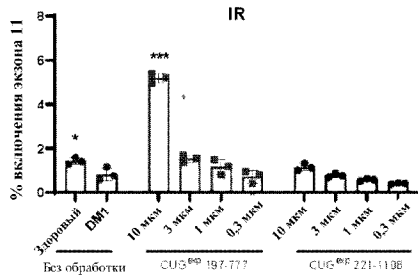
ФИГ. 10D



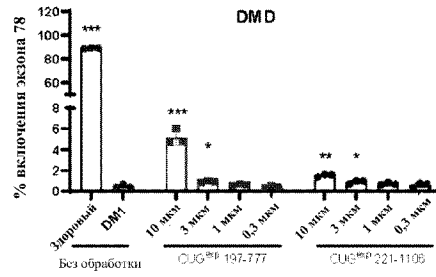
ФИГ. 11А



ФИГ. 11В

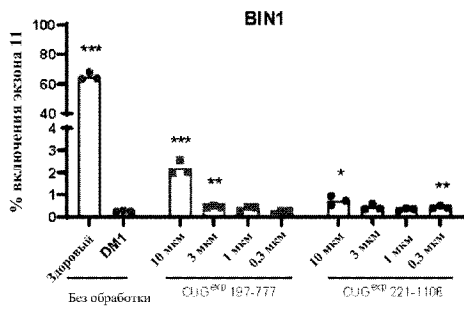


ФИГ. 11С

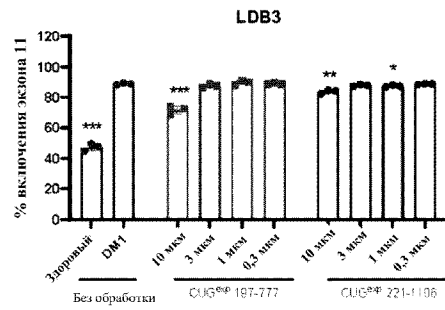


ФИГ. 11D

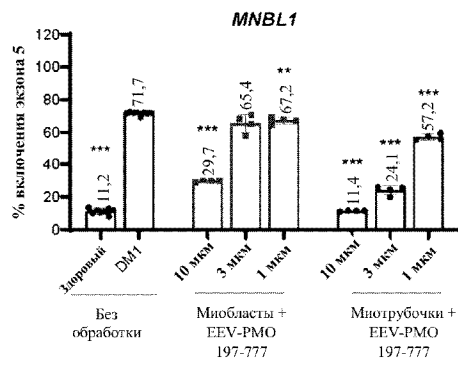




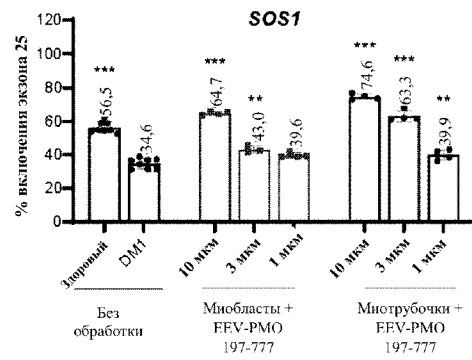
ФИГ. 11E



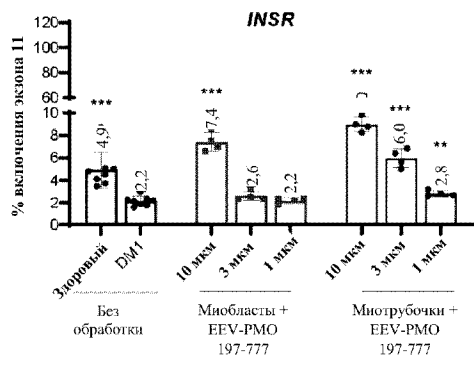
ФИГ. 11F



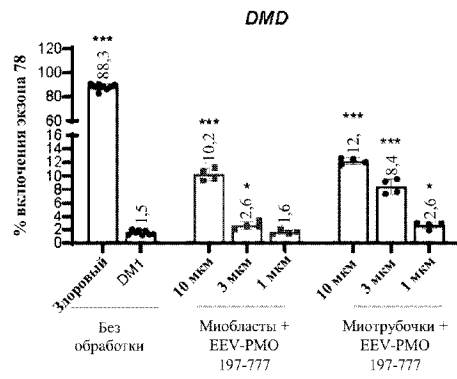
ФИГ. 12А



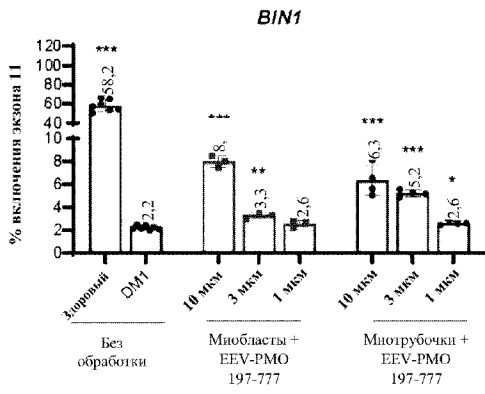
ФИГ. 12В



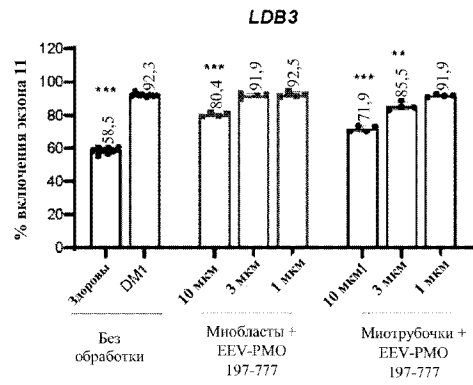
ФИГ. 12С



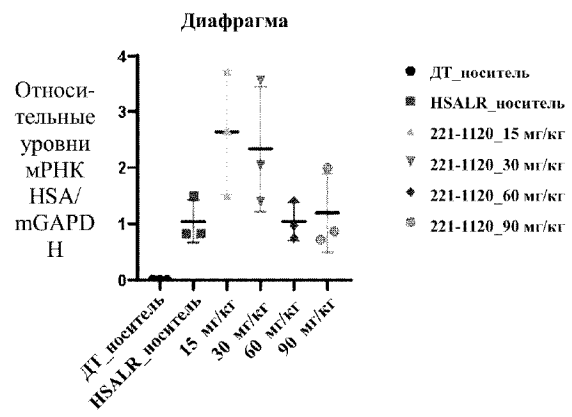
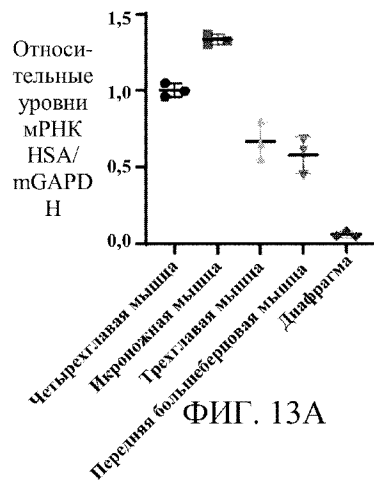
ФИГ. 12D



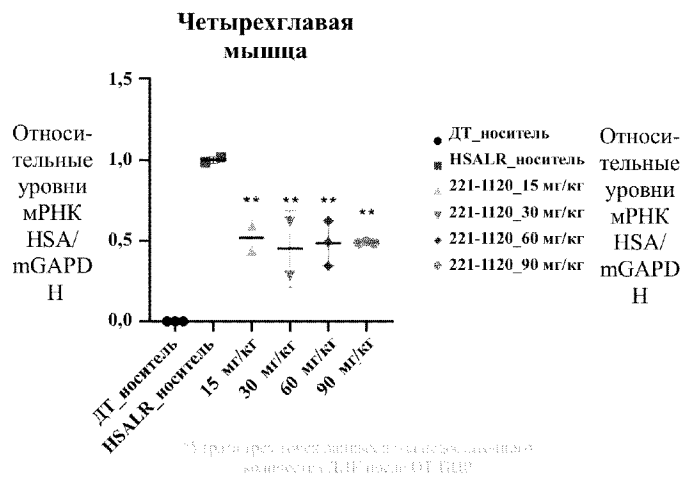
ФИГ. 12Е



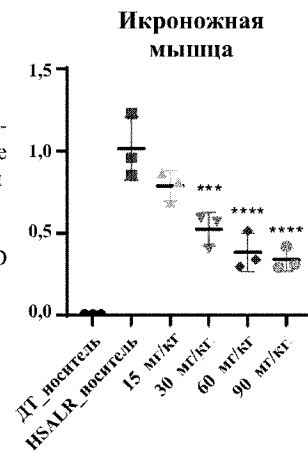
ФИГ. 12F



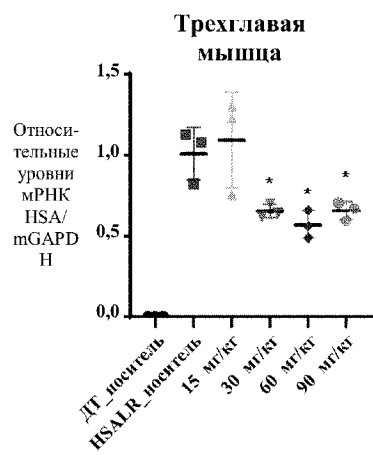
ФИГ. 13В



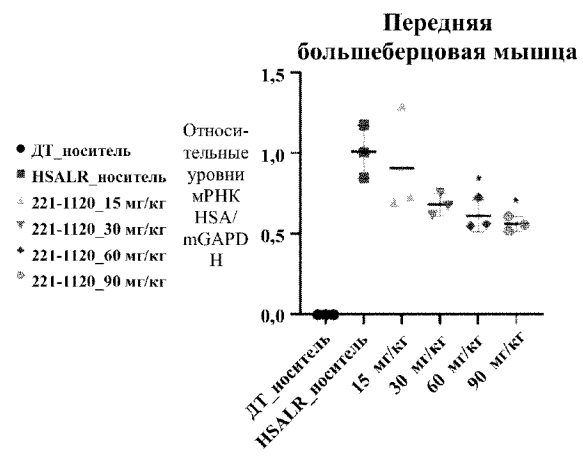
ФИГ. 14А



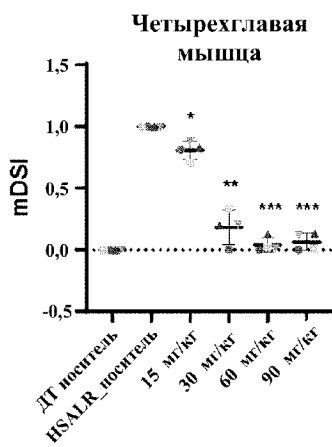
ФИГ. 14В



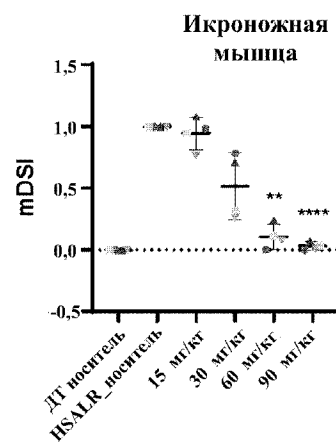
ФИГ. 14С



ФИГ. 14D



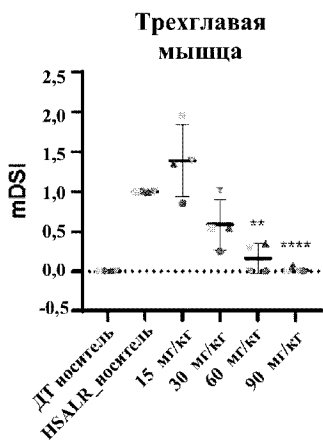
ФИГ. 15А



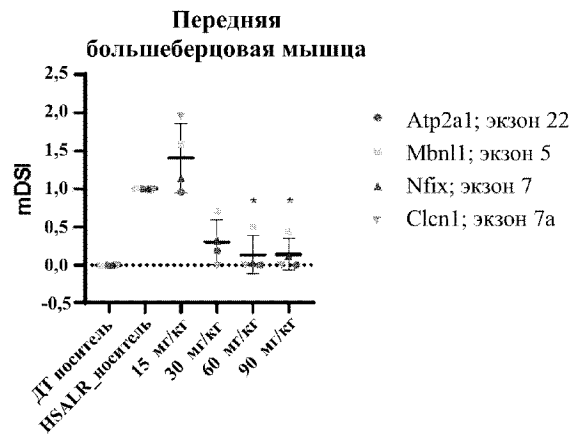
ФИГ. 15В

- ◆ Atp2a1; экзон 22
- ⊗ Mbnl1; экзон 5
- ▲ Nfix; экзон 7
- ▼ Clcn1; экзон 7a





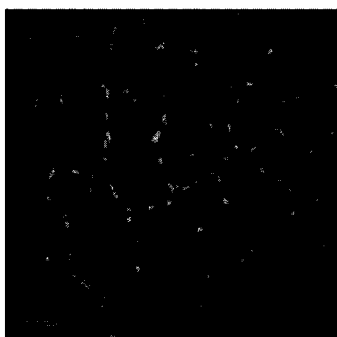
ФИГ. 15C



ФИГ. 15D

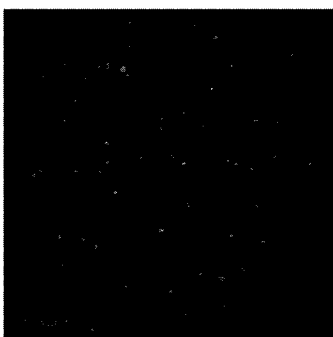
- ◆ Atp2a1; экзон 22
- ◇ Mbnl1; экзон 5
- ▲ Nfix; экзон 7
- ▽ Cln1; экзон 7a

HSA-LR носитель

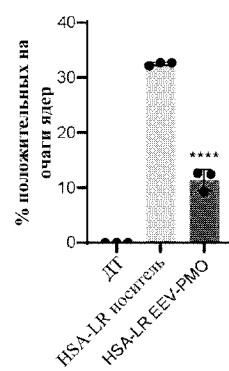


ФИГ. 16А

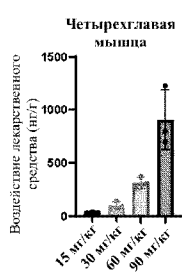
HSA-LR EEV-PMO



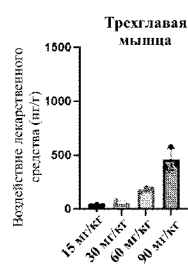
ФИГ. 16В



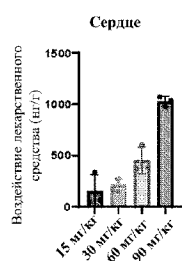
ФИГ. 16С



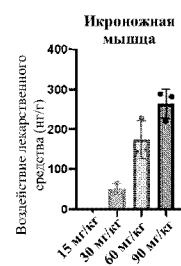
ФИГ. 17А



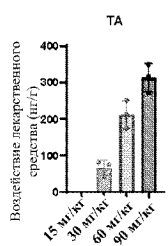
ФИГ. 17В



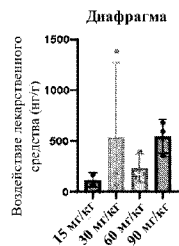
ФИГ. 17С



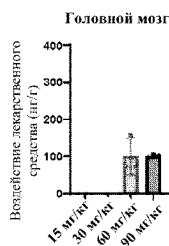
ФИГ. 17Е



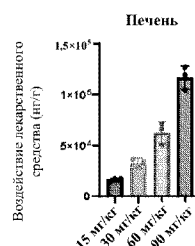
ФИГ. 17F



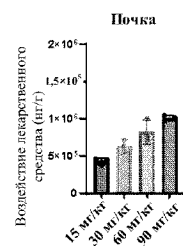
ФИГ. 17G



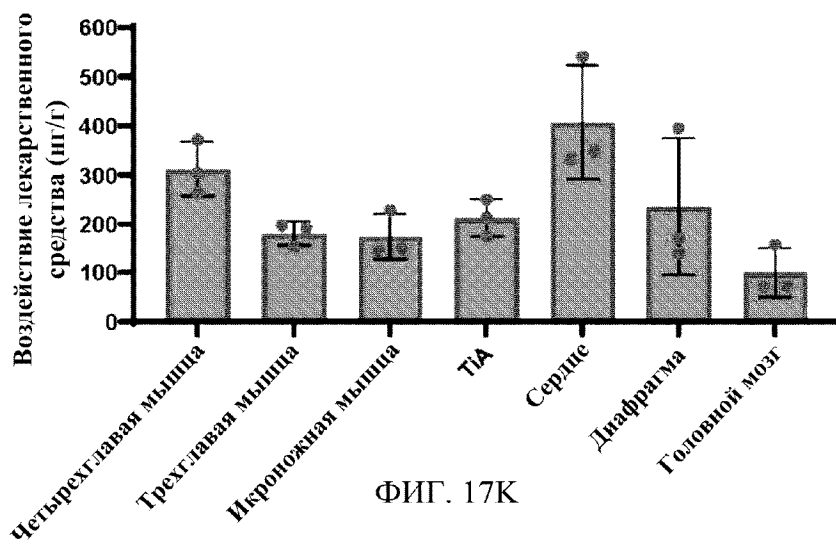
ФИГ. 17H

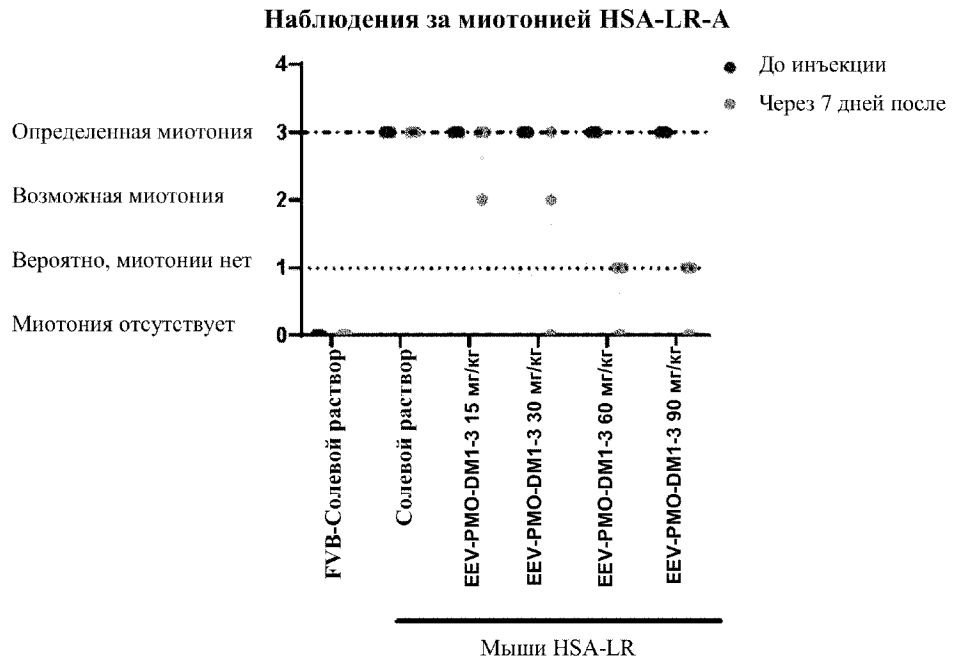


ФИГ. 17I

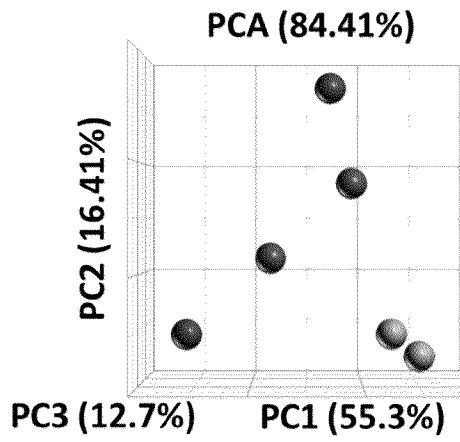


ФИГ. 17J

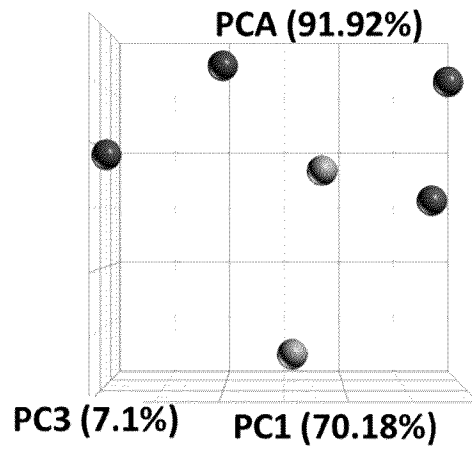




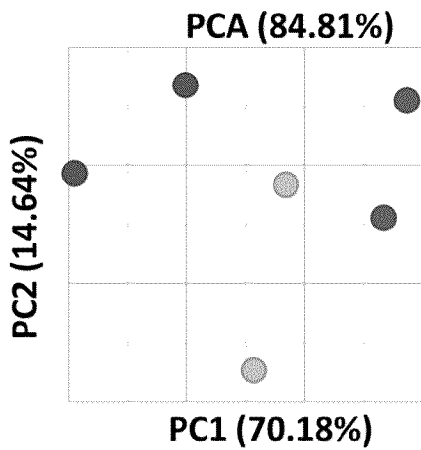
ФИГ. 18А



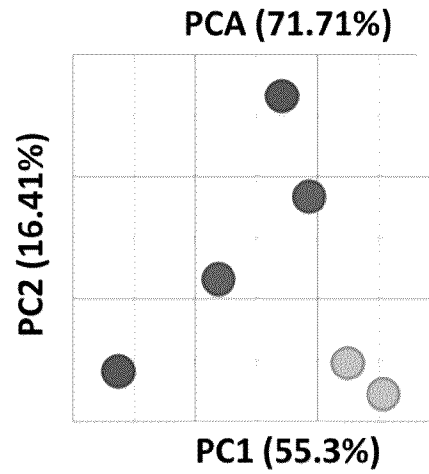
ФИГ. 19А



ФИГ. 19В



ФИГ. 19С

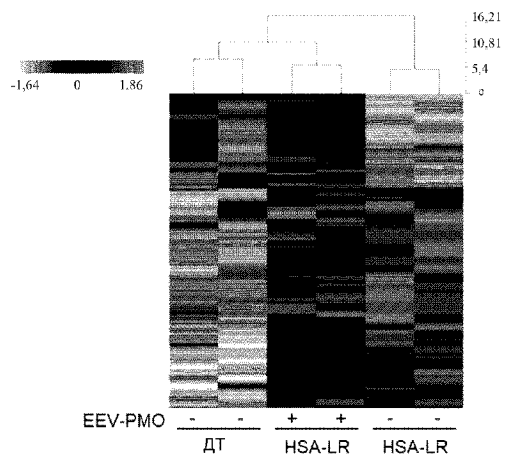


ФИГ. 19D

Обработка

- WT
- HSA-LR
- 221-1120

6 баллов

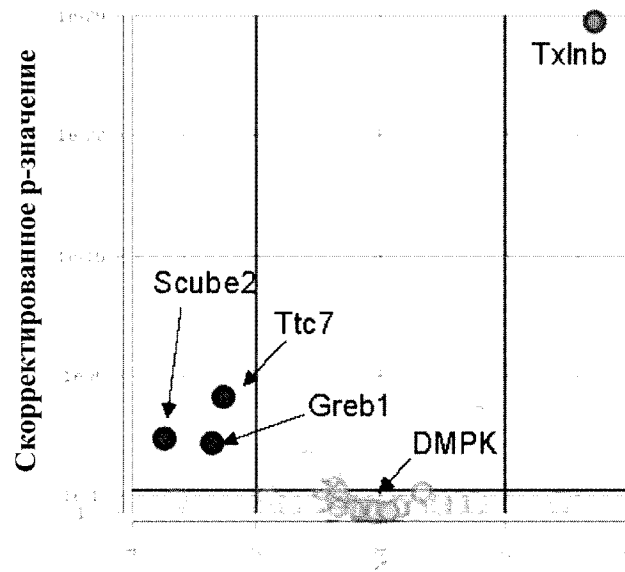


ФИГ. 20А



ФИГ. 20В

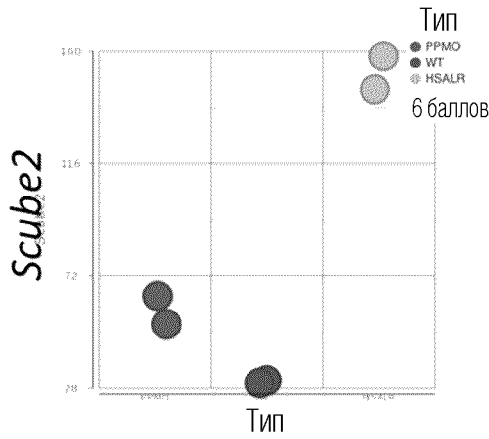
**Общее изменение транскрипции  
(гены с 40 повторами CTG) HAS-LA-PPMO  
в сравнении с HAS-LA**



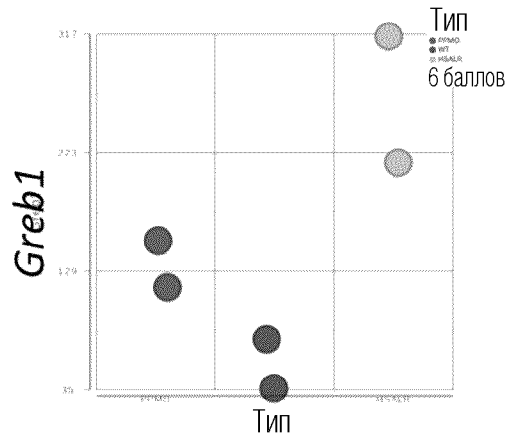
Кратное изменение (EEV-PMO по сравнению с HSA-LR)

ФИГ. 21

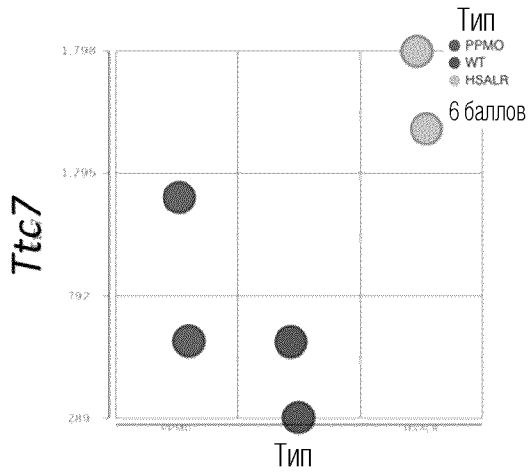




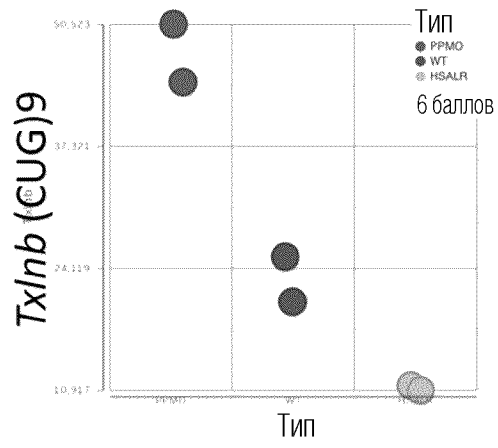
ФИГ. 22А



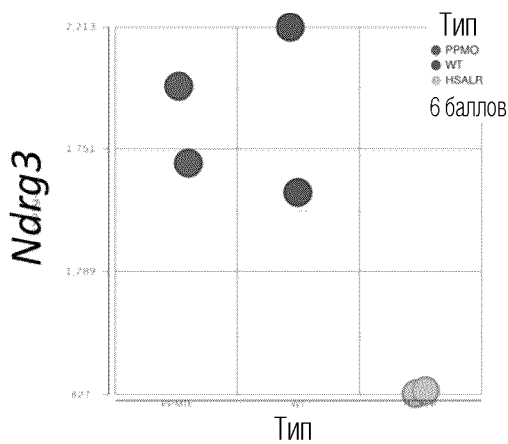
ФИГ. 22В



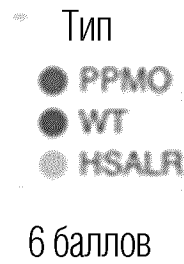
ФИГ. 22С

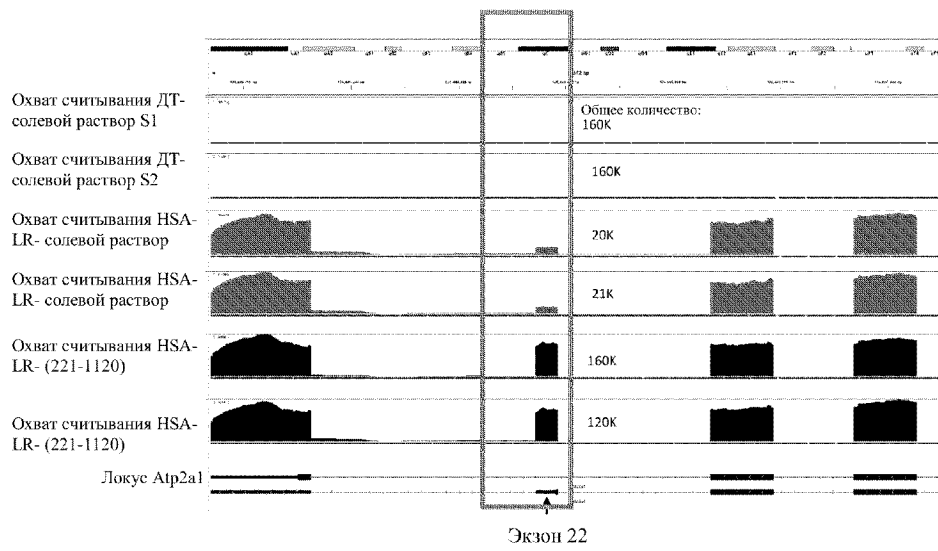


ФИГ. 22D

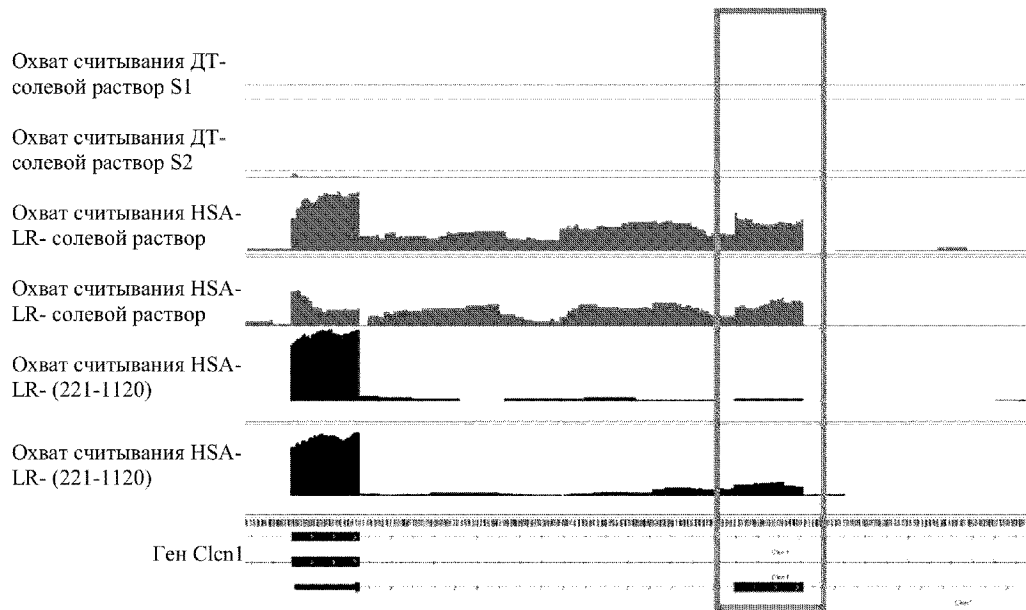


ФИГ. 22Е

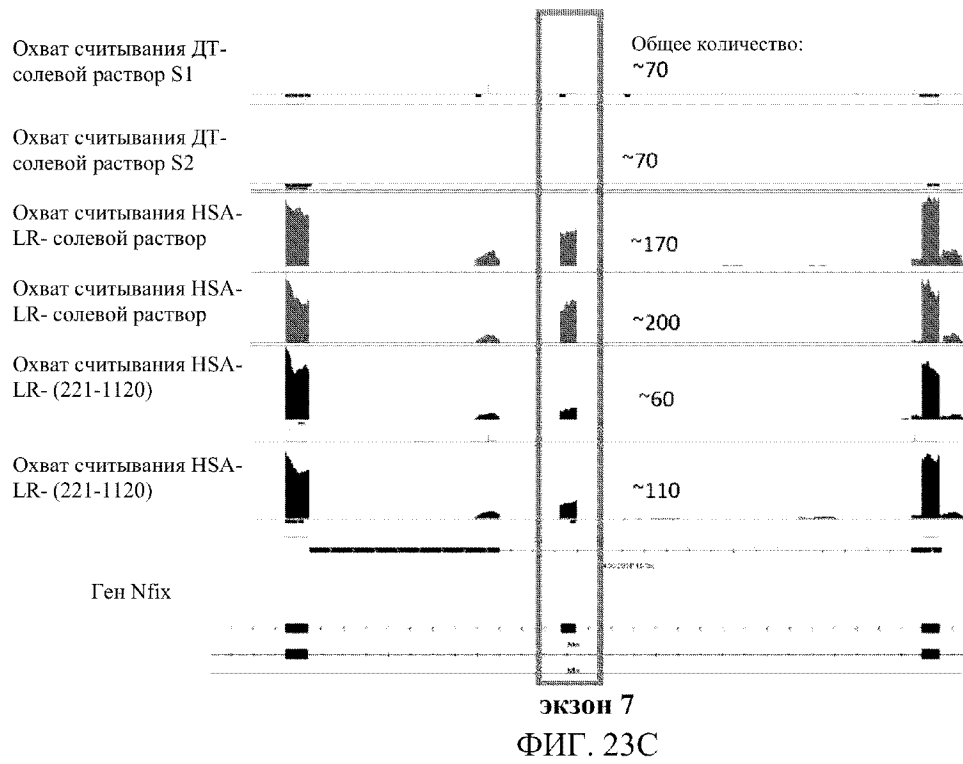


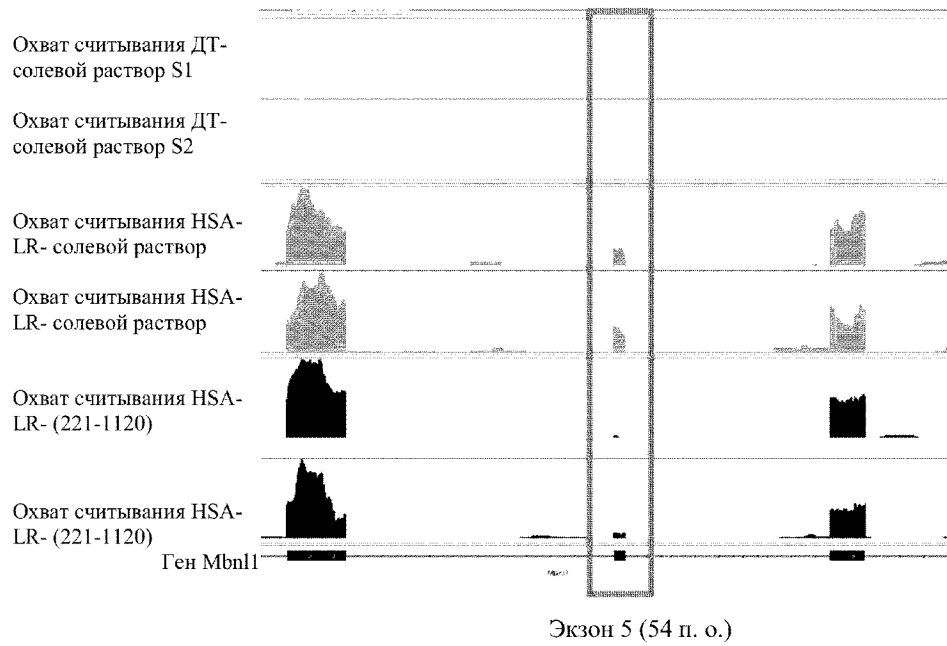


ФИГ. 23А

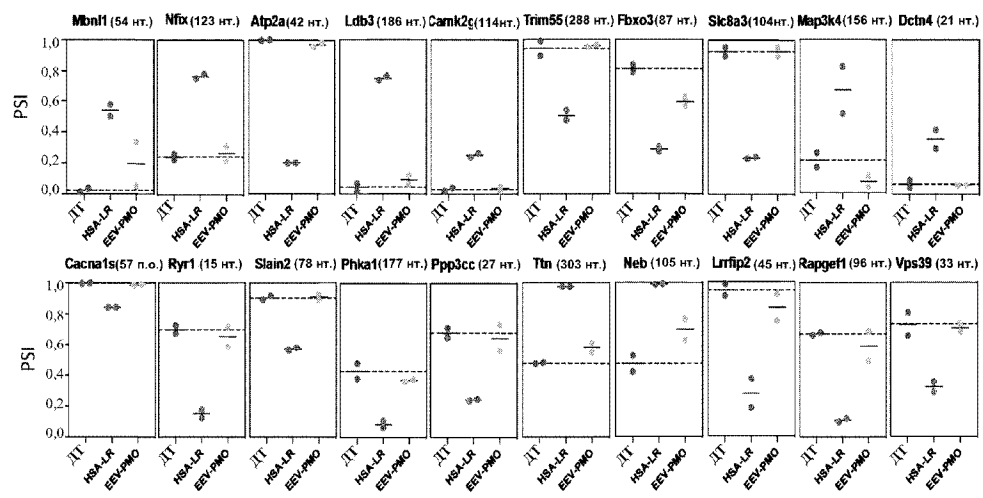


ФИГ. 23В

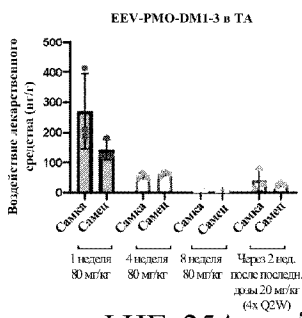




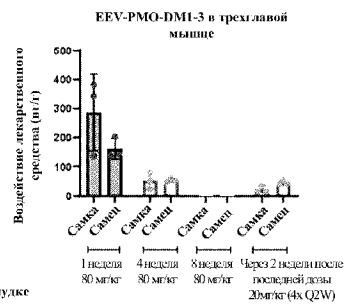
ФИГ. 23D



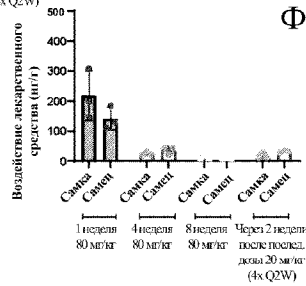
ФИГ. 24



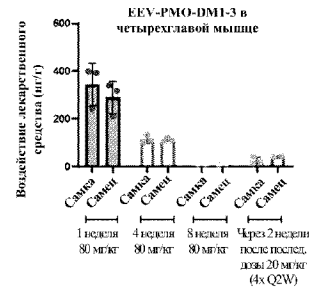
ФИГ. 25А



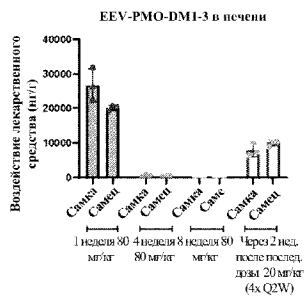
ФИГ. 25С



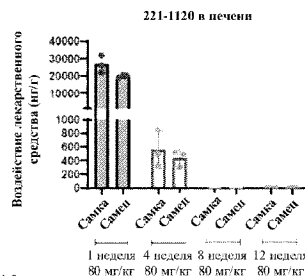
ФИГ. 25В



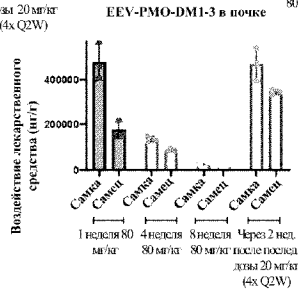
ФИГ. 25D



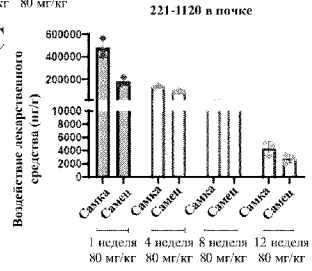
ФИГ. 26А



ФИГ. 26С

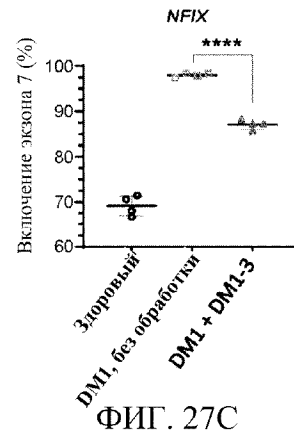
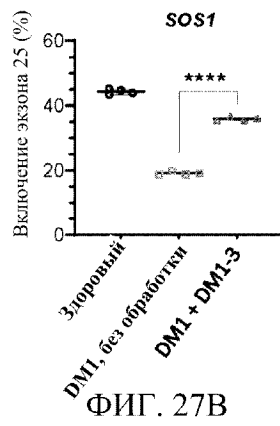
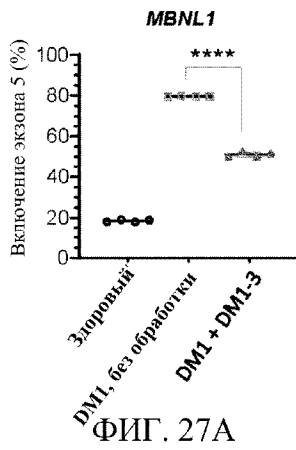


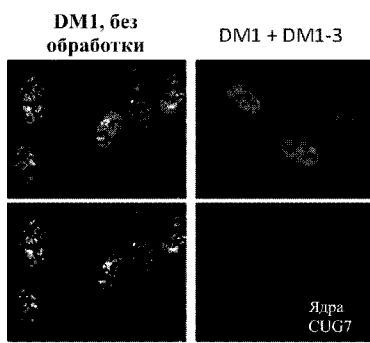
ФИГ. 26В



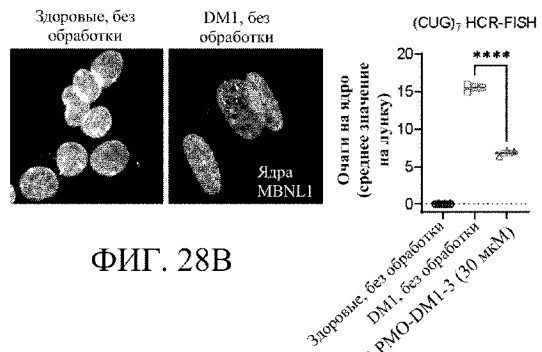
ФИГ. 26D





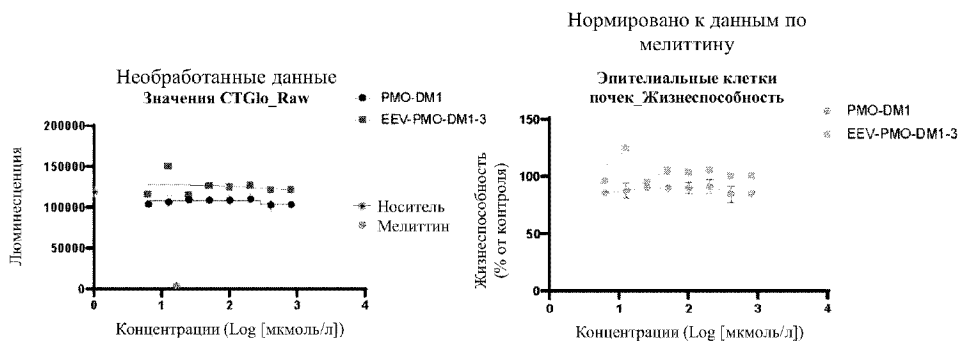


ФИГ. 28А



ФИГ. 28В

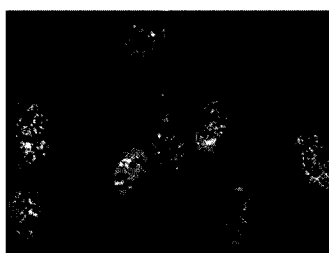
ФИГ. 28С



ФИГ. 29А

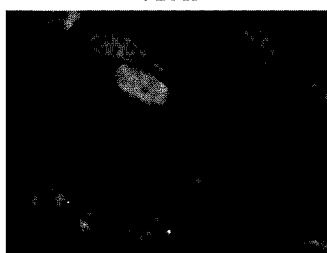
ФИГ. 29В

Необработанные клетки  
пациента с DM1

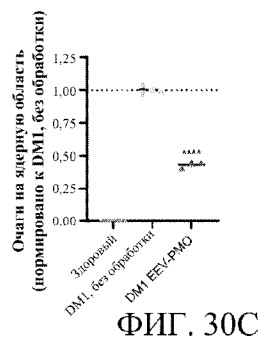


ФИГ. 30А

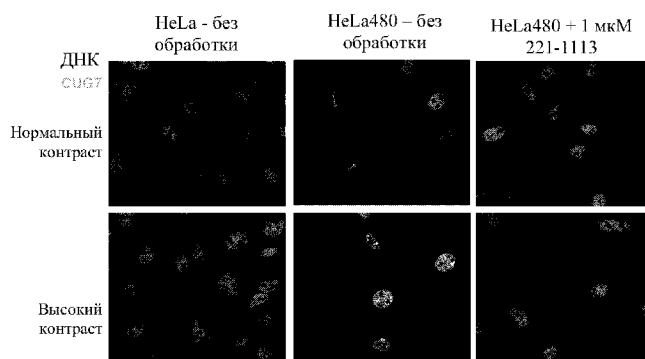
Обработанные клетки пациента  
с DM1



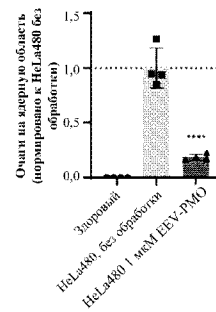
ФИГ. 30В



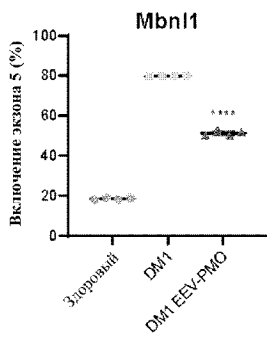
ФИГ. 30С



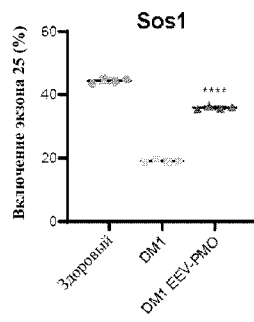
ФИГ. 31А



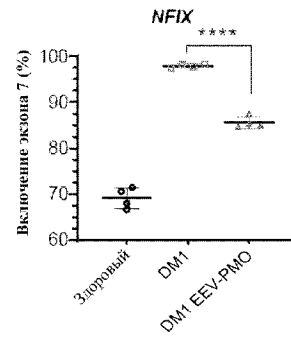
ФИГ. 31В



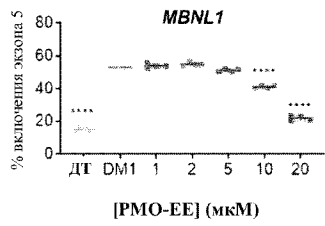
ФИГ. 32А



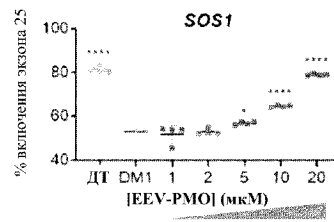
ФИГ. 32В



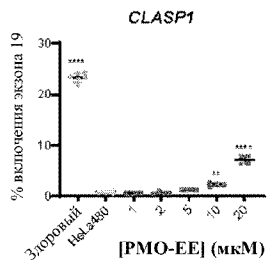
ФИГ. 32С



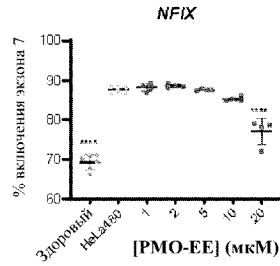
ФИГ. 33А



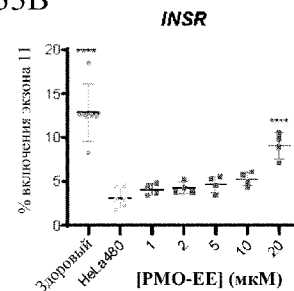
ФИГ. 33В



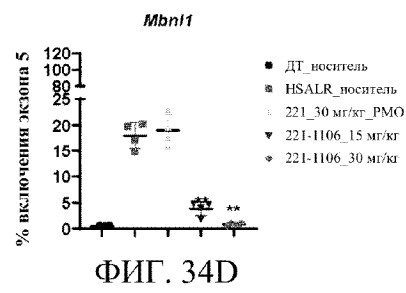
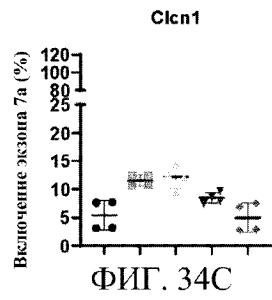
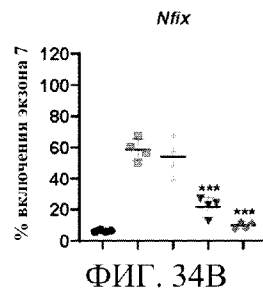
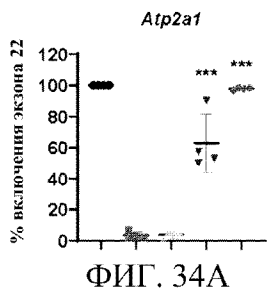
ФИГ. 33С



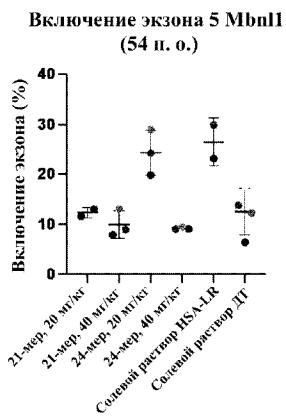
ФИГ. 33Д



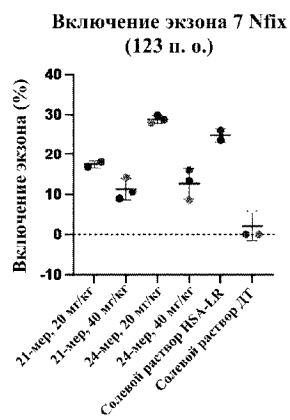
ФИГ. 33Е



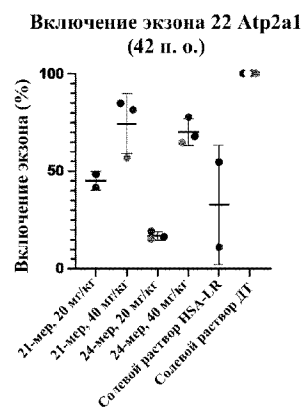




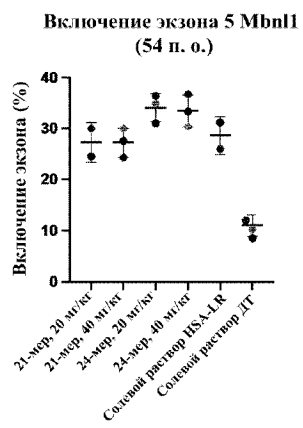
ФИГ. 35А



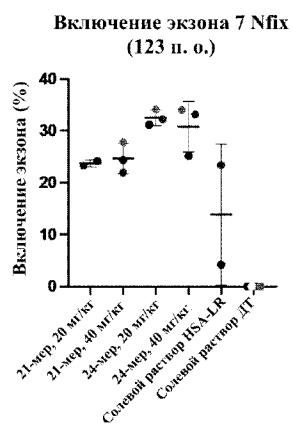
ФИГ. 35В



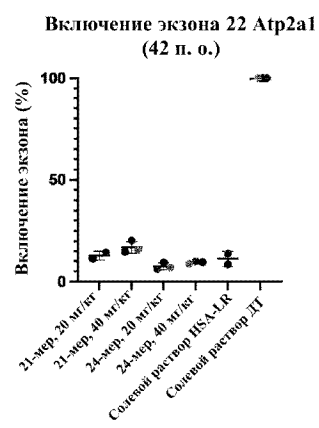
ФИГ. 35С



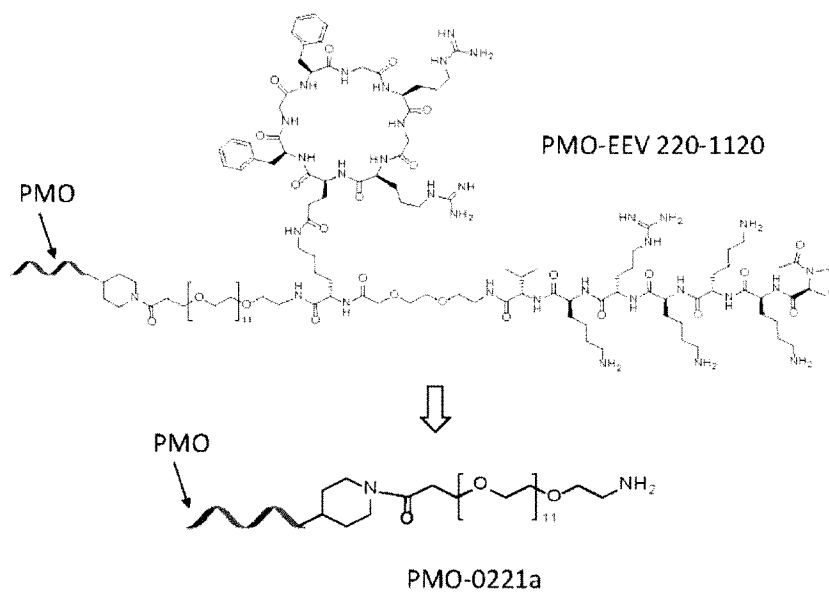
ФИГ. 36А



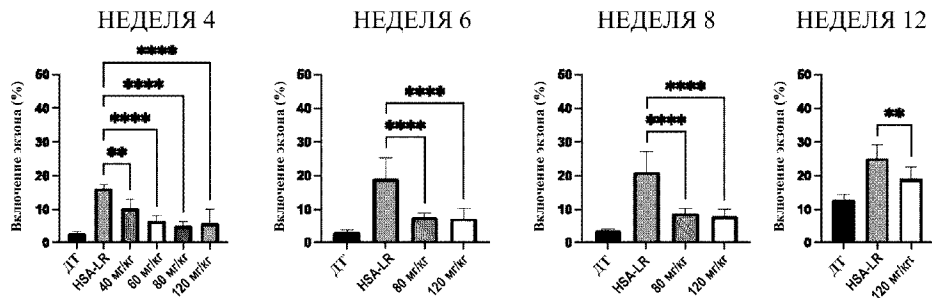
ФИГ. 36В



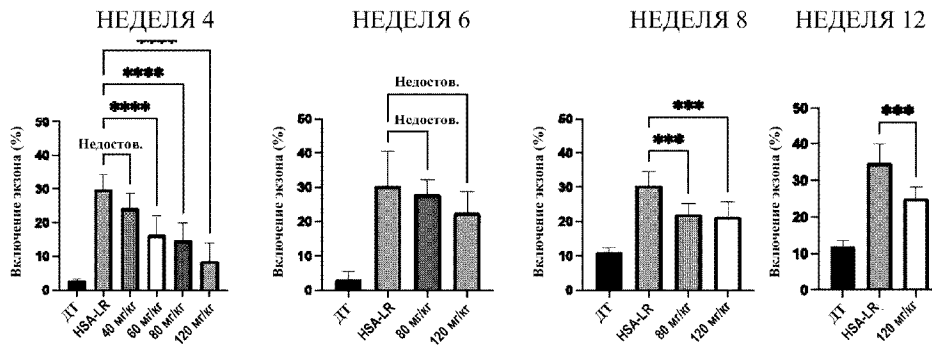
ФИГ. 36С



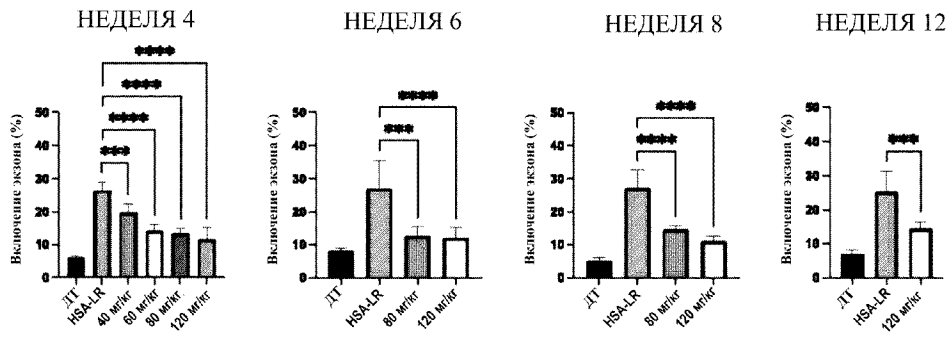
ФИГ. 37



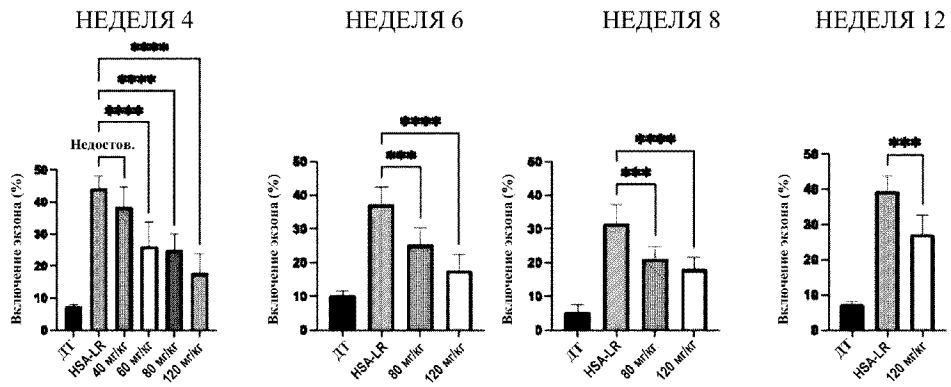
ФИГ. 38А



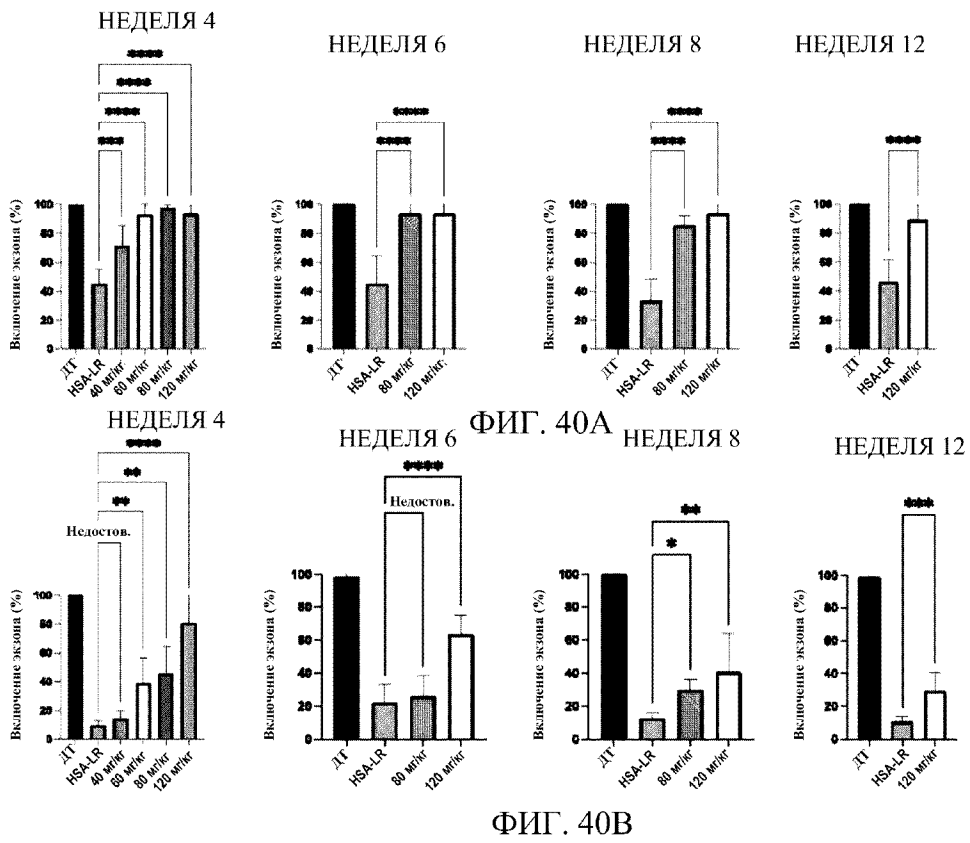
ФИГ. 38В



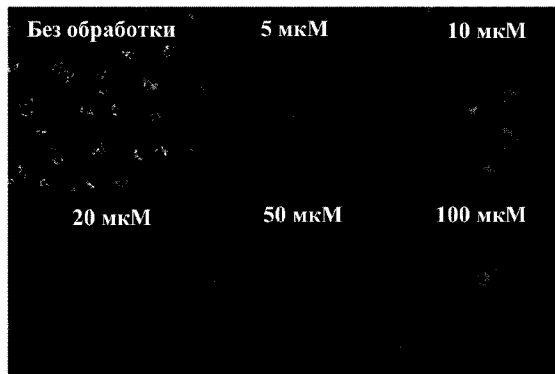
ФИГ. 39А



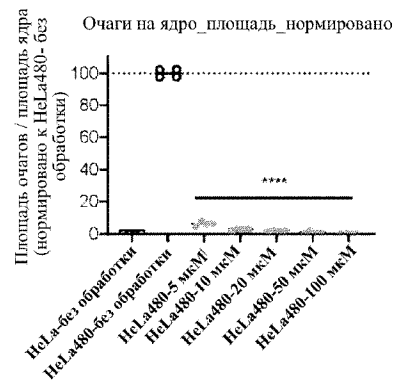
ФИГ. 39В



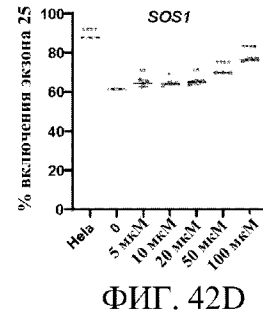
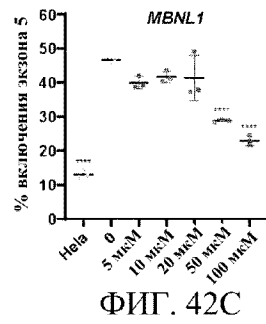
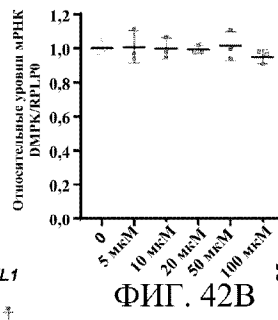
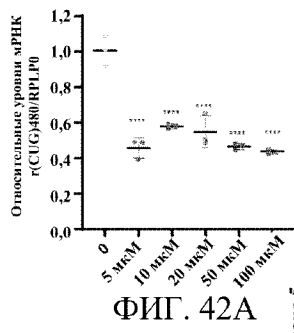
ФИГ. 40В



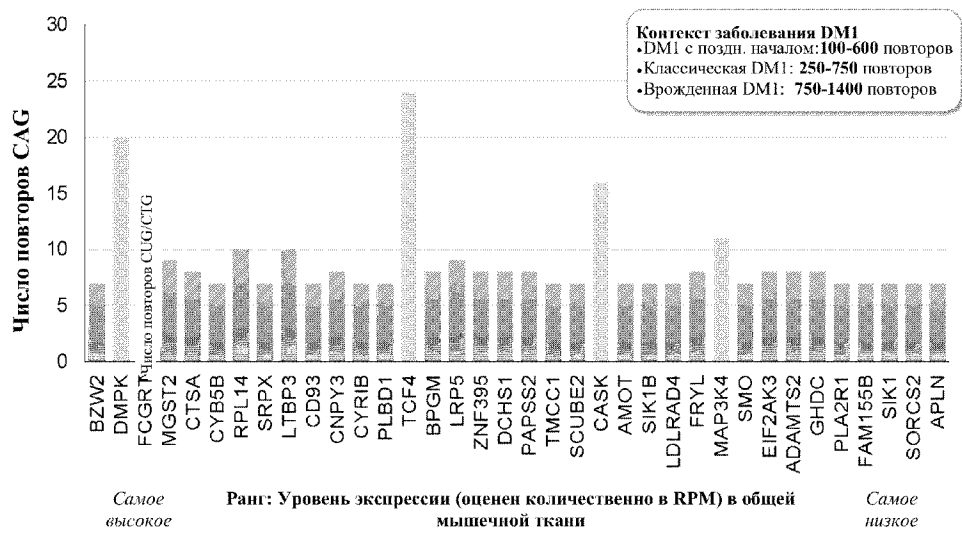
ФИГ. 41А



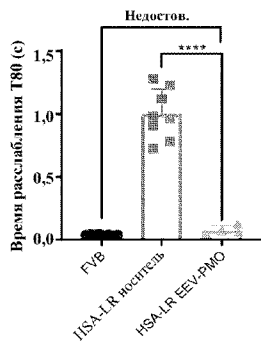
ФИГ. 41В



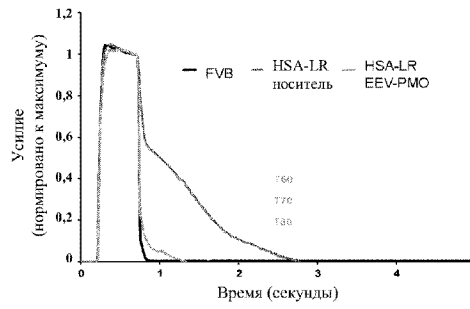




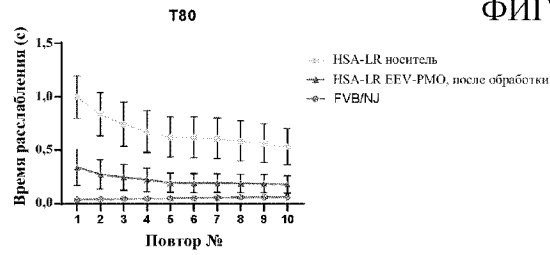
ФИГ. 43



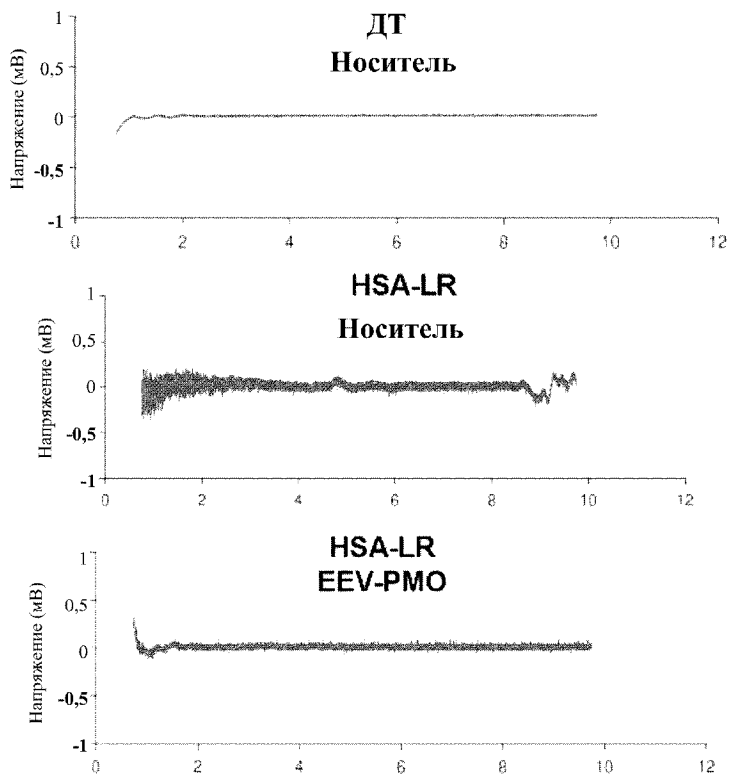
ФИГ. 44А



ФИГ. 44В



ФИГ. 44С



ФИГ. 44D