

(19)



Евразийское
патентное
ведомство

(21) 202291718 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.08.16

(51) Int. Cl. *A61K 31/7088* (2006.01)
C12N 5/10 (2006.01)
C12N 15/10 (2006.01)

(22) Дата подачи заявки
2020.09.29

(54) ЭУКАРИОТИЧЕСКИЕ ПОЛУСИНТЕТИЧЕСКИЕ ОРГАНИЗМЫ

(31) 62/908,421

(32) 2019.09.30

(33) US

(86) PCT/US2020/053339

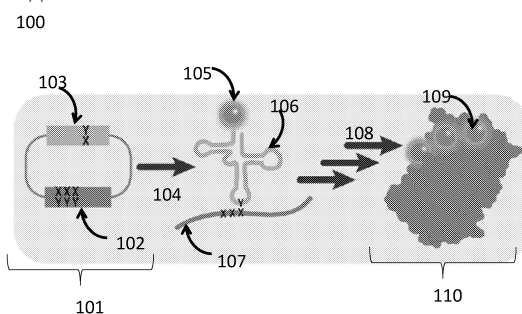
(87) WO 2021/067313 2021.04.08

(71) Заявитель:
ДЗЕ СКРИППС РИСЕРЧ
ИНСТИТУТ (US)

(72) Изобретатель:
Ромесберг Флойд И., Чжоу Энн
Сяочжоу, Шен Кай (US)

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

(57) В настоящем документе представлены эукариотические полусинтетические организмы и способы их использования и производства.



202291718 A1

202291718 A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-573669EA/085

ЭУКАРИОТИЧЕСКИЕ ПОЛУСИНТЕТИЧЕСКИЕ ОРГАНИЗМЫ

ПЕРЕКРЕСТНАЯ ССЫЛКА НА РОДСТВЕННЫЕ ЗАЯВКИ

По настоящей заявке испрашивается приоритет предварительной заявки США № 62/908,421, поданной 30 сентября 2019 г.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Настоящая заявка содержит Перечень последовательностей, который был представлен в электронном виде в формате ASCII и полностью включен в настоящий документ посредством ссылки. Указанная копия ASCII, созданная 24 сентября 2020 г., называется 36271-810_601_SL.txt и имеет размер 19000 байт.

ЗАЯВЛЕНИЕ О ФЕДЕРАЛЬНО СПОНСИРУЕМОМ ИССЛЕДОВАНИИ

Это изобретение было сделано при поддержке правительства по гранту № GM118178, выданному Национальными институтами здравоохранения (NIH). Правительство имеет определенные права на это изобретение.

УРОВЕНЬ ТЕХНИКИ

Каждый белок, когда-либо произведенный в клетке, кодируется четырехбуквенным генетическим алфавитом для двух пар оснований. Это обычно ограничивает аминокислоты, из которых могут быть построены белки, каноническими 20 протеогенными аминокислотами. Хотя это позволило разнообразие жизни, многие потенциальные функциональные возможности недоступны, и, таким образом, расширение за счет включения не канонических аминокислот (ncAA), в том числе тех, которые выбраны для обеспечения желаемой активности, может позволить создать новые белки с улучшенными свойствами для различных применений, от материалов до терапевтических агентов. Усилия по включению ncAA в основном основывались на расширении генетического алфавита за счет подавления стоп-кодона (UAG) или четырехбуквенного кодона (квадруплетные кодоны), хотя в этих случаях включение ncAA должно конкурировать с естественными функциями кодонов. Чтобы обойти это ограничение, усилия были сосредоточены на синтезе геномов с устраненными естественными стоп-кодонами или редкими кодонами, что освобождает их для переназначения на ncAA. Тем не менее, редкие кодоны потенциально могут играть важную роль в регуляции трансляции и свертывания белков, и синтез генома нецелесообразен в качестве общей стратегии, особенно с большими эукариотическими геномами.

Альтернативный подход основан на использовании пары оснований не природного происхождения (UBP), что, в принципе, с практической точки зрения позволило бы создать практически неограниченное количество новых совершенно новых кодонов, не обремененных какой-либо естественной функцией. Следуя примеру медицинской химии, было разработано семейство UBPs, типовым примером которого является dNaM-dTPT3 (фигура 1B), которые использовались в качестве основы полусинтетического организма (SSO) *E. coli*. SSO *E. coli* хранит UBPs в своем геноме или на плазмиде, транскрибирует его

в иРНК и тРНК, и с тРНК, заряженной ncAA с помощью ортогональной синтетазы, транслирует белки, содержащие ncAA. SSO *E. coli* имеет важное практическое применение, поскольку в настоящее время используется для получения новых терапевтических средств.

Широкий спектр ncAA и возникающих в результате полипептидов не природного происхождения, которые могут быть получены, диктуется, по меньшей мере, частично, используемым SSO. На сегодняшний день использование UBP, таких как dNAM-dTPT3, не было показано в эукариотических SSO или системе. Доказательство концепции подхода, кратко описанного в настоящем документе, в эукариотических клетках, позволит получать более широкий спектр ncAA и получающихся в результате полипептидов не природного происхождения, которые могут быть полезны для важных практических применений, таких как создание новых терапевтических средств.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

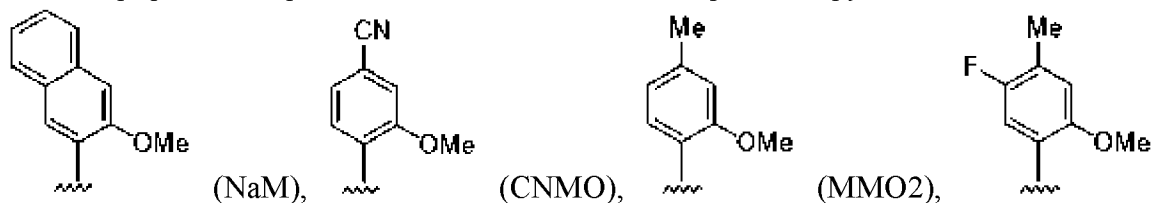
В некоторых вариантах осуществления, в настоящем документе представлены эукариотические полусинтетические организмы (SSO), которые были получены путем изучения трансляции не кодонов природного происхождения. Продуцирование белка характеризовали после прямой, временной, тройной трансфекции иРНК, содержащей кодон не природного происхождения, тРНК, содержащей родственный кодон не природного происхождения и ДНК, кодирующей подходящую синтетазу для зарядки тРНК не канонической аминокислотой (ncAA).

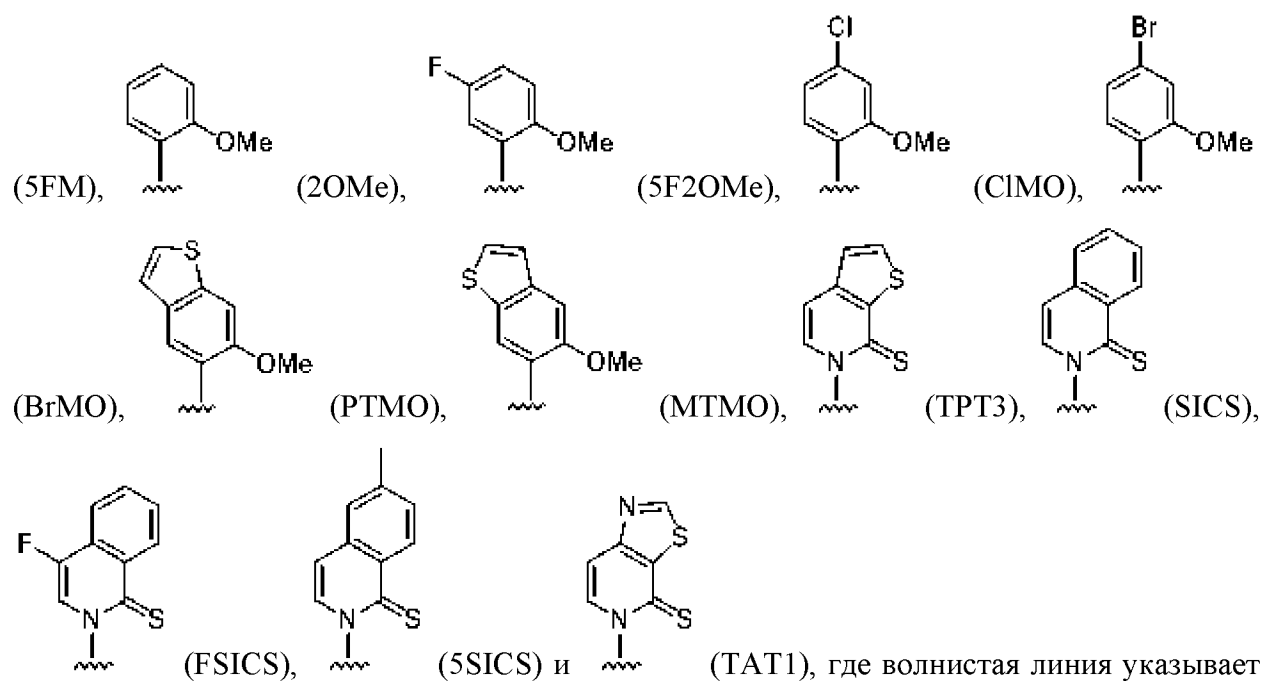
Аспекты, описанные в настоящем документе, представляют эукариотические клетки, содержащие (а) информационную РНК (иРНК) с кодоном, содержащим первое не основание природного происхождения, и (b) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения образуют пару оснований не природного происхождения (UBP) в эукариотической клетке, и где иРНК способна транслироваться в клетке с образованием полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения. В некоторых вариантах осуществления, тРНК заряжена аминокислотой не природного происхождения. В некоторых вариантах осуществления эукариотическая клетка дополнительно содержит полипептид, транслируемый из иРНК, где полипептид содержит, по меньшей мере, одну аминокислоту не природного происхождения. В некоторых вариантах осуществления, эукариотическая клетка дополнительно содержит рибосому, которая способна транслировать полипептид, содержащий, по меньшей мере, одну аминокислоту не природного происхождения из иРНК с использованием тРНК.

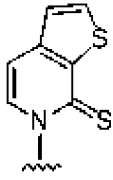
Аспекты, описанные в настоящем документе, также представляют эукариотические клетки, содержащие пару оснований не природного происхождения (UBP), содержащие: (а) первый рибонуклеотид не природного происхождения, содержащий первое основание не природного происхождения; (b) второй рибонуклеотид не природного происхождения, содержащий второе основание не природного происхождения, где первое и второе основания не природного происхождения образуют пару оснований не природного

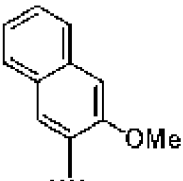
происхождения (UBP) в эукариотической клетке.

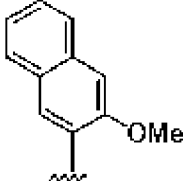
В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из: (i) 2-тиоурацила, 2-тиотимина, 2'-дезоксинуридина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-галорацила, 5-пропинилурацила, 6-азотимина, 6-азоурацила, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила, псевдоурацила, метилового эфира урацил-5-оксуксусной кислоты, урацил-5-оксуксусной кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, 5-метил-2-тиоурацила, 4-тиоурацила, 5-метилурацила, 5'-метоксикарбоксиметилурацила, 5-метоксиурацила, урацил-5-оксиуксусной кислоты, 5-(карбоксигидроксиметил)урацила, 5-карбоксиметиламинометил-2-тиуридина, 5-карбоксиметиламинометилурацила или дигидроурацила; (ii) 5-гидроксиметилцитозина, 5-трифторметилцитозина, 5-галоцитозина, 5-пропинилцитозина, 5-гидроксцитозина, циклоцитозина, цитозина арабинозида, 5,6-дигидроцитозина, 5-нитроцитозина, 6-азоцитозина, азацитозина, N4-этилцитозина, 3-метилцитозина, 5-метилцитозина, 4-ацетилцитозина, 2-тиоцитозина, феноксазина цитидина ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазина цитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), феноксазина цитидина (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазола цитидина (2H-пиримидо[4,5-b]индол-2-она) или пиридоиндола цитидина (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она); (iii) 2-аминоаденина, 2-пропиладенина, 2-аминоаденина, 2-F-аденина, 2-аминопропиладенина, 2-амино-2'-дезоксияденозина, 3-дезааденина, 7-метиладенина, 7-дезааденина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидрокси-замещенных аденинов, N6-изопентениладенина, 2-метиладенина, 2,6-диаминопурина, 2-метилтио-N6-изопентениладенина или 6-азааденина; (iv) 2-метилгуанина, 2-пропила и алкильных производных гуанина, 3-дезагуанина, 6-тиогуанина, 7-метилгуанина, 7-дезагуанина, 7-дезагуанозина, 7-деза-8-азагуанина, 8-азагуанина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидрокси-замещенных гуанинов, 1-метилгуанина, 2,2-диметилгуанина, 7-метилгуанина или 6-азагуанина и (v) гипоксантина, ксантина, 1-метилюридина, кеозина, бета-D-галактозилкеозина, инозина, бета-D-маннозилкеозина, вибутоксозина, гидроксимочевина, (аср3)w, 2-аминопиридина или 2-пиридона. В некоторых вариантах осуществления, каждое из первого основания не природного происхождения и второго основания не природного происхождения независимо выбрано из группы, состоящей из

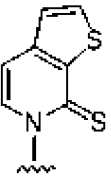


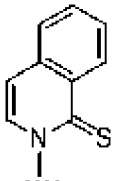


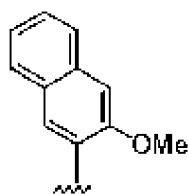
основание природного происхождения представляет собой  (TPT3), второе

основание не природного происхождения представляет собой  (NaM), и когда

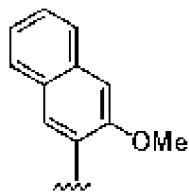
первое основание не природного происхождения представляет собой  (NaM),

второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой

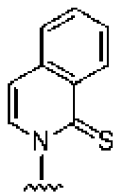
 (5SICS), второе основание не природного происхождения представляет собой



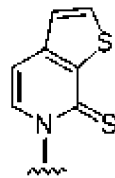
(NaM), и когда первое основание не природного происхождения



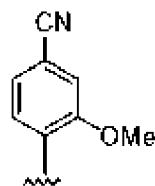
представляет собой (NaM), второе основание не природного происхождения



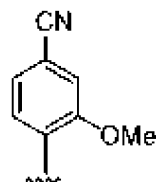
представляет собой (5SICS), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не



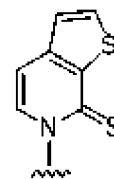
природного происхождения представляет собой (TPT3), второе основание не



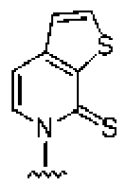
природного происхождения представляет собой (CNMO), и когда первое



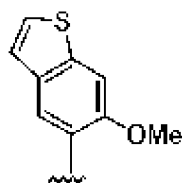
основание не природного происхождения представляет собой (CNMO), второе



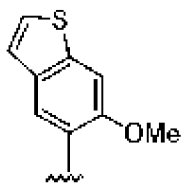
основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой



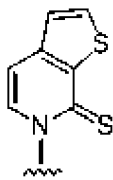
(TPT3), второе основание не природного происхождения представляет собой



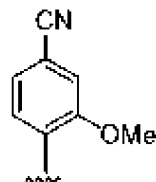
(PTMO), и когда первое основание не природного происхождения



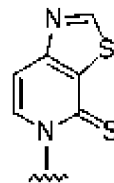
представляет собой (PTMO), второе основание не природного



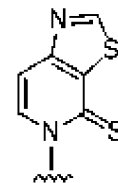
происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое



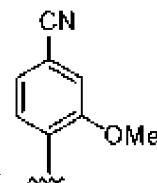
основание не природного происхождения представляет собой (CNMO), второе



основание не природного происхождения представляет собой (TAT1), и когда

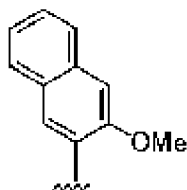


первое основание не природного происхождения представляет собой (TAT1),

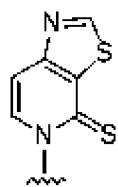


второе основание не природного происхождения представляет собой (CNMO),

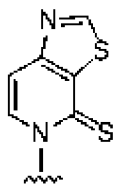
где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой



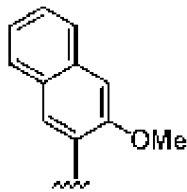
(NaM), второе основание не природного происхождения представляет собой



(TAT1), и когда первое основание не природного происхождения представляет



собой (TAT1), второе основание не природного происхождения представляет



собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения содержат модифицированную сахарную группу, выбранную из группы, состоящей из: модификации в положении 2':

ОН, замещенный низший алкил, алкарил, аралкил, О-алкарил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F;

О-алкил, S-алкил, N-алкил;

О-алкенил, S-алкенил, N-алкенил;

О-алкинил, S-алкинил, N-алкинил;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, где алкил, алкенил и алкинил могут быть замещены или не замещены C₁-C₁₀ алкилом, C₂-C₁₀ алкенилом, C₂-C₁₀ алкинилом, O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-NH₂ и -O(CH₂)_nON[(CH₂)_nCH₃]₂, где n и m составляют от 1 до примерно 10;

и/или модификации в положении 5':

5'-винил, 5'-метил (R или S);

модификации в положении 4':

4'-S, гетероциклоалкил, гетероциклоалкарил, аминоалкиламино, полиалкиламино, замещенный силлил, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для улучшения фармакодинамических свойств олигонуклеотида, и любая их комбинация.

В некоторых вариантах осуществления эукариотическая клетка дополнительно содержит: (a) транспортную РНК (тРНК) с антикодоном, содержащим первое основание не природного происхождения; (b) информационную РНК (иРНК) с кодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке. В некоторых вариантах осуществления эукариотическая клетка дополнительно содержит: (a) транспортную РНК (тРНК) с

антикодоном, содержащим второе основание не природного происхождения; (b) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке. В некоторых вариантах осуществления кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК. В некоторых вариантах осуществления, эукариотическая клетка дополнительно содержит полипептид, транслированный из иРНК, где полипептид содержит, по меньшей мере, одну аминокислоту не природного происхождения. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения: (a) представляет собой лизиновый аналог; (b) содержит ароматическую боковую цепь; (c) содержит азидогруппу; (d) содержит алкиновую группу; или (e) содержит альдегидную или кетонную группу. В некоторых вариантах осуществления, одну или несколько аминокислот не природного происхождения выбрано из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилфенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацетил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, O-аллилтирозина, O-метил-L-тирозина, O-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-O-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((азидоэтокси)-карбонил)-L-лизин (AzK). В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((2-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-

(((3-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((4-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, эукариотическая клетка представляет собой клетку человека. В некоторых вариантах осуществления, клетка человека представляет собой клетку НЕК293Т. В некоторых вариантах осуществления, клетка представляет собой клетку хомяка. В некоторых вариантах осуществления, клетка хомяка представляет собой клетку яичника китайского хомяка (СНО). В некоторых вариантах осуществления, иРНК и тРНК стабилизированы для деградации в эукариотической клетке.

Аспекты, описанные в настоящем документе, относятся к полусинтетическим организмам, содержащим эукариотическую клетку, описанную в настоящем документе.

Аспекты, описанные в настоящем документе, представляют линии эукариотических клеток, содержащие множество эукариотических клеток по настоящему изобретению.

Аспекты, описанные в настоящем документе, представляют способы получения полипептида, содержащего одну или несколько аминокислот не природного происхождения, в эукариотической клетке, включающие: (а) введение в клетку: (i) информационной РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; и (ii) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения в эукариотической клетке, где первое и второе содержащие основания не природного происхождения образуют пару оснований не природного происхождения (UBP) в эукариотической клетке; и (b) трансляцию полипептида, содержащего одну или несколько аминокислот не природного происхождения из иРНК с использованием тРНК. В некоторых вариантах осуществления, тРНК заряжена аминокислотой не природного происхождения.

Аспекты, описанные в настоящем документе, также относятся к способам получения полипептида, содержащего одну или несколько аминокислот не природного происхождения, в эукариотической клетке, включающим: (а) получение эукариотической клетки, содержащей: (i) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; (ii) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения, образуют пару оснований не природного происхождения (UBP) в эукариотической клетке; (b) трансляцию полипептида, содержащего одну или несколько аминокислот не природного происхождения из иРНК с использованием тРНК с помощью рибосомы, которая является эндогенной для эукариотической клетки. В некоторых вариантах осуществления, полипептид содержит эукариотический профиль гликозилирования. Профиль гликозилирования может соответствовать клетке, в которой он продуцируется (например, быть профилем гликозилирования млекопитающих, если клетка принадлежит млекопитающему, профилем гликозилирования человека, если клетка принадлежит человеку, и т.д.).

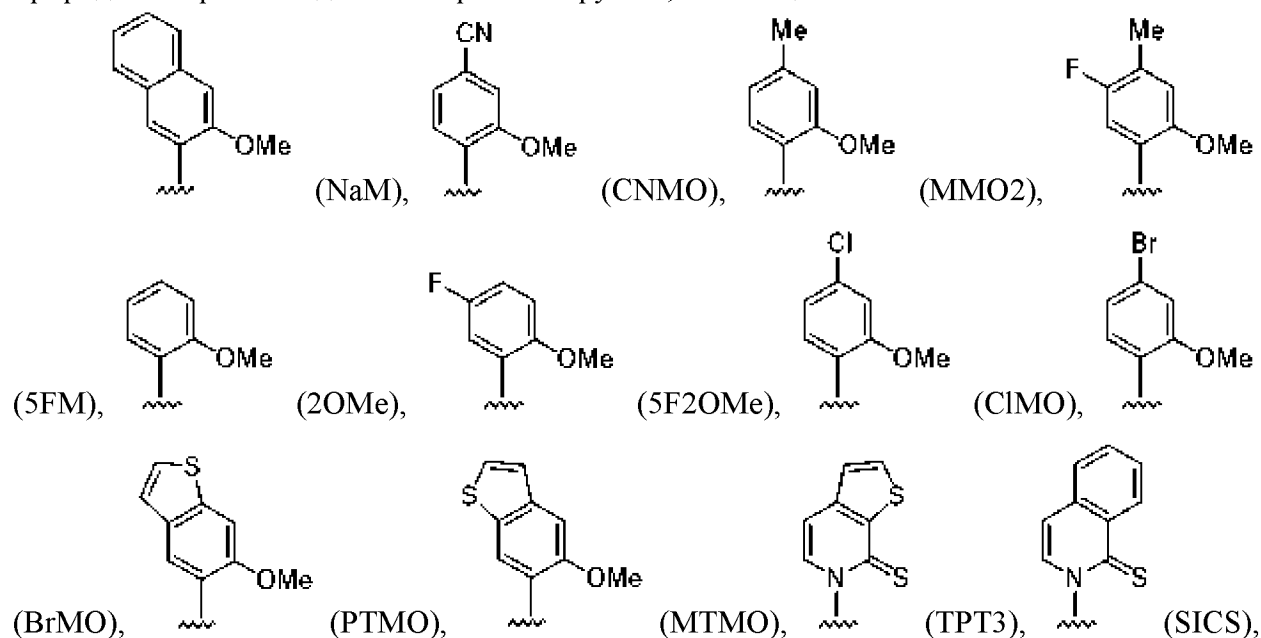
Аспекты, описанные в настоящем документе, также относятся к способам получения

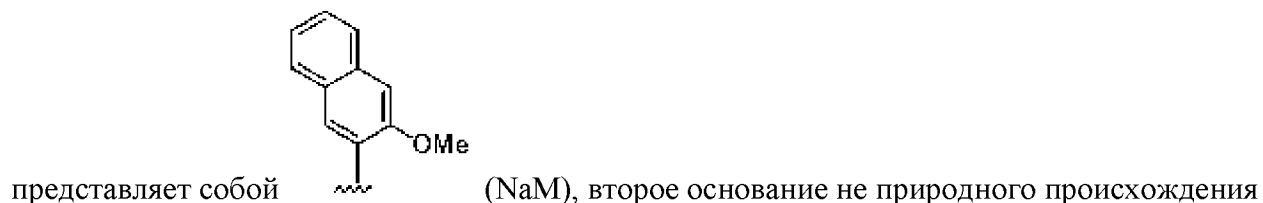
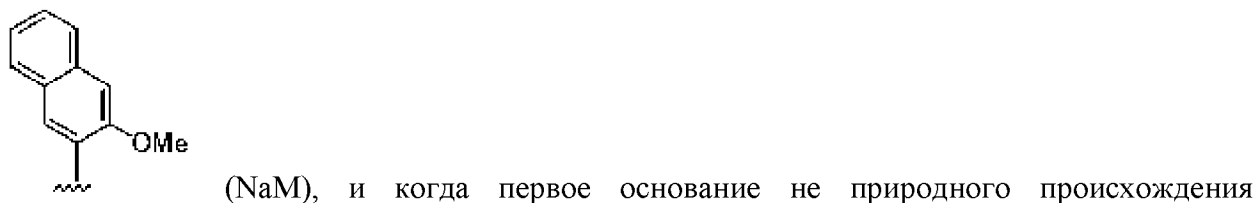
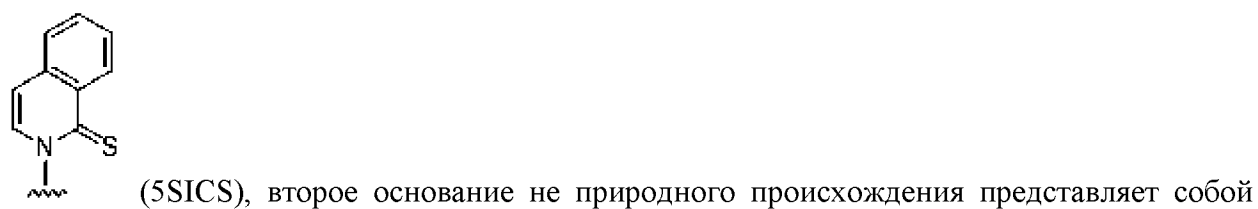
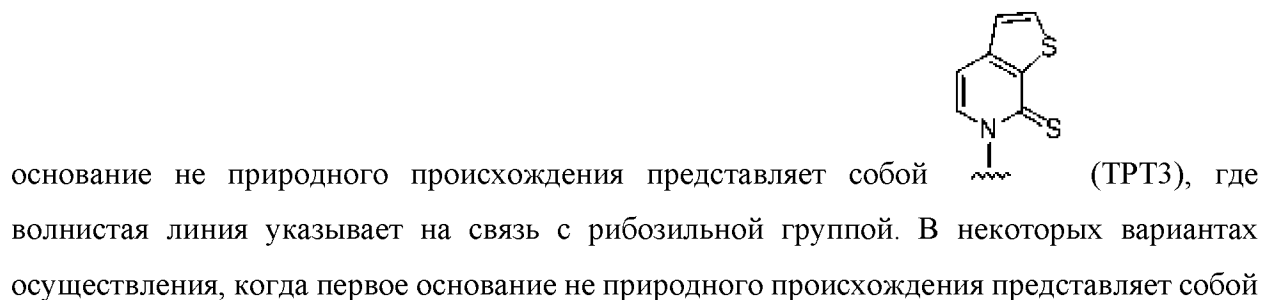
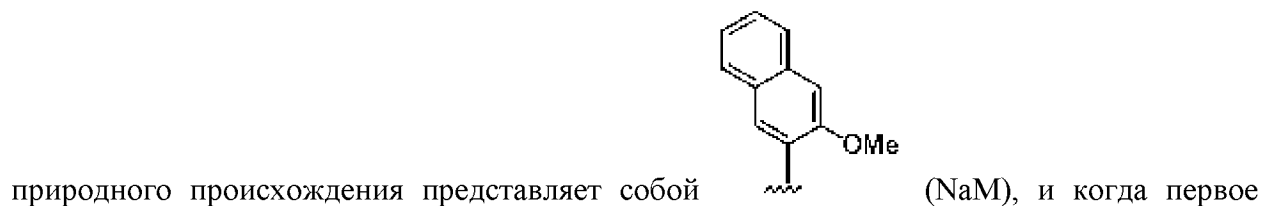
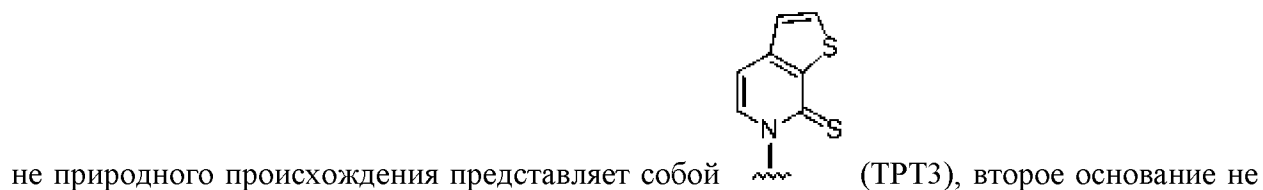
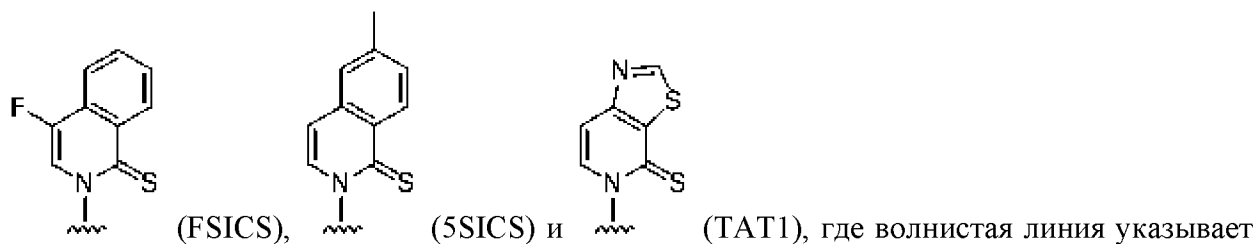
полипептида в эукариотической клетке, где полипептид содержит одну или несколько аминокислот не природного происхождения, где способ включает, где способ включает: (a) получение эукариотической клетки, где эукариотическая клетка содержит: (i) иРНК, содержащую кодон, где кодон содержит первое основание не природного происхождения; (ii) тРНК, содержащую антикодон, где антикодон содержит второе основание не природного происхождения, и где первое и второе основания не природного происхождения образуют комплементарную пару оснований и (iii) тРНК синтетазу, где тРНК синтаза предпочтительно аминоацилирует тРНК с помощью одной или нескольких аминокислот не природного происхождения по сравнению с природной аминокислотой; и (b) предоставление еще одной аминокислоты не природного происхождения в эукариотическую клетку, где эукариотическая клетка продуцирует полипептид, содержащий одну или несколько аминокислот не природного происхождения.

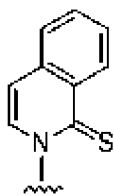
Аспекты, описанные в настоящем документе, также относятся к способам получения полипептида, содержащего одну или несколько аминокислот не природного происхождения, в эукариотической клетке, включающим: (a) получение эукариотической клетки, содержащей: (i) транспортную РНК (тРНК) с антикодоном, содержащим первое основание не природного происхождения; (ii) информационную РНК (иРНК) с кодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения образуют пару оснований не природного происхождения (UBP) в эукариотической клетке; и (c) трансляцию полипептида, содержащего одну или несколько аминокислот не природного происхождения из иРНК с использованием тРНК, через рибосому, которая является эндогенной для эукариотической клетки.

В некоторых вариантах осуществления, кодон мРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) находится в среднем положении (N-X-N) в кодоне иРНК. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК. В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из: (a) 2-тиоурацила, 2-тиотимина, 2'-дезоксинуридина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-галоурацила, 5-пропинилурацила, 6-азотимина, 6-азоурацила, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила, псевдоурацила, метилового эфира урацил-5-оксуксусной кислоты, урацил-5-оксуксусной кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, 5-метил-2-тиоурацила, 4-тиоурацила, 5-метилурацила, 5'-метоксикарбоксиметилурацила, 5-метоксинурацила, урацил-5-оксуксусной кислоты, 5-(карбоксигидроксиметил)урацила,

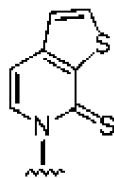
5-карбоксиметиламинометил-2-тиуридина, 5-карбоксиметиламинометилурацила или дигидроурацила; (b) 5-гидроксиметилцитозина, 5-трифторметилцитозина, 5-галоцитозина, 5-пропинилцитозина, 5-гидроксцитозина, циклоцитозина, цитозина арабинозида, 5,6-дигидроцитозина, 5-нитроцитозина, 6-азоцитозина, азацитозина, N4-этилцитозина, 3-метилцитозина, 5-метилцитозина, 4-ацетилцитозина, 2-тиоцитозина, феноксазина цитидина ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазина цитидина (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), феноксазина цитидина (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазола цитидина (2H-пиримидо[4,5-b]индол-2-она) или пиридоиндолцитидина (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она); (c) 2-аминоаденина, 2-пропиладенина, 2-аминоаденина, 2-F-аденина, 2-аминопропиладенина, 2-амино-2'-дезоксияденозина, 3-дезааденина, 7-метиладенина, 7-дезааденина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидрокси-замещенных аденинов, N6-изопентениладенина, 2-метиладенина, 2,6-диаминопурина, 2-метилтио-N6-изопентениладенина или 6-азааденина; (d) 2-метилгуанина, 2-пропила и алкильных производных гуанина, 3-дезагуанина, 6-тиогуанина, 7-метилгуанина, 7-дезагуанина, 7-дезагуанозина, 7-деза-8-азагуанина, 8-азагуанина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидрокси-замещенных гуанинов, 1-метилгуанина, 2,2-диметилгуанина, 7-метилгуанина или 6-азагуанина и (e) гипоксантина, ксантина, 1-метилюозина, кеозина, бета-D-галактозилкеозина, инозина, бета-D-маннозилкеозина, вибутоксозина, гидроксимочевина, (аср3)w, 2-аминопиридина или 2-пиридона. В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из



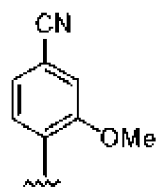




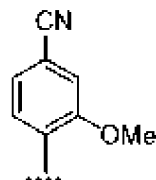
представляет собой (5SICS), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не



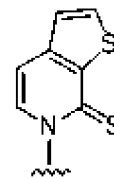
природного происхождения представляет собой (TPT3), второе основание не



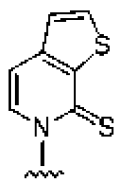
природного происхождения представляет собой (CNMO), и когда первое



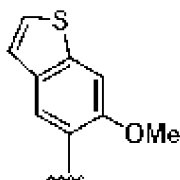
основание не природного происхождения представляет собой (CNMO), второе



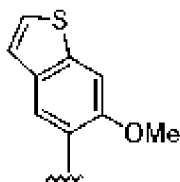
основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой



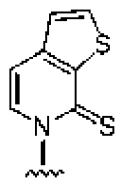
(TPT3), второе основание не природного происхождения представляет собой



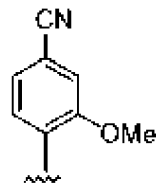
(PTMO), и когда первое основание не природного происхождения



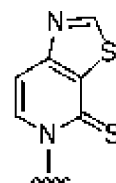
представляет собой (PTMO), второе основание не природного



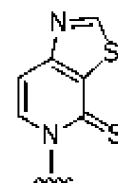
происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое



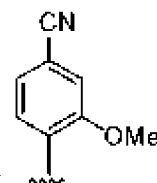
основание не природного происхождения представляет собой (CNMO), второе



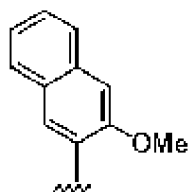
основание не природного происхождения представляет собой (TAT1), и когда



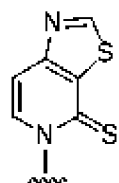
первое основание не природного происхождения представляет собой (TAT1),



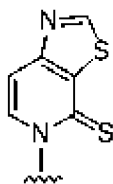
второе основание не природного происхождения представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой



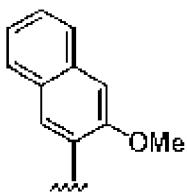
(NaM), второе основание не природного происхождения представляет собой



(TAT1), и когда первое основание не природного происхождения представляет



собой (TAT1), второе основание не природного происхождения представляет



собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения содержит модифицированную сахарную группу, выбранную из группы, состоящей из: модификации в положении 2':

ОН, замещенный низший алкил, алкарил, аралкил, О-алкарил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F;

О-алкил, S-алкил, N-алкил;

О-алкенил, S-алкенил, N-алкенил;

О-алкинил, S-алкинил, N-алкинил;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, где алкил, алкенил и алкинил могут быть замещены или не замещены C₁-C₁₀ алкилом, C₂-C₁₀ алкенилом, C₂-C₁₀ алкинилом, O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-NH₂ и -O(CH₂)_nON[(CH₂)_nCH₃]₂, где n и m составляют от 1 до примерно 10;

и/или модификация в положении 5':

5'-винил, 5'-метил (R или S);

модификация в положении 4':

4'-S, гетероциклоалкил, гетероциклоалкарил, аминокламино, полиалкиламино, замещенный силлил, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для улучшения фармакодинамических свойств олигонуклеотида, и любая их комбинация.

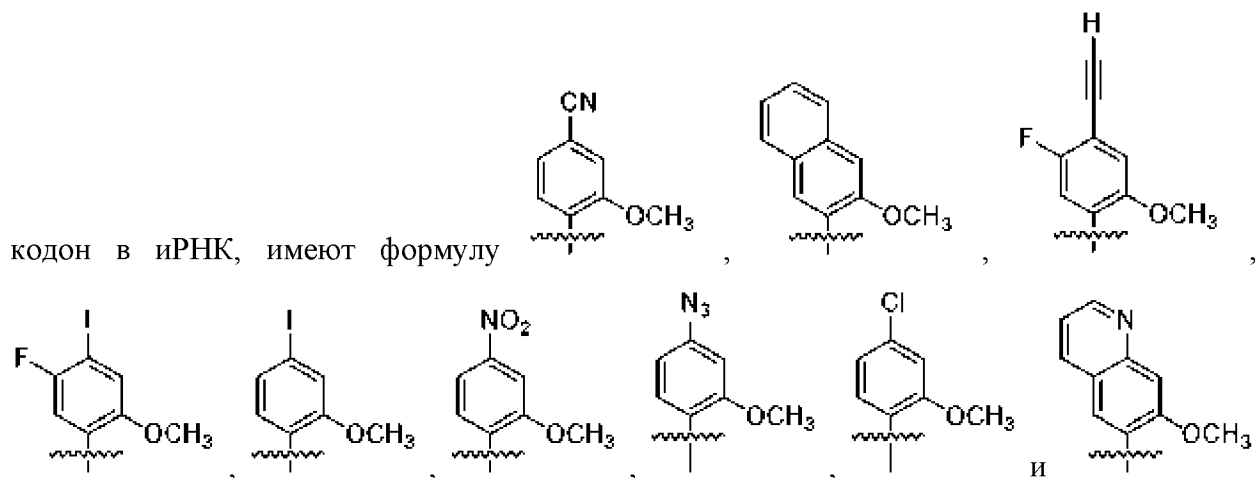
В некоторых вариантах осуществления, эукариотическая клетка представляет собой клетку человека. В некоторых вариантах осуществления, клетка человека представляет собой клетку НЕК293Т. В некоторых вариантах осуществления, клетка представляет собой клетку хомяка. В некоторых вариантах осуществления, клетка хомяка представляет собой клетку яичника китайского хомяка (СНО). В некоторых вариантах осуществления, аминокислота не природного происхождения: (а) представляет собой аналог лизина; (b) содержит ароматическую боковую цепь; (с) содержит азидогруппу; (d) содержит алкиновую группу; или (е) содержит альдегидную или кетонную группу. В некоторых вариантах осуществления, аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, ТСО-лизина, метилтетразинлизина, аллилоксикарбонилизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина, изопропил-L-фенилаланина,

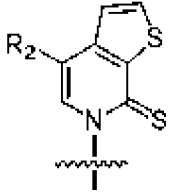
п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонотигозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK). В некоторых вариантах осуществления, одна или несколько аминокислот не природного происхождения представляет собой N6-(((2-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, одна или несколько аминокислот не природного происхождения представляют собой N6-(((3-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, одна или несколько аминокислот не природного происхождения представляют собой N6-(((4-азидобензил)окси)карбонил)-L-лизин.

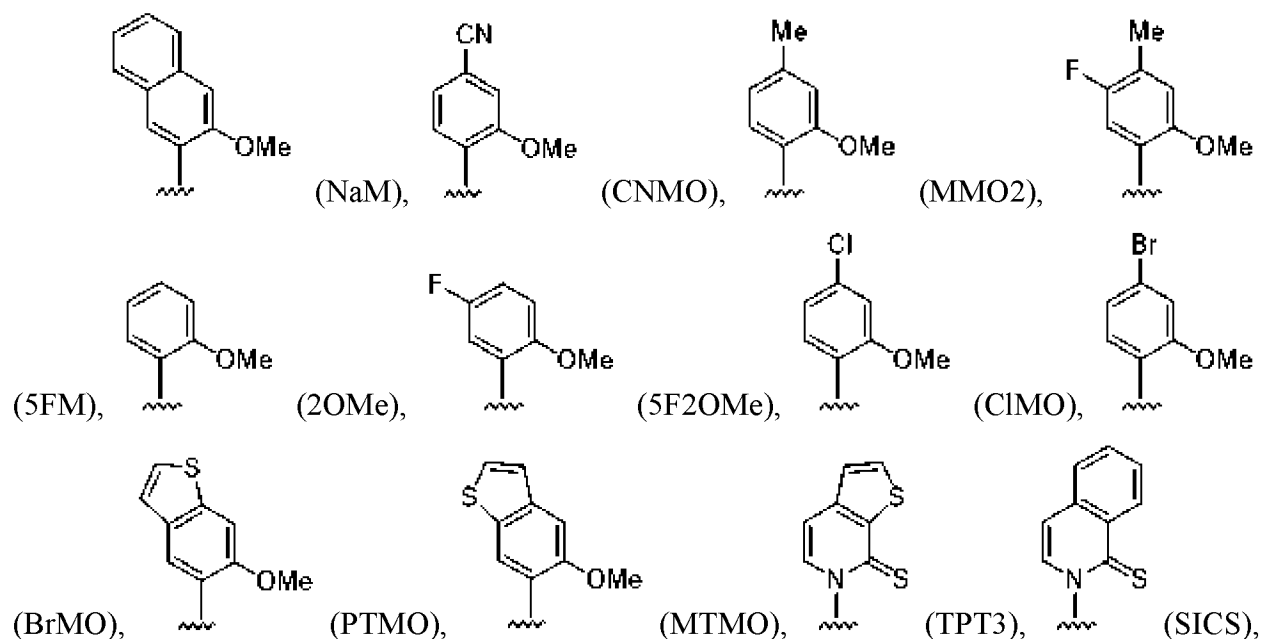
Аспекты, описанные в настоящем документе, относятся к способам получения полипептида в эукариотической клетке, где полипептид содержит одну или несколько аминокислот не природного происхождения, где способ включает: (a) получение эукариотической клетки, где эукариотическая клетка содержит: (i) иРНК, содержащую кодон, где кодон содержит одно или несколько оснований не природного происхождения; (ii) тРНК, содержащую антикодон, где антикодон содержит одно или несколько оснований не природного происхождения, и где одно или несколько оснований не природного происхождения, содержащих кодон в иРНК, и одно или несколько оснований не природного происхождения, содержащих антикодон в тРНК, образуют комплементарную пару оснований и (iii) тРНК-синтетазу, где тРНК-синтетаза предпочтительно аминоацилирует тРНК с помощью одной или нескольких аминокислот не природного происхождения по сравнению с природной аминокислотой; и (b) предоставление одной или нескольких аминокислот не природного происхождения эукариотической клетке, где эукариотическая клетка продуцирует полипептид, содержащий одну или несколько аминокислот не природного происхождения. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК. В

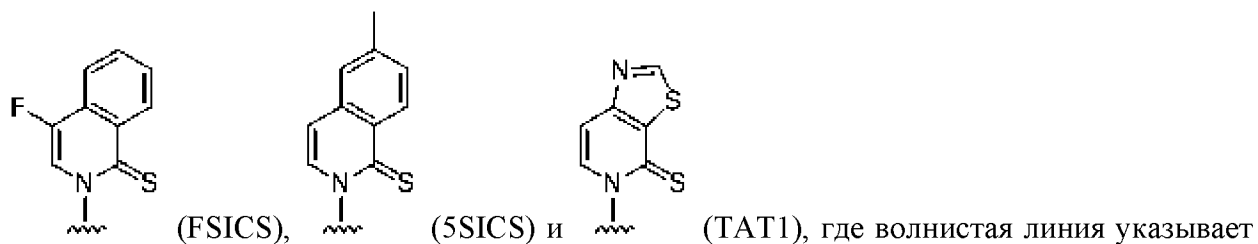
некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК. В некоторых вариантах осуществления, одно или несколько оснований не природного происхождения, содержащих

кодон в иРНК, имеют формулу

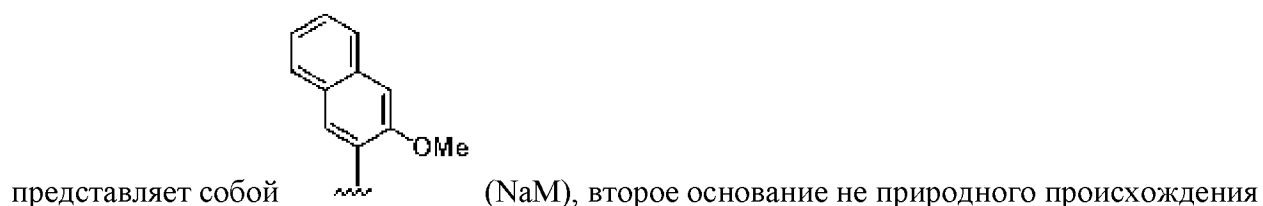
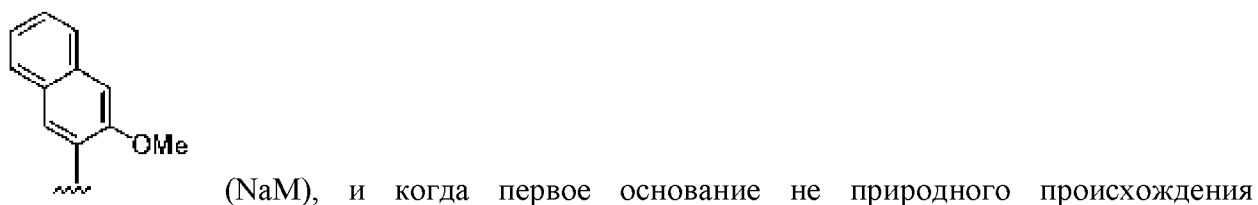
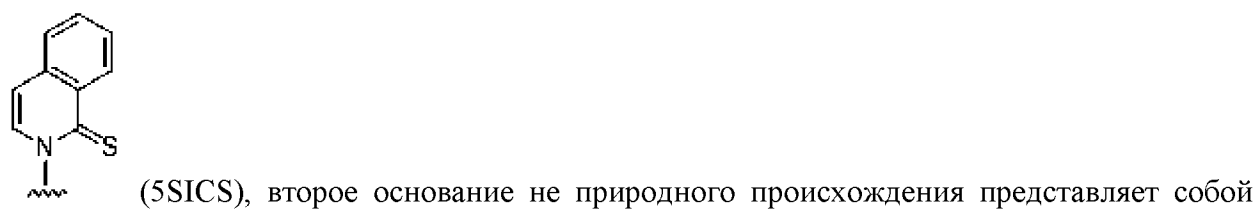
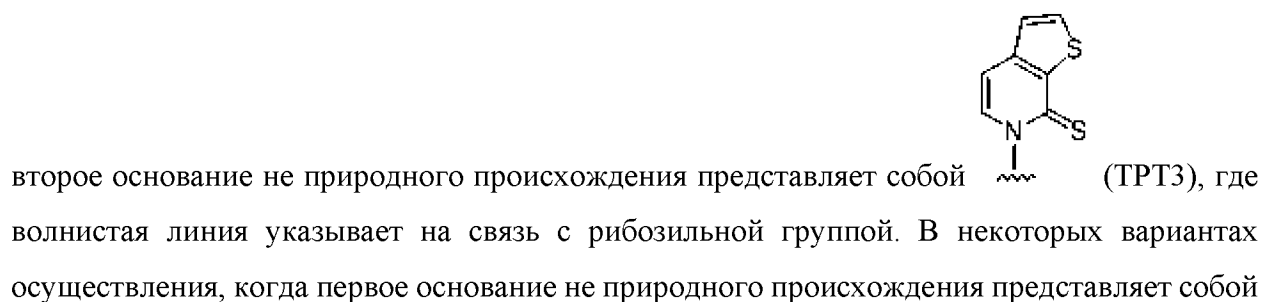
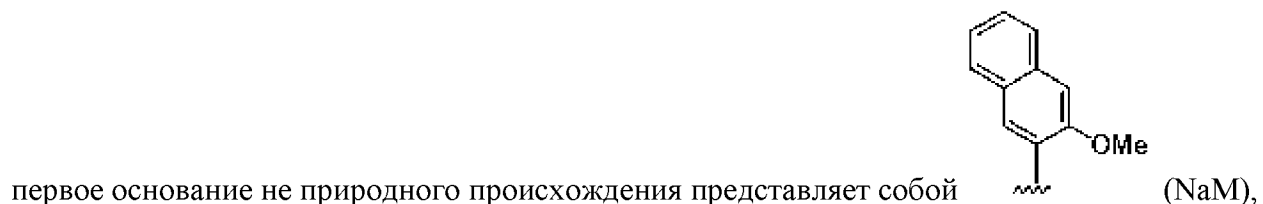


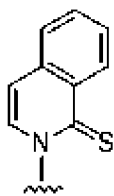
или , где R₂ выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азидо, и волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из



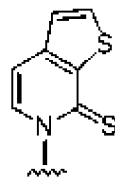


В некоторых вариантах осуществления, когда первое

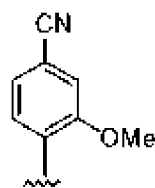




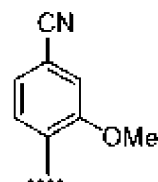
представляет собой (5SICS), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не



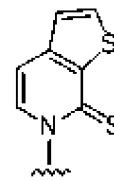
природного происхождения представляет собой (TPT3), второе основание не



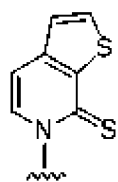
природного происхождения представляет собой (CNMO), и когда первое



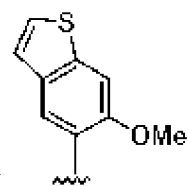
основание не природного происхождения представляет собой (CNMO), второе



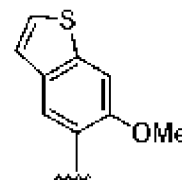
основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения представляет собой



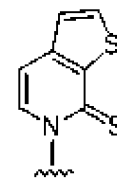
(TPT3), второе основание не природного происхождения



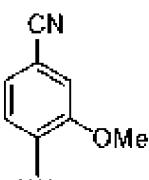
(PTMO), и

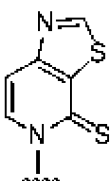


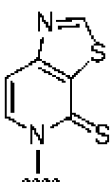
когда первое основание не природного происхождения представляет собой

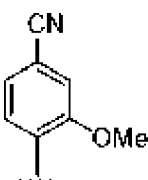


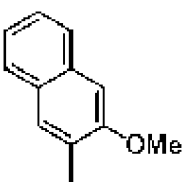
(PTMO), второе основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения

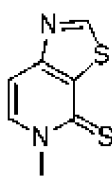
представляет собой  (CNMO), второе основание не природного происхождения

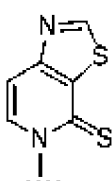
представляет собой  (TAT1), и когда первое основание не природного

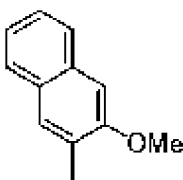
происхождения представляет собой  (TAT1), второе основание не природного

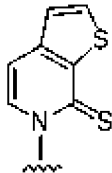
происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое

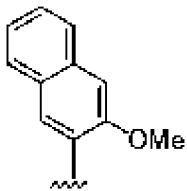
основание не природного происхождения представляет собой  (NaM), второе

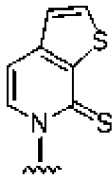
основание не природного происхождения представляет собой  (TAT1), и когда

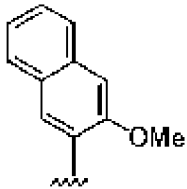
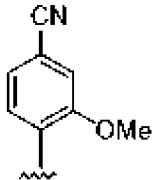
первое основание не природного происхождения представляет собой  (TAT1),

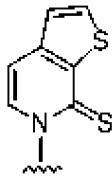
второе основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда первое основание не природного происхождения

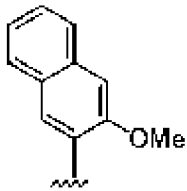
представляет собой  (TPT3), и второе основание не природного происхождения

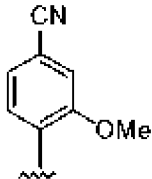
представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, где нуклеотид не

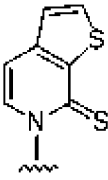
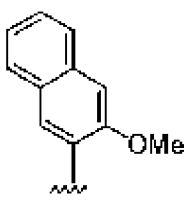
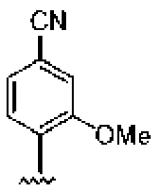
природного происхождения, содержащий кодон в иРНК, выбран из  (TPT3),

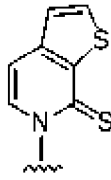
 (NaM) и  (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, нуклеотид не природного

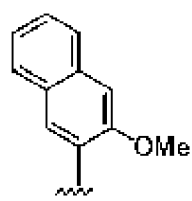
происхождения, содержащий кодон в иРНК, представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, нуклеотид не природного происхождения, содержащий кодон в иРНК,

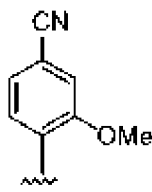
представляет собой  (NaM), где волнистая линия указывает В некоторых вариантах осуществления, нуклеотид не природного происхождения, содержащий кодон в

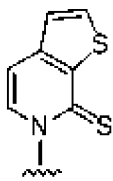
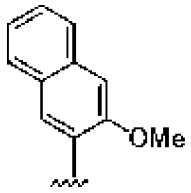
иРНК, представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание (X) расположено в первом положении (X-N-N) в кодоне иРНК, где основание не природного происхождения выбрано

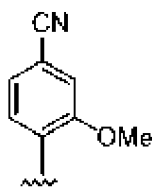
из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильную группу. В некоторых вариантах осуществления,

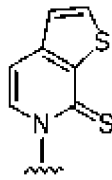
основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения представляет собой

 (NaM), где волнистая линия указывает на связь с рибозильной группой, В некоторых вариантах осуществления, основание не природного происхождения

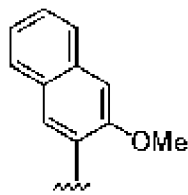
представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК, где

основание не природного происхождения выбрано из  (TPT3),  (NaM)

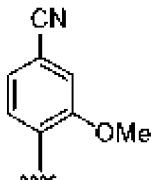
и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения

представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной

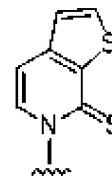
группой. В некоторых вариантах осуществления, основание не природного происхождения



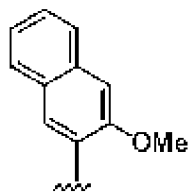
представляет собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного



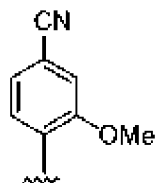
происхождения представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне



иРНК, где основание не природного происхождения выбрано из (TPT3),

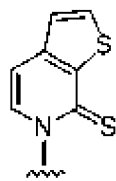


(NaM) и

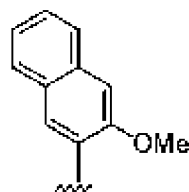


(CNMO), и где волнистая линия указывает на связь с

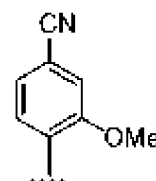
рибозильной группой. В некоторых вариантах осуществления, основание не природного



происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не

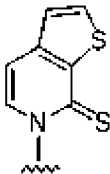
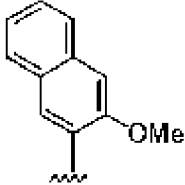


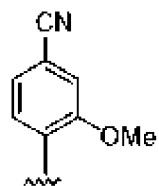
природного происхождения представляет собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления,



основание не природного происхождения представляет собой (CNMO), где

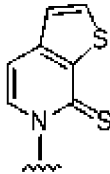
волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N) и где первое основание не природного происхождения (**X**) расположено в первом положении (**X**-N-N) в антикодоне тРНК. В некоторых вариантах осуществления, основание

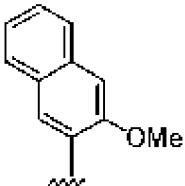
не природного происхождения выбрано из  (TPT3),  (NaM) и

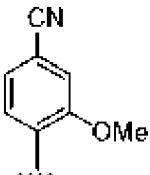


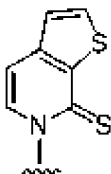
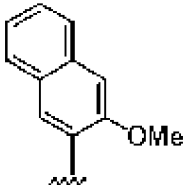
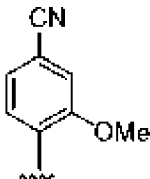
(CNMO), и где волнистая линия указывает на связь с рибозильной группой. В

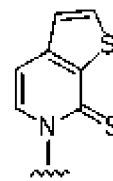
некоторых вариантах осуществления, основание не природного происхождения

представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения

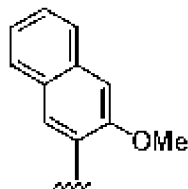
представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного

происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (**X**) расположено в среднем положении (N-**X**-N) в антикодоне тРНК. В некоторых вариантах осуществления, основание не природного происхождения

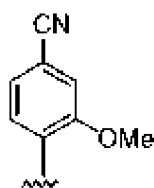
выбрано из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления,



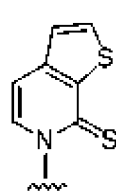
основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения представляет собой



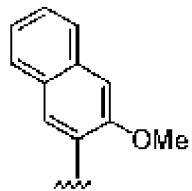
(NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения



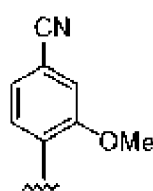
представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в антикодоне тРНК. В некоторых вариантах осуществления, основание не природного происхождения выбрано из



(TPT3),

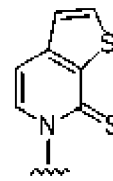


(NaM) и

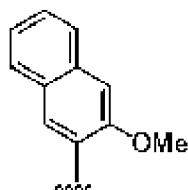


(CNMO), и где волнистая линия

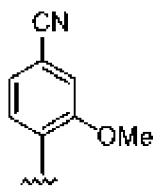
указывает на связь с рибозильной группой. В некоторых вариантах осуществления,



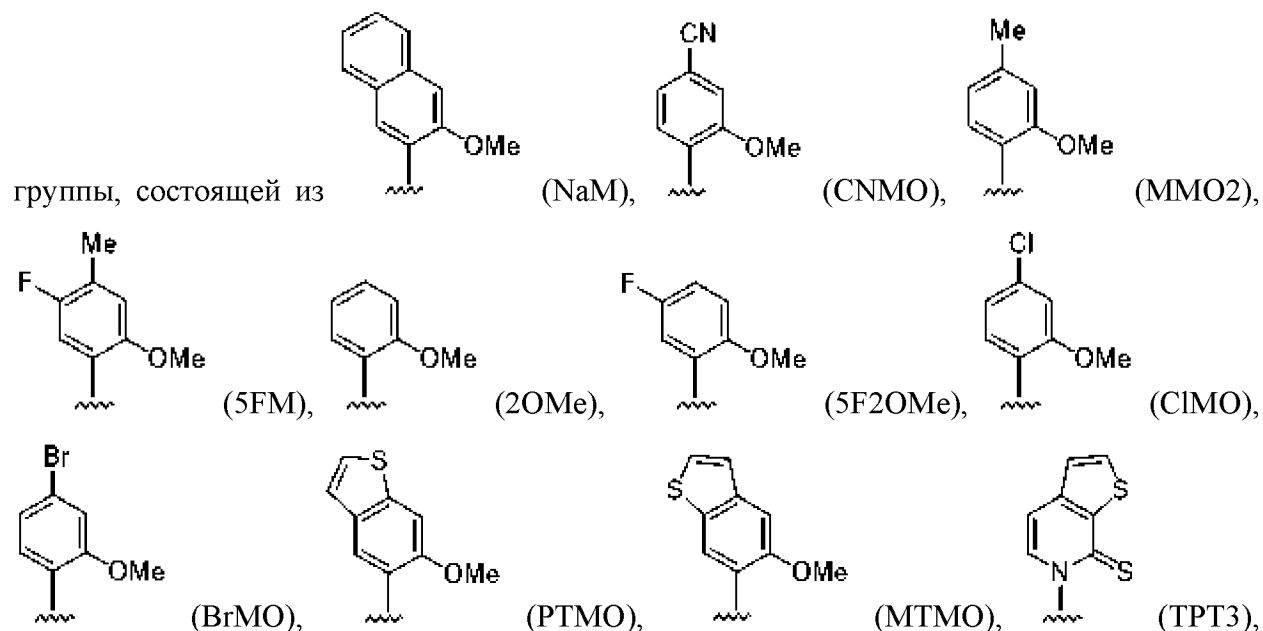
основание не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения представляет собой

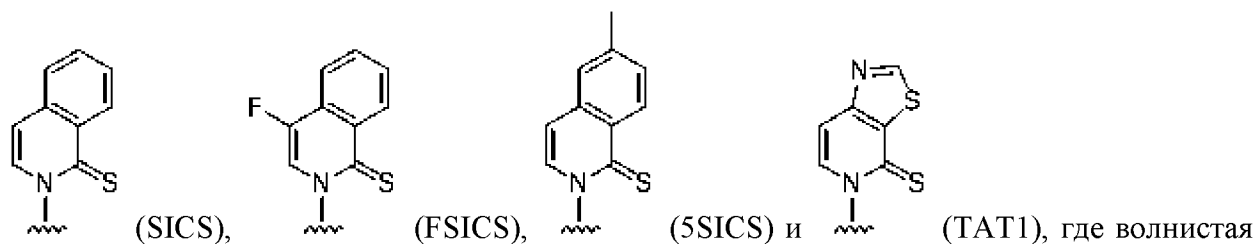


(NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения

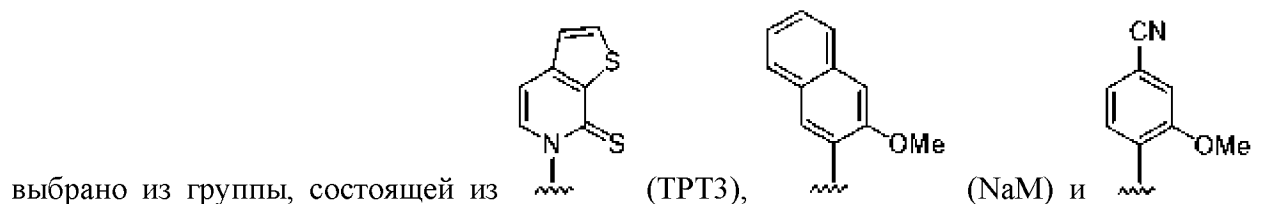


представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в первом положении (X-N-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в последнем положении (N-N-Y) антикодона. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК являются одинаковыми или разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК. тРНК выбраны из

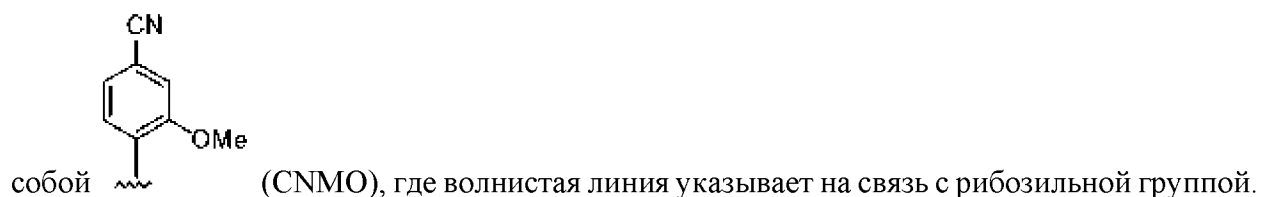
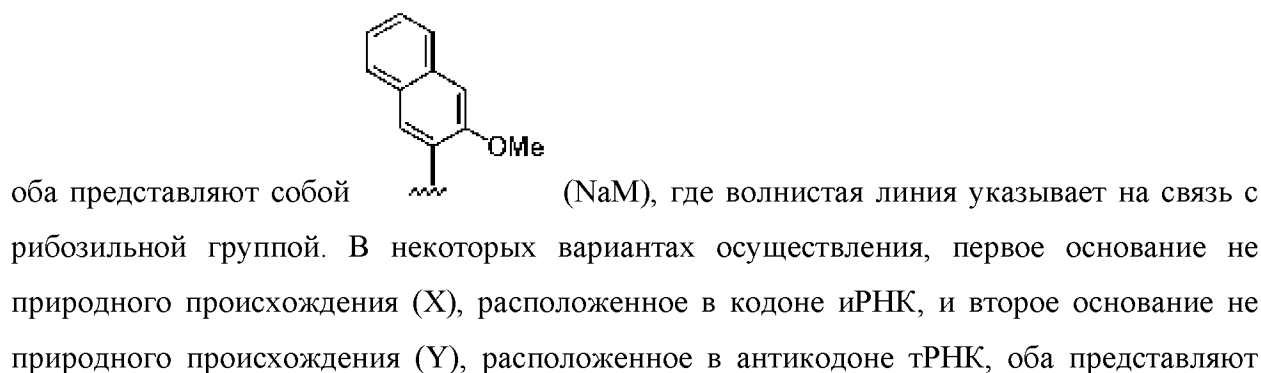
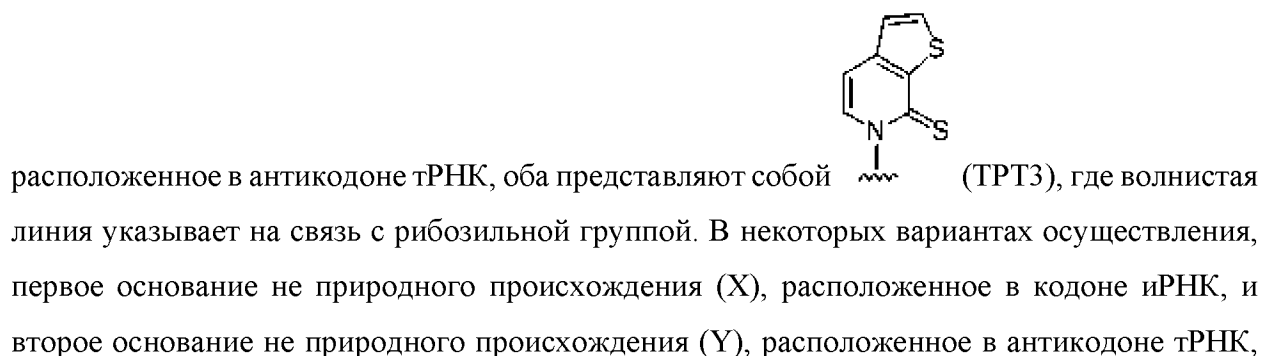




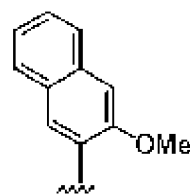
линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК,



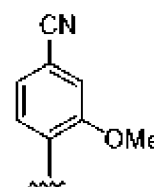
вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y),



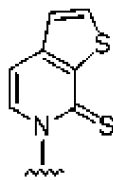
В некоторых вариантах осуществления, первое основание не природного происхождения



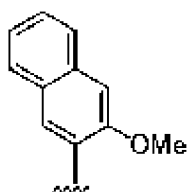
(NaM)



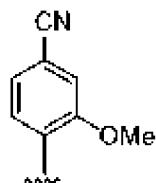
(X), расположенное в кодоне иРНК, выбрано из (NaM) и (CNMO), и второе основание не природного происхождения (Y), расположенное в



антикодоне тРНК, представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет

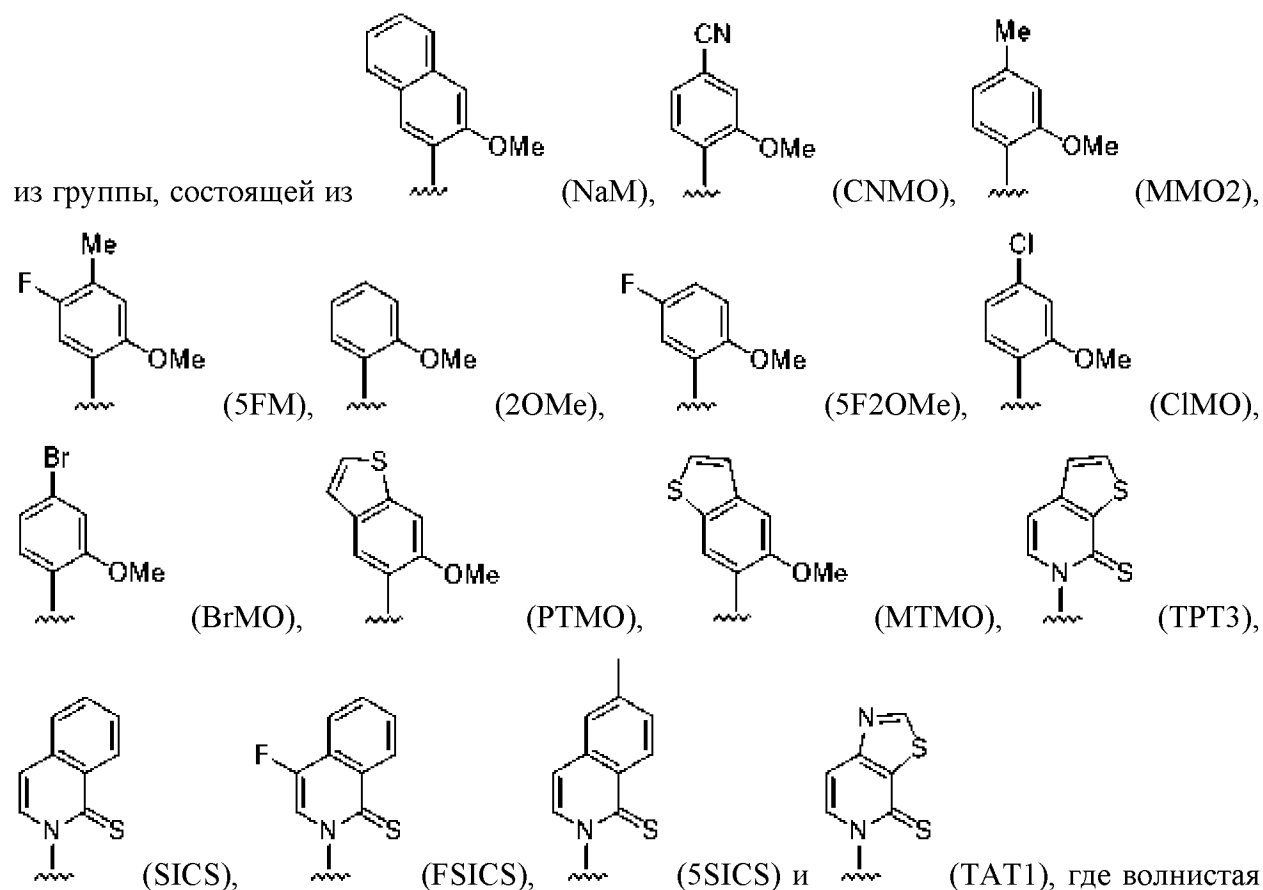


собой (NaM). В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет собой

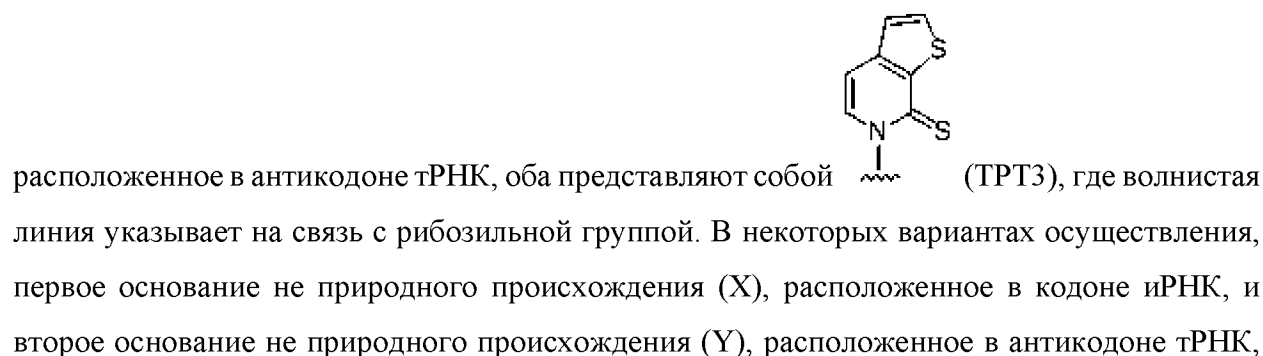
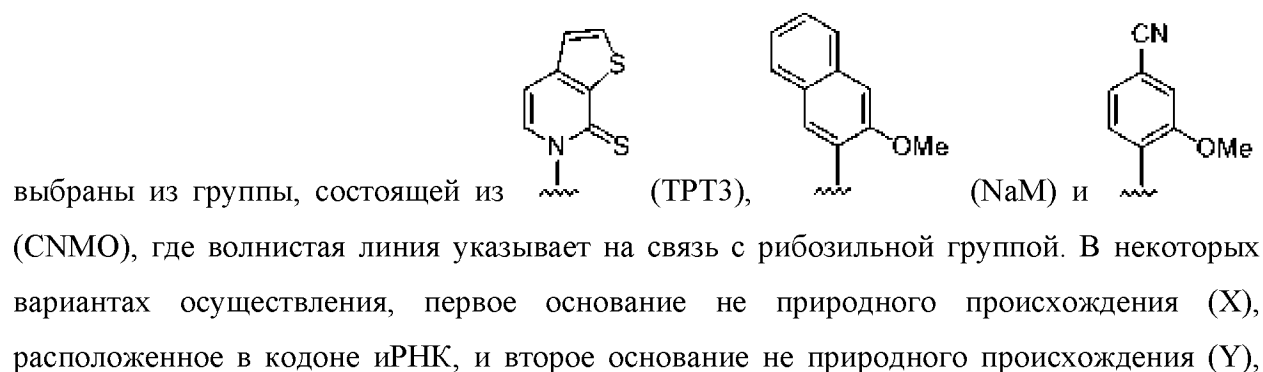


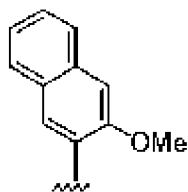
(CNMO). В некоторых вариантах осуществления, каждый из кодона и антикодона содержит три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в среднем положении (N-X-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в среднем положении (N-Y-N) антикодона.

В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК выбраны

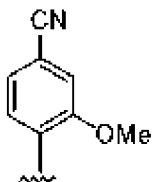


линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК,

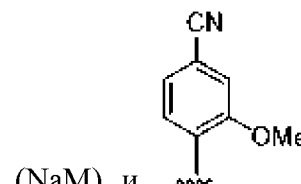
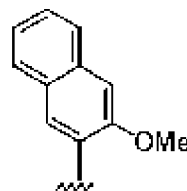




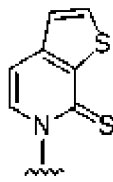
оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба представляют



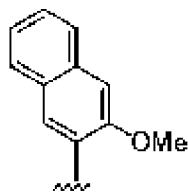
собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения



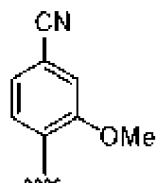
(X), расположенное в кодоне иРНК, выбрано из (NaM) и (CNMO), и второе основание не природного происхождения (Y), расположенное в



антикодоне тРНК, представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет

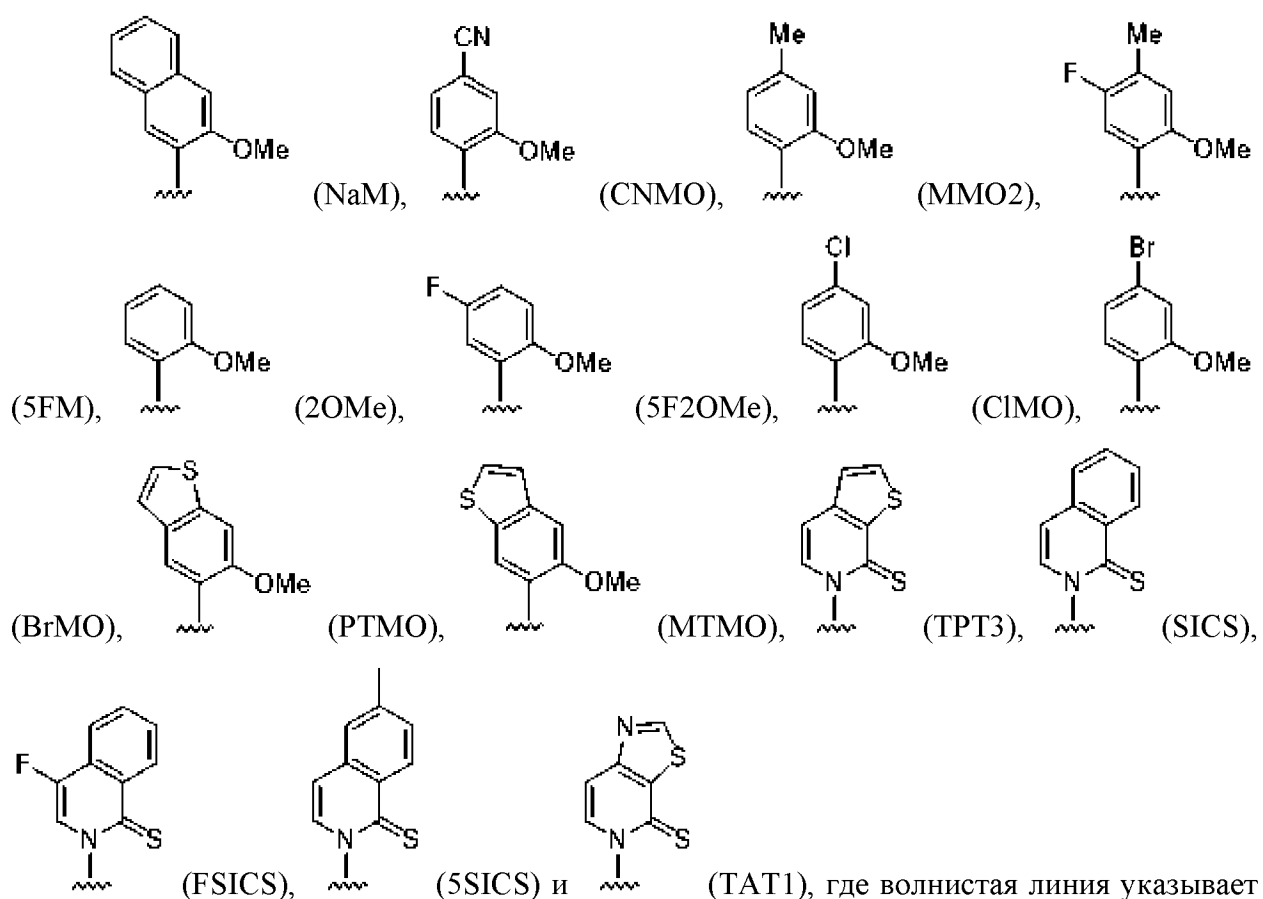


собой (NaM). В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет собой

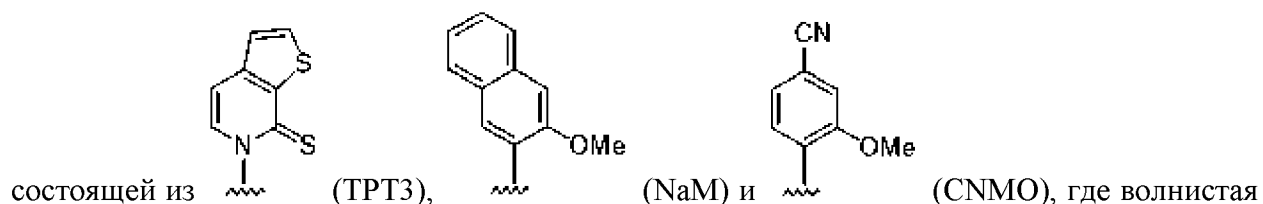


(CNMO). В некоторых вариантах реализации кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в последнем положении (N-N-X) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в первом положении (Y-N-N) антикодона. В некоторых

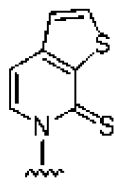
вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано из группы, состоящей из



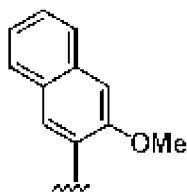
В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК выбрано из группы,



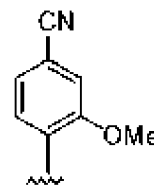
линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК,



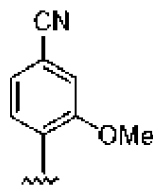
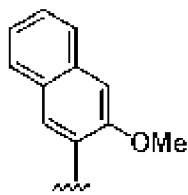
оба представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба представляют



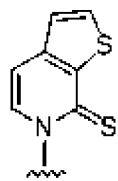
собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X) расположенное в кодоне иРНК, и второе основание не природного происхождения (Y),



расположенное в антикодоне тРНК, оба представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, первое основание не природного происхождения (X), расположенное в

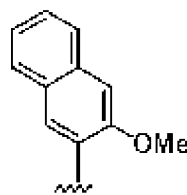


кодоне иРНК, выбрано из (NaM) и (CNMO), и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, представляет собой



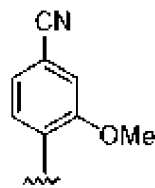
, где в каждом случае волнистая линия указывает на связь с рибозильной группой.

В некоторых вариантах осуществления, первое основание не природного происхождения



(X), расположенное в кодоне иРНК, представляет собой (NaM). В некоторых

вариантах осуществления, первое основание не природного происхождения (X),



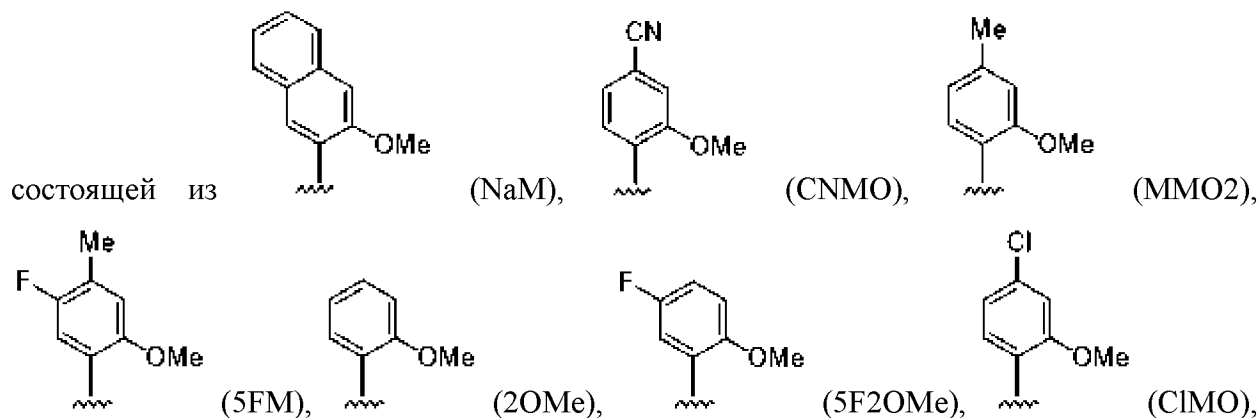
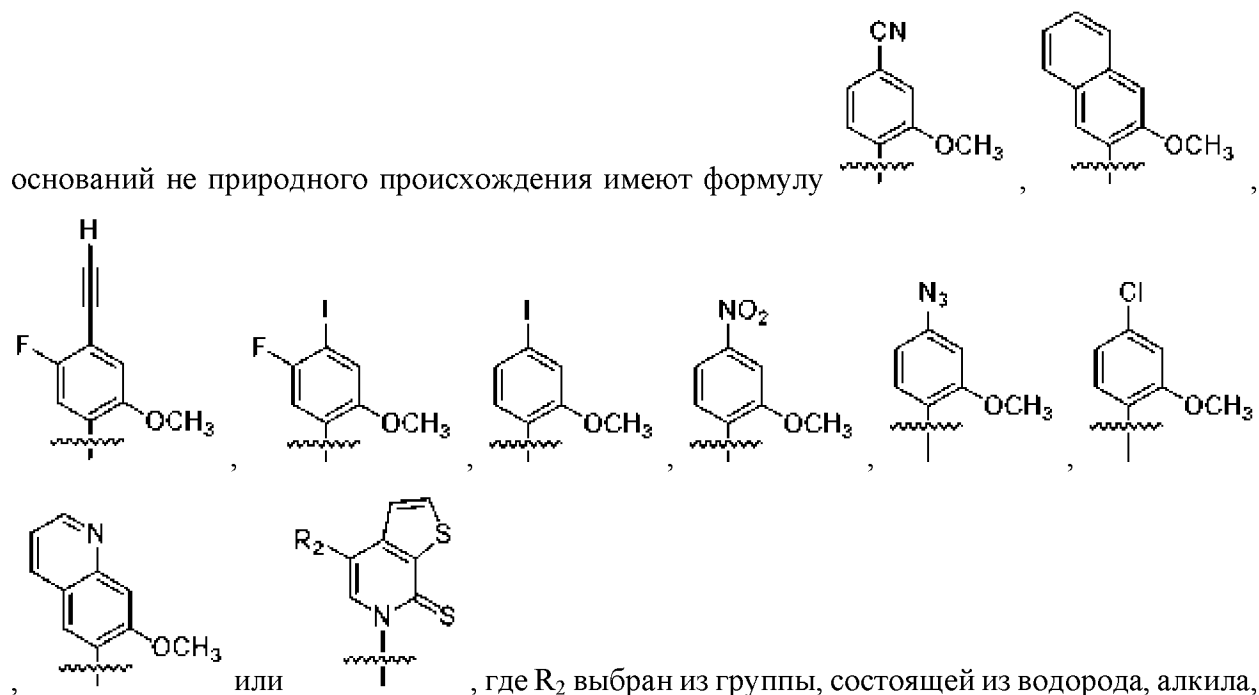
расположенное в кодоне иРНК, представляет собой (CNMO). В некоторых вариантах осуществления, кодон в иРНК выбран из AXC, GXC или GXU, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, кодон в иРНК представляет собой AXC, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, кодон в иРНК представляет собой GXC, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, кодон в иРНК представляет собой GXU, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, кодон в иРНК выбран из AXC, GXC или GXU, где антикодон в тРНК выбран из GYU, GYC и AYC, где X представляет собой первое основание не природного происхождения, и Y представляет собой второе основание не природного происхождения. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, кодон в иРНК A представляет собой AXC, и антикодон в тРНК представляет собой GYU. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, кодон в иРНК представляет собой GXC, и антикодон в тРНК представляет собой GYC. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, кодон в иРНК представляет собой GXU, и антикодон представляет собой AYC. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, кодон в иРНК представляет собой GXU, и антикодон представляет собой AYC. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, тРНК происходит из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, аминоксил тРНК синтетаза (также называемая в настоящем документе просто тРНК синтетазой) происходит из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых

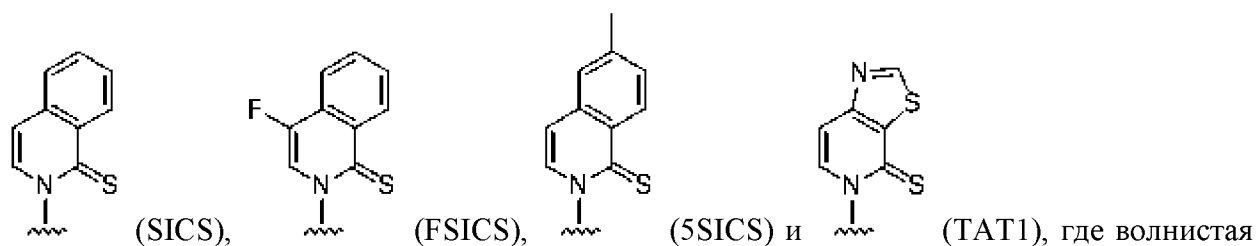
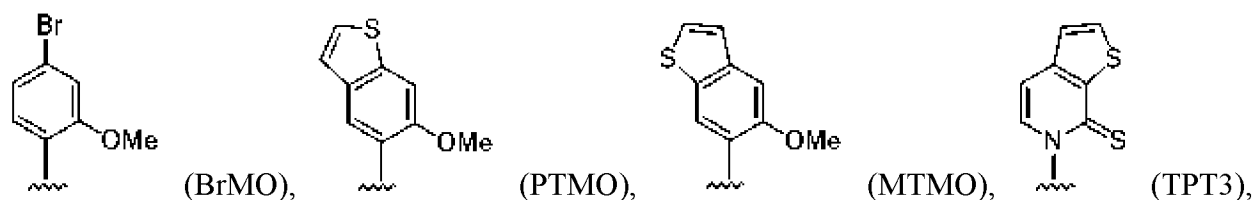
вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanococcus jannaschii*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina barkeri*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina mazei*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получена из *Methanococcus jannaschii*, и тРНК синтетазу получают из *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получают из *Methanosarcina barkeri*, и тРНК синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получают из *Methanosarcina mazei*, и тРНК синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получают из *Methanosarcina acetivorans*, и тРНК синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina mazei*. В некоторых вариантах осуществления, тРНК получают из *Methanosarcina mazei*, и тРНК синтетазу получают из *Methanosarcina barkeri*. В некоторых вариантах осуществления, клетка представляет собой клетку человека. В некоторых вариантах осуществления, клетка человека представляет собой клетку НЕК293Т. В некоторых вариантах осуществления, клетка представляет собой клетку хомяка. В некоторых вариантах осуществления, клетка хомяка представляет собой клетку яичника китайского хомяка (СНО). В некоторых вариантах осуществления, аминокислота не природного происхождения: (а) представляет собой аналог лизина; (б) содержит ароматическую боковую цепь; (с) содержит азидогруппу; (d) содержит алкиновую группу или (е) содержит альдегидную или кетонную группу. В некоторых вариантах осуществления, аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонила)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонотигозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-

нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина или N6-(((4-азидобензил)окси)карбонил)-L-лизина. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK). В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((2-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((3-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((4-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, иРНК и тРНК стабилизированы к деградации в эукариотической клетке. В некоторых вариантах осуществления, полипептид получают путем трансляция иРНК с использованием тРНК через рибосому, эндогенную для эукариотической клетки.

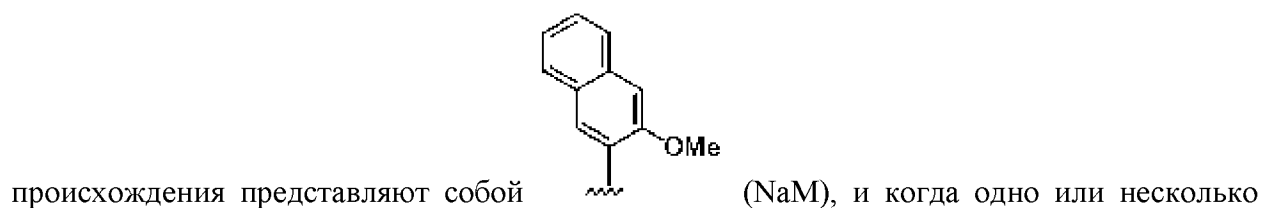
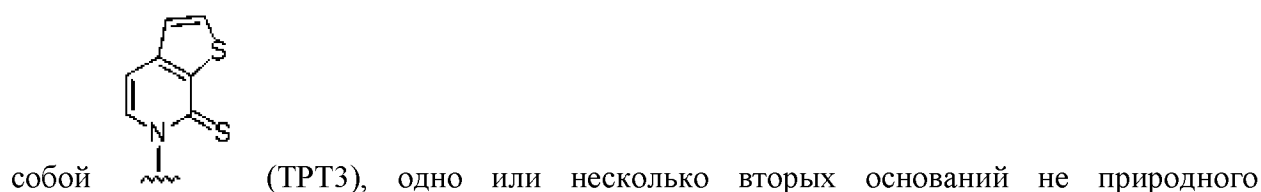
Аспекты, описанные в настоящем документе, представляют системы для экспрессии полипептида не природного происхождения, содержащие: (a) по меньшей мере, одну аминокислоту не природного происхождения; (b) иРНК, кодирующую полипептид не природного происхождения, где указанная иРНК содержит, по меньшей мере, один кодон, содержащий одно или несколько первых оснований не природного происхождения; (c) тРНК, содержащую, по меньшей мере, один антикодон, содержащий один или нескольких вторых оснований не природного происхождения, где одно или несколько первых оснований не природного происхождения и одно или несколько вторых оснований не природного происхождения образуют одну или несколько комплементарных пар оснований; и (d) эукариотическую рибосому, способную к трансляции иРНК в полипептид, содержащий аминокислоту не природного происхождения, с использованием тРНК и тРНК синтетазы. тРНК может быть заряжена аминокислотой не природного происхождения, и/или система может дополнительно содержать тРНК синтетазу и/или одну или несколько конструкций нуклеиновых кислот, содержащих последовательность нуклеиновой кислоты, кодирующая тРНК синтетазу, где тРНК синтетаза преимущественно аминоацилирует тРНК с, по меньшей мере, одной не встречающейся в природе аминокислотой. Система может быть создана *in vitro* (например, бесклеточная, такая как клеточный лизат или восстановленная система очищенных компонентов) или в эукариотической клетке. В некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три

непрерывных нуклеотидных основания (N-N-N); и где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N), по меньшей мере, в одном кодоне иРНК. В некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК. В некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X), по меньшей мере, в одном кодоне иРНК. В некоторых вариантах осуществления, одно или несколько

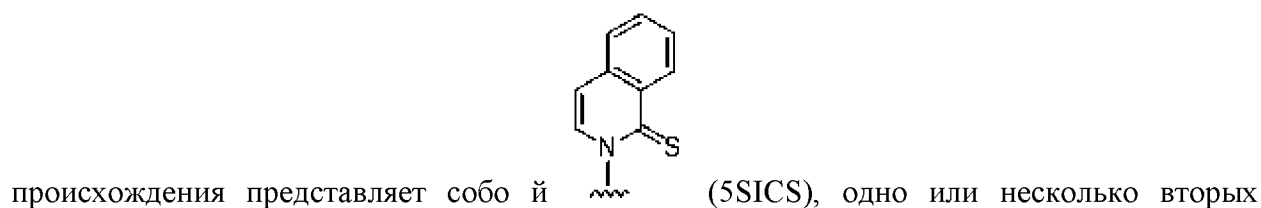




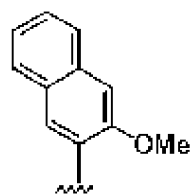
линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда одно или несколько первых оснований не природного происхождения представляют



вариантах осуществления, когда одно или несколько первых оснований не природного

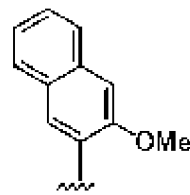


ненатуральных оснований представляет собой



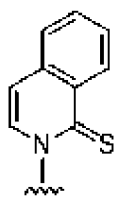
(NaM), и когда одно или

несколько первых ненатуральных оснований представляет собой



(NaM),

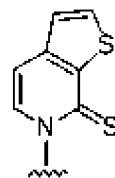
одно или несколько вторых оснований не природного происхождения представляет собой



(5SICS), где волнистая линия указывает связь с рибозильной группой. В

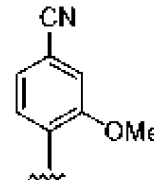
некоторых вариантах осуществления, когда одно или несколько первых оснований не

природного происхождения представляют собой



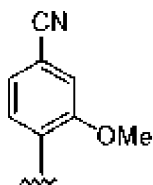
(TPT3), одно или несколько

вторых оснований не природного происхождения представляют собой



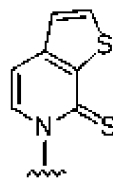
(CNMO), и когда одно или несколько первых оснований не природного происхождения

представляют собой



(CNMO), одно или несколько вторых оснований не

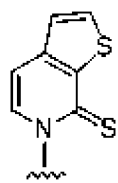
природного происхождения представляют собой



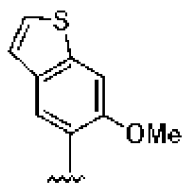
(TPT3), где волнистая линия

указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда

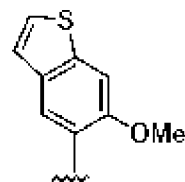
одно или несколько первых оснований не природного происхождения представляют собой



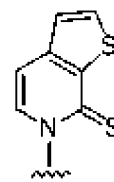
(TPT3), одно или несколько вторых оснований не природного происхождения



представляют собой (PTMO), и когда одно или несколько первых оснований

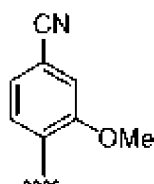


не природного происхождения представляют собой (PTMO), одно или

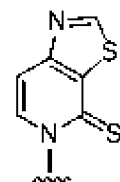


несколько вторых оснований не природного происхождения представляют собой

(TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда одно или несколько первых оснований не природного

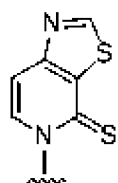


происхождения представляют собой (CNMO), одно или несколько вторых

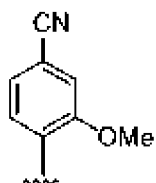


оснований не природного происхождения представляют собой (TAT1), и когда

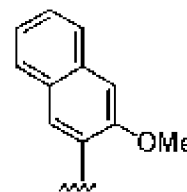
одно или несколько первых оснований не природного происхождения представляют собой



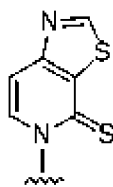
(TAT1), одно или несколько вторых оснований не природного происхождения



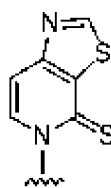
представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, когда одно или несколько



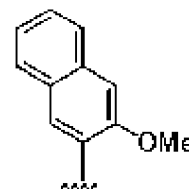
первых оснований не природного происхождения представляют собой (NaM), одно или несколько вторых оснований не природного происхождения представляют



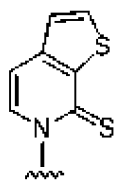
собой (TAT1), и когда одно или несколько первых оснований не природного



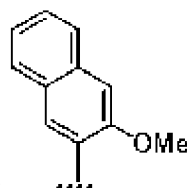
происхождения представляют собой (TAT1), одно или несколько вторых



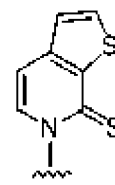
оснований не природного происхождения представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения



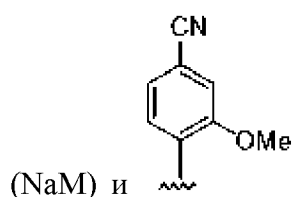
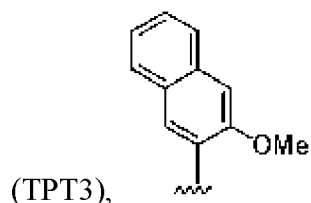
представляют собой (TPT3) и одно или несколько вторых оснований не



природного происхождения представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно

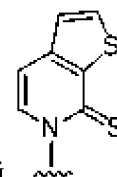


или несколько первых оснований не природного происхождения выбрано из

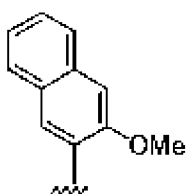


(CNMO), где волнистая линия указывает на

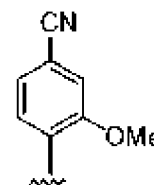
связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько



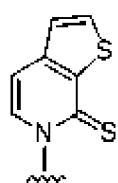
первых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения



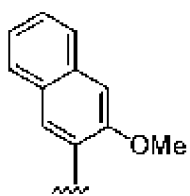
представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых



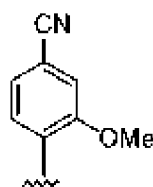
оснований не природного происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных оснований (N-N-N), где одно или несколько первых оснований не природного происхождения (**X**) расположены в первом положении (**X**-N-N) в кодоне иРНК, где одно или несколько первых оснований не природного происхождения выбраны из



(TPT3),

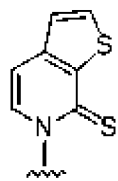


(NaM) и



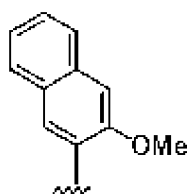
(CNMO), и где волнистая линия

указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения представляют собой



(TPT3), где волнистая линия указывает на связь с рибозильной группой

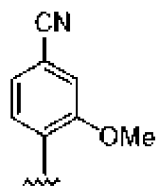
В некоторых вариантах осуществления, одно или несколько первых оснований не природного



происхождения представляют собой

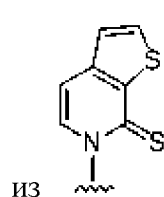
(NaM), где волнистая линия указывает

на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения представляют собой



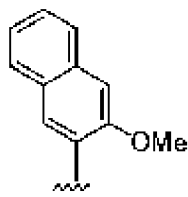
(CNMO), где волнистая линия указывает на связь с рибозильной группой. В

некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК, где одно или несколько первых оснований не природного происхождения выбраны

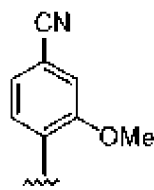


из

(TPT3),

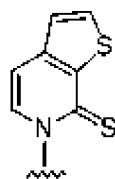


(NaM) и



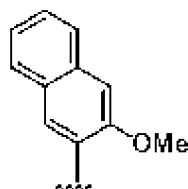
(CNMO), и где волнистая линия

указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения представляют собой



(TPT3), где волнистая линия указывает на связь с рибозильной группой. В

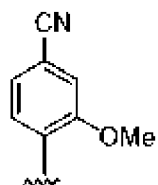
некоторых вариантах осуществления, одно или несколько первых оснований не природного



происхождения представляют собой (NaM), где волнистая линия указывает

на связь с рибозильной группой. В некоторых вариантах осуществления, одно или

несколько первых оснований не природного происхождения представляют собой

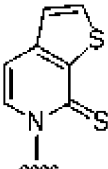
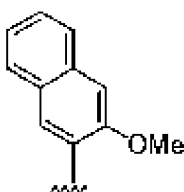
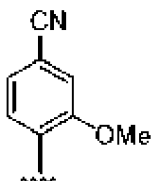


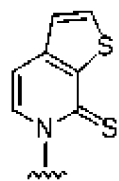
(CNMO), где волнистая линия указывает на связь с рибозильной группой. В

некоторых вариантах осуществления, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований

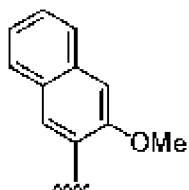
не природного происхождения (X) расположены в последнем положении (N-N-X) в кодоне

иРНК, где одно или несколько первых оснований не природного происхождения выбраны

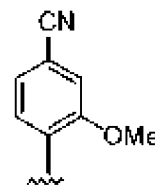
из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения представляют собой



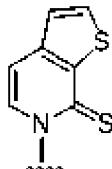
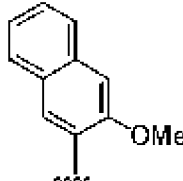
(TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного

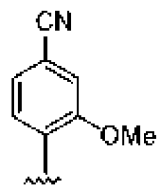


происхождения (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых

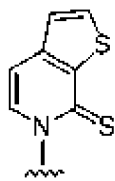


оснований не природного происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько вторых оснований не природного происхождения (X) расположены в первом положении (X-N-N) антикодона тРНК. В некоторых вариантах осуществления, одно или несколько вторых оснований не

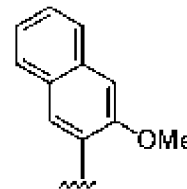
природного происхождения выбраны из  (TPT3),  (NaM) и



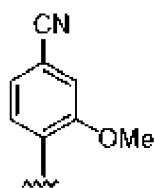
(CNMO), и где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько вторых оснований не природного



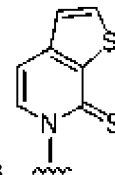
происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько



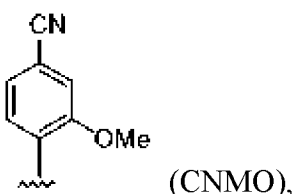
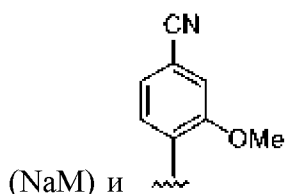
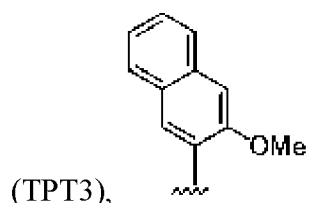
вторых оснований не природного происхождения представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько вторых оснований не природного



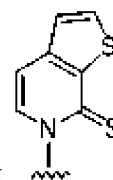
происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в антикодоне тРНК. В некоторых вариантах осуществления,



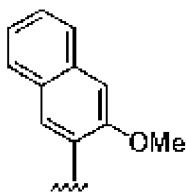
одно или несколько вторых оснований не природного происхождения выбраны из



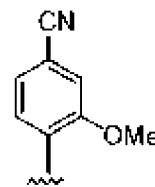
связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько



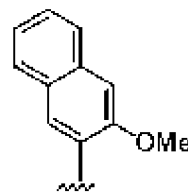
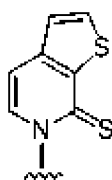
вторых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько вторых оснований не природного происхождения



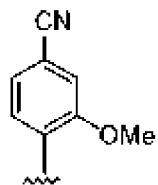
представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой В некоторых вариантах осуществления, одно или несколько вторых



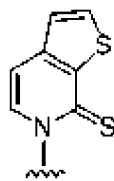
оснований не природного происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько вторых оснований не природного происхождения (X) расположены в последней позиции n(N-N-X) в антикодоне тРНК. В некоторых вариантах осуществления, одно или несколько вторых оснований не



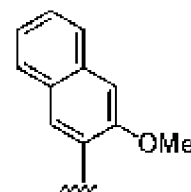
природного происхождения выбраны из (TPT3), (NaM) и



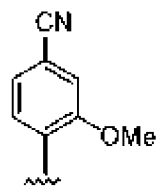
(CNMO), и где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько вторых оснований не природного



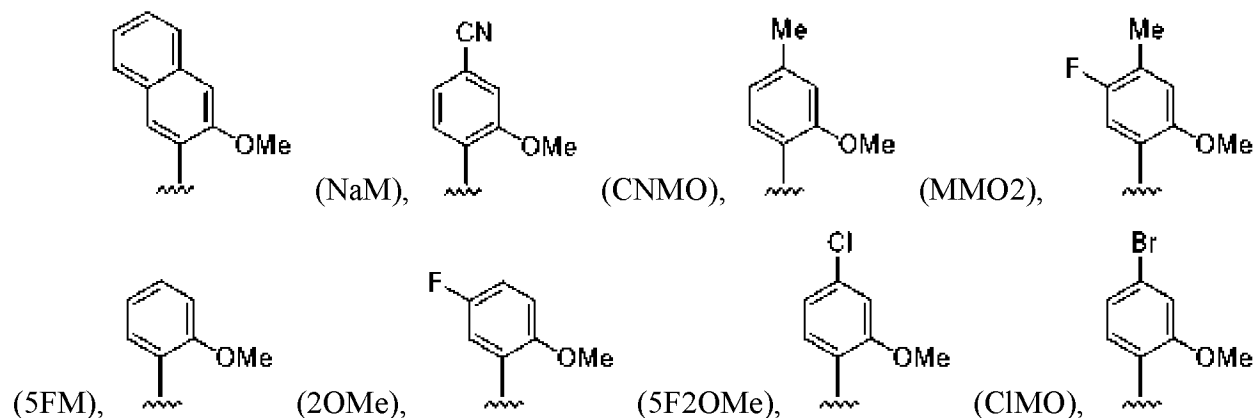
происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько

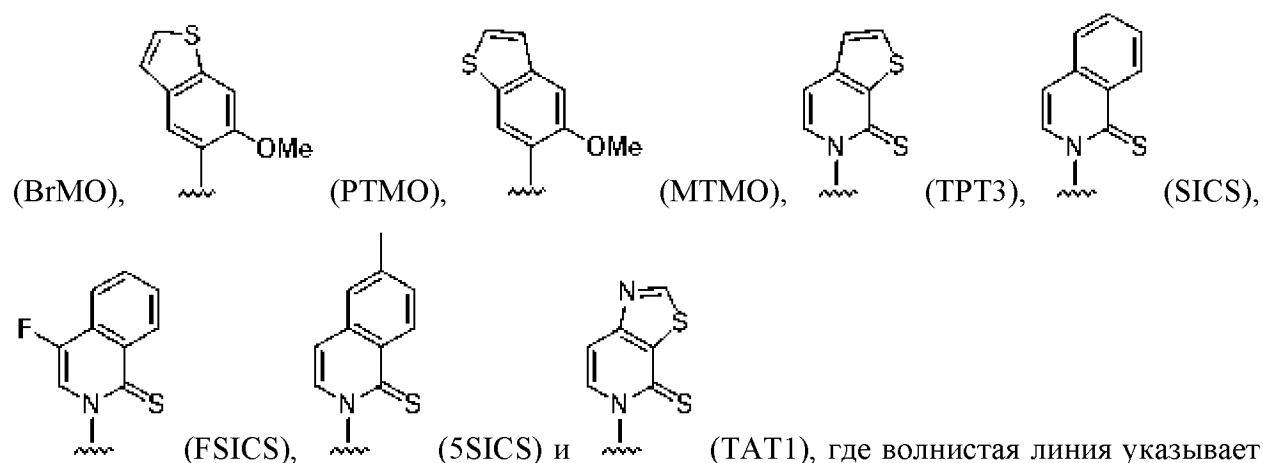


вторых оснований не природного происхождения представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько вторых оснований не природного

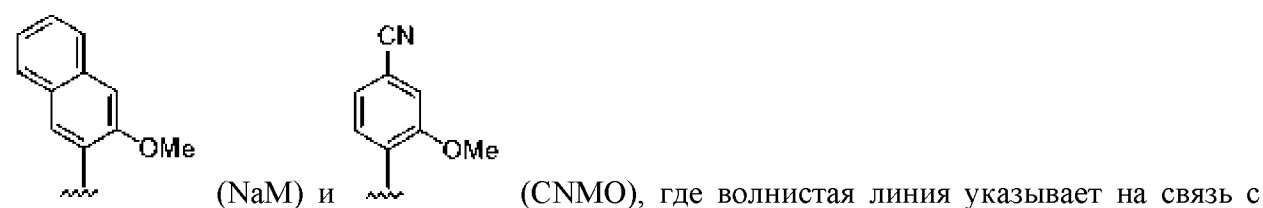
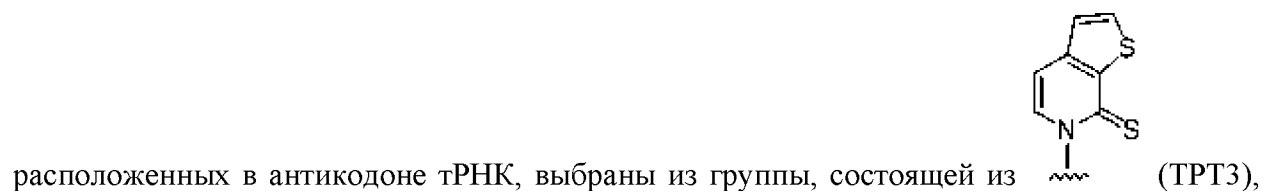


происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, по меньшей мере, один кодон и, по меньшей мере, один антикодон, каждый независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон содержит одно или несколько первых оснований не природного происхождения (X), расположенных в первом положении (X-N-N) кодона, и, по меньшей мере, одно антикодона в тРНК содержат одно или несколько вторых оснований не природного происхождения (Y), расположенных в последнем положении (N-N-Y) антикодона. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными. В некоторых вариантах осуществления, одно или несколько первых не природного происхождения основания (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

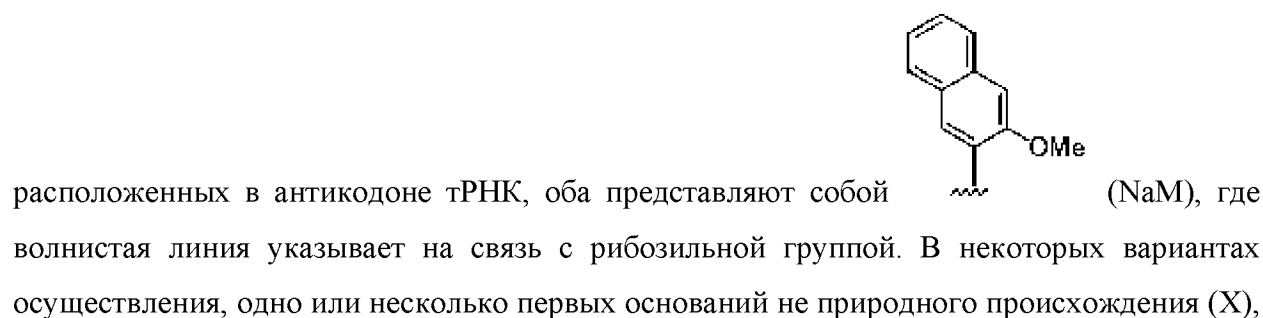
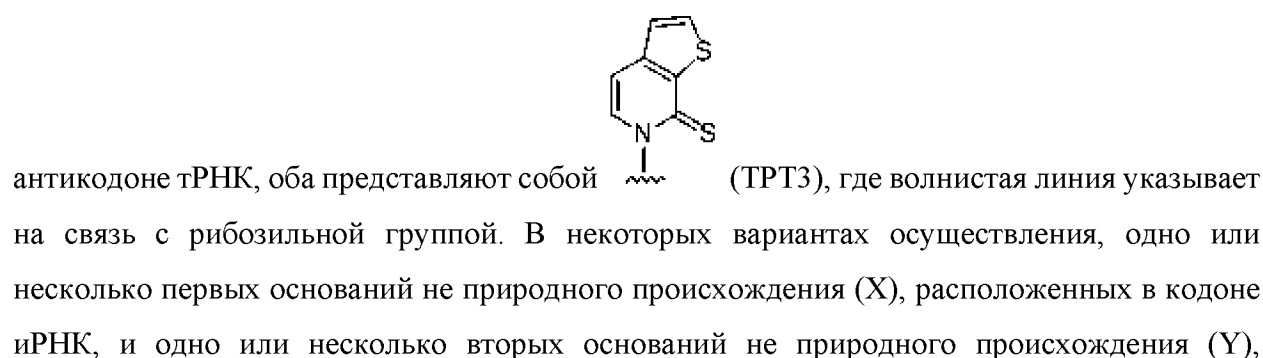




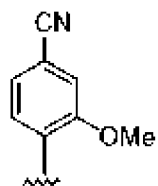
В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



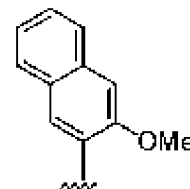
В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в



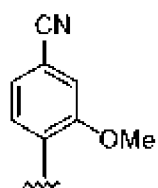
расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, оба представляют собой



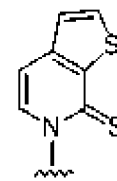
(CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного



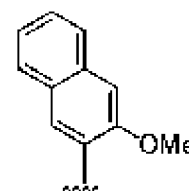
происхождения (X), расположенных в кодоне иРНК, выбрано из (NaM) и



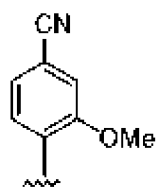
(CNMO), и одно или несколько вторых оснований не природного



происхождения (Y), расположенных в антикодоне тРНК, представляют собой (TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного

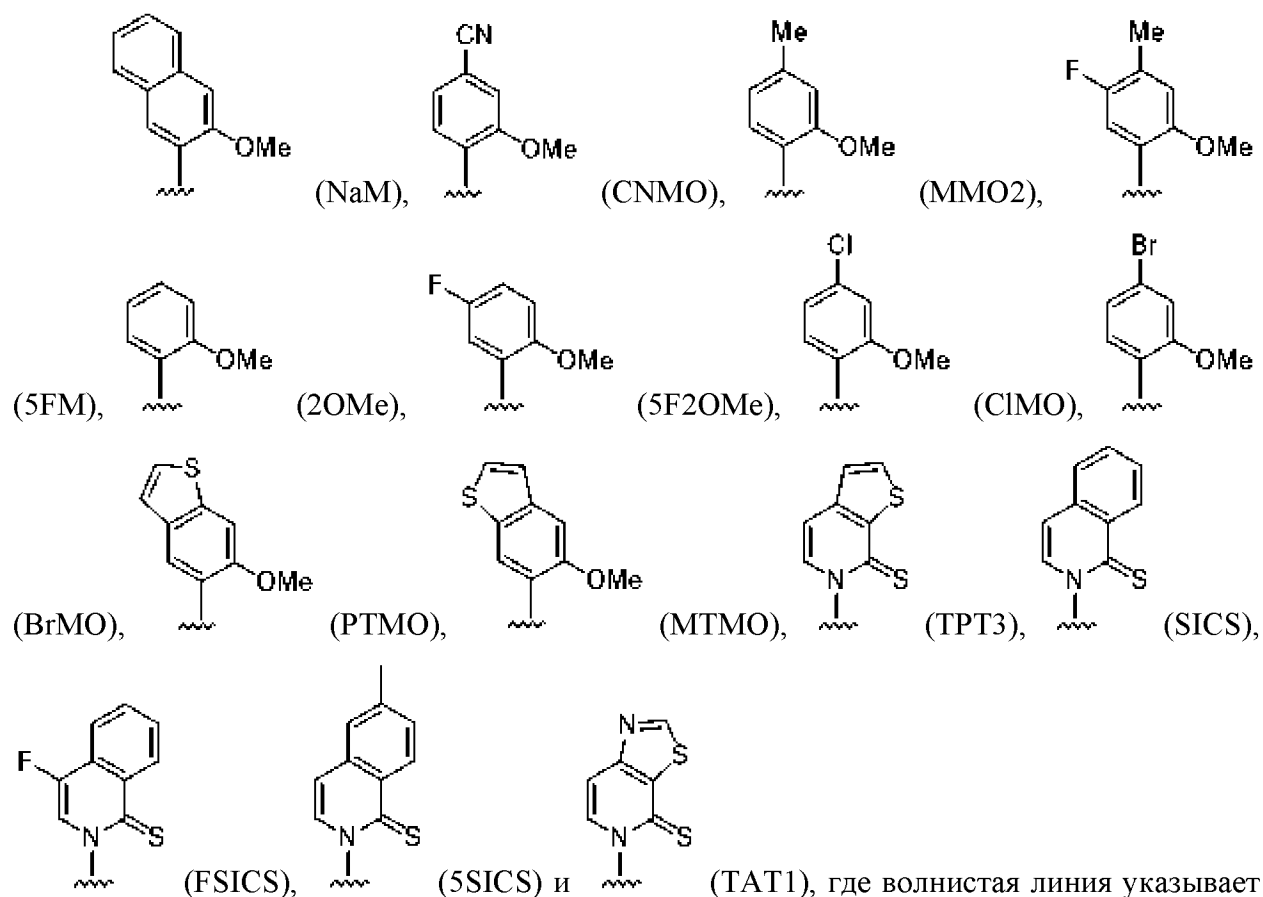


происхождения (X), расположенных в кодоне иРНК представляет собой (NaM). В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, представляют собой

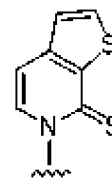


(CNMO). В некоторых вариантах осуществления, по меньшей мере, один кодон и, по меньшей мере, один антикодон каждый, независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и, где по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в среднем положении (N-X-N), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного

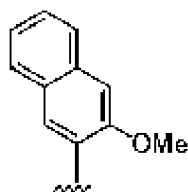
происхождения (Y), расположенных в среднем положении (N-Y-N) антикодона. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из



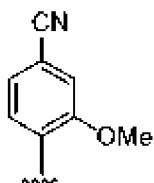
В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



расположенных в антикодоне тРНК, выбраны из группы, состоящей из (TPT3),

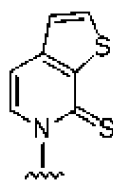


(NaM) и

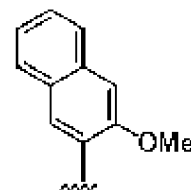


(CNMO), где волнистая линия указывает на связь с

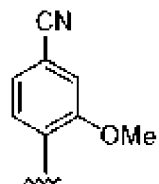
рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в



антикодоне тРНК, оба представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

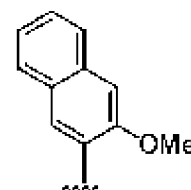


расположенных в антикодоне тРНК, оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенные в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, оба представляют собой

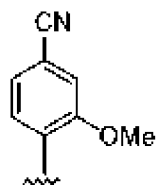


(CNMO), где волнистая линия указывает на связь с рибозильной группой. В

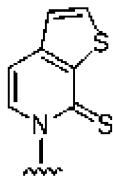
некоторых вариантах осуществления, одно или несколько первых оснований не природного

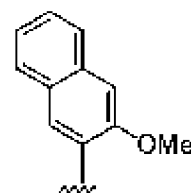


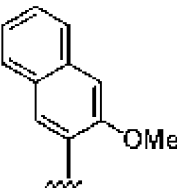
происхождения (X), расположенных в кодоне иРНК, выбраны из (NaM) и

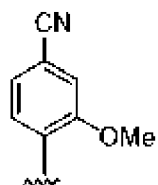


(CNMO), и одно или несколько вторых оснований не природного

происхождения (Y), расположенных в антикодоне тРНК, представляет собой , где волнистая линия в каждом случае указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного



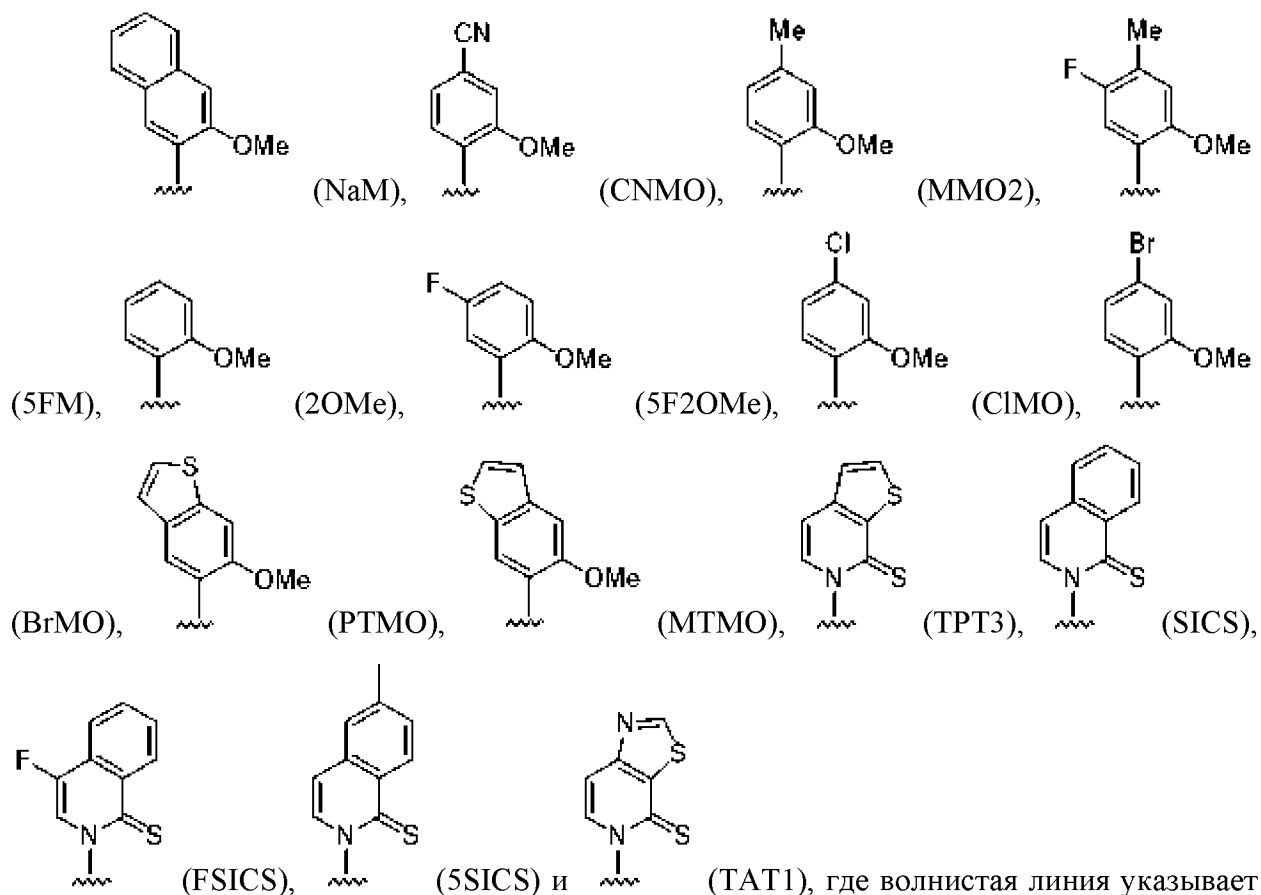
происхождения (X), расположенных в кодоне иРНК представляют собой  (NaM). В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, представляют собой



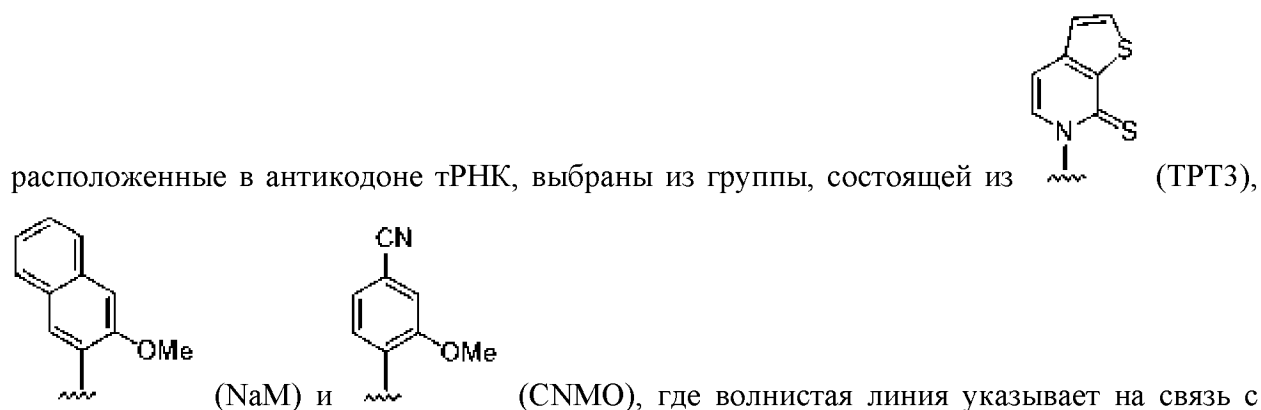
(CNMO). В некоторых вариантах осуществления, по меньшей мере, один кодон

и по меньшей мере, один антикодон каждый, независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в последнем положении (N-N-X), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в первом положении (Y-N-N) антикодона. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными. В некоторых

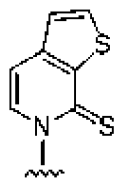
вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из



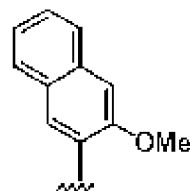
В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



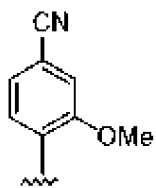
В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в



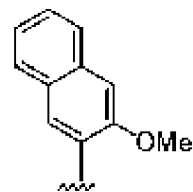
антикодоне тРНК, оба представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



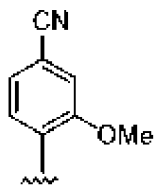
расположенных в антикодоне тРНК, оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенные в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, оба представляют собой



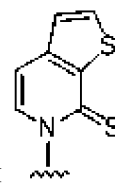
(CNMO), где волнистая линия указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного



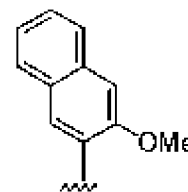
происхождения (X), расположенных в кодоне иРНК, выбрано из (NaM) и



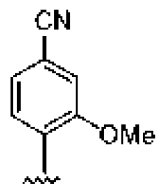
(CNMO), и одно или несколько вторых оснований не природного



происхождения (Y), расположенных в антикодоне тРНК, представляют собой , где волнистая линия в каждом случае указывает на связь с рибозильной группой. В некоторых вариантах осуществления, одно или несколько первых оснований не природного



происхождения (X), расположенных в кодоне иРНК, представляют собой (NaM). В некоторых вариантах осуществления, одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, представляют собой



(CNMO). В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК выбран из AXC, GXC или GXU, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой AXC, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой GXC, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой GXU, где X представляет собой основание не природного происхождения. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК выбран из AXC, GXC или GXU, где по меньшей мере, один антикодон в тРНК выбран из GYU, GYC и AYC, где X представляет собой одно или несколько первых оснований не природного происхождения и Y представляет собой одно или несколько вторых оснований не природного происхождения. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой AXC, а по меньшей мере, один антикодон в тРНК представляет собой GYU. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой GXC, а по меньшей мере, один антикодон в тРНК представляет собой GYC. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, по меньшей мере, один кодон в иРНК представляет собой GXU, и, по меньшей мере, один антикодон представляет собой

АУС. В некоторых вариантах осуществления, X и Y являются одинаковыми или разными. В некоторых вариантах осуществления, X и Y являются одинаковыми. В некоторых вариантах осуществления, X и Y являются разными. В некоторых вариантах осуществления, тРНК получают из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanococcus jannaschii*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina barkeri*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina mazei*. В некоторых вариантах осуществления, тРНК и тРНК синтетаза получены из *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получены из *Methanococcus jannaschii* и тРНК синтетаза получена из *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получены из *Methanosarcina mazei*, и тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina acetivorans*. В некоторых вариантах осуществления, тРНК получена из *Methanosarcina acetivorans*, и тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina mazei*. В некоторых вариантах осуществления, тРНК получена из *Methanosarcina mazei*, и тРНК синтетаза получена из *Methanosarcina barkeri*. В некоторых вариантах осуществления, клетка представляет собой клетку человека. В некоторых вариантах осуществления, клетка человека представляет собой клетку HEK293T. В некоторых вариантах осуществления, клетка представляет собой клетку хомяка. В некоторых вариантах осуществления, клетка хомяка представляет собой клетку яичника китайского хомяка (СНО). В некоторых вариантах осуществления, аминокислота не природного происхождения: (a) представляет собой аналог лизина; (b) содержит ароматическую боковую цепь; (c) содержит азидогруппу; (d) содержит алкиновую группу или (e) содержит альдегидную или кетонную группу. В некоторых вариантах осуществления, аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононаноевой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина,

изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтирозина, три-О-ацетил-GlcNAc-серины, L-фосфосерина, фосфоносерины, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK). В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((2-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((3-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-(((4-азидобензил)окси)карбонил)-L-лизин. В некоторых вариантах осуществления, иРНК и тРНК стабилизированы к деградации в эукариотической клетке. В некоторых вариантах осуществления, полипептид получают путем трансляции иРНК с использованием тРНК через рибосому, эндогенную для эукариотической клетки.

В одном варианте осуществления, эукариотическая клетка содержит иРНК, кодирующую усиленный зеленый флуоресцентный белок (EGFP) с кодоном не природного происхождения в положении 151 (EGFP151(NXN); где N относится к одному из природных нуклеотидных оснований, и X относится к NaM), tRNAPyl *Methanosarcina mazei*, перекодированный родственным антикодом не природного происхождения (tRNAPyl(NYN), где Y относится к TPT3), и химерную пирролизил-тРНК синтетазу *Methanosarcina barkeri* (ChPylRS), которая может заряжать tRNAPyl не природного происхождения с помощью N6-(2-азидозтокси)-карбонил-L-лизин (AzK).

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Различные аспекты изобретения подробно изложены в прилагаемой формуле изобретения. Лучшее понимание особенностей и преимуществ настоящего изобретения будет получено путем обращения к следующему подробному описанию, в котором представлены иллюстративные варианты осуществления, в которых используются принципы изобретения, и прилагаемым чертежам, в которых:

На **фигурах 1А-1С** показаны UBP и рабочий процесс с использованием UBP по настоящему варианту осуществления. На **фигуре 1А** изображены типовые пары оснований не природного происхождения (UBP) dNaM и dTPT3. На **фигуре 1В** показан рабочий

процесс с использованием UBP для сайт-специфического включения не канонических аминокислот (ncAA) в белок с использованием пары оснований не природного происхождения XY. Включение трех ncAA в белок показано только в качестве примера; может быть включено любое количество ncAA. На **фигуре 1С** изображены типовые UBP.

На **фигуре 2** показаны аналоги dXTP. Рибоза и фосфаты не показаны для ясности.

На **фигурах 3А-3В** показаны типовые основания не природного происхождения.

На **фигурах 4А-4Г** показаны типовые аминокислоты не природного происхождения. Эти аминокислоты не природного происхождения (UAA) генетически закодированы в белках (**фиг. 4D** - UAA # 1-42; **фиг. 4E** - UAA # 43-89; **фиг. 4F** - UAA # 90-128; **фиг. 4G** - UAA # 129-167). **Фиг. 4D-4G** взяты из таблицы 1 из Dumas *et al.*, *Chemical Science* 2015, 6, 50-69.

На **фигурах 5А-5В** показана трансляция кодонов не природного происхождения в клетках НЕК293Т. На **фигуре 5А** показан средний сигнал флуоресценции EGFP клеток НЕК293Т, трансфицированных кодонами не природного происхождения с или без родственных тРНК, измеренный с помощью проточной цитометрии. На **фигуре 5В** показан анализ белкового сдвига для клеток НЕК293Т, трансфицированных кодоном не природного происхождения GXC с использованием клеточного лизата.

На **фигурах 6А-6В** показана трансляция кодонов не природного происхождения в клетках СНО. На **фигуре 6А** показан средний сигнал флуоресценции EGFP клеток СНО, трансфицированных кодонами не природного происхождения (представленными ДНК, кодирующей кодон не природного происхождения) с или без родственных тРНК (и само-спаривающейся тРНК для кодона AGX), измеренный с помощью проточной цитометрии. На **фигуре 6В** показан анализ сдвига белка для клеток СНО, трансфицированных кодонами не природного происхождения AXC, GXC, GXT, GYC и AGX (представленными ДНК, кодирующей кодон не природного происхождения) с использованием очищенного EGFP.

На **Фигурах 7А-7В** показана трансляция кодонов не природного происхождения в контексте CYBA UTR в клетках СНО. **Фигура 7А**: Средний сигнал флуоресценции EGFP клеток СНО, трансфицированных кодонами не природного происхождения в контексте CYBA UTR, с или без родственных тРНК (и само-спаривающейся тРНК для кодона AGX), измеренный с помощью проточной цитометрии. * $P < 0,05$, ** $P < 0,005$, *** $P < 0,0005$, **** $P < 0,00005$ (двухсторонний парный t-критерий). **Фиг. 7В**: Анализ белкового сдвига для клеток СНО, трансфицированных кодонами не природного происхождения GXC и GYC в контексте CYBA UTR с использованием очищенного EGFP.

На **фигурах 7С-7D** показано соотношение экспрессии белка между иРНК с CYBA UTR и иРНК с CS2 UTR. На **фигуре 7С** показано соотношение уровней экспрессии EGFP для различных кодонов не природного происхождения в пределах CYBA UTR и CS2 UTR. Уровень экспрессии измеряют с помощью проточной цитометрии. На **фигуре 7D** показано, используя РВ-кПЦР, количество иРНК, измеренное через 4 ч после транскрипции и через 8 ч после транскрипции. Соотношение иРНК, оставшейся через 8 ч, по сравнению с иРНК, оставшейся через 4 ч, сравнивается для различных конструкций иРНК. Обратите внимание,

что кодоны не природного происхождения на фигурах 7А и 7В представлены кодирующей последовательностью ДНК, кодирующей иРНК.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Определенная терминология

Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют то же значение, которое обычно понимается специалистом в области техники, к которой относится заявленный объект изобретения. Следует понимать, что предшествующее общее описание и следующее подробное описание являются только примерными и объясняющими, и не ограничивают какой-либо заявленный объект. В этой заявке использование единственного числа включает множественное число, если специально не указано иное. Следует отметить, что, как используется в описании и прилагаемой формуле изобретения, формы единственного числа «a», «an» и «the» включают ссылки во множественном числе, если контекст явно не указывает на иное. В этой заявке использование «или» означает «и/или», если не указано иное. Кроме того, использование термина «включая» а также другие формы, такие как «включать», «включает» и «включается», не является ограничивающим.

Используемые в настоящем документе диапазоны и количества могут быть выражены как «приблизительно» для конкретного значения или диапазона. Примерно также включает точное количество. Следовательно, «примерно 5 мкл» означает «примерно 5 мкл», а также «5 мкл». В общем, термин «примерно» включает количество, которое, как ожидается, будет находиться в пределах ошибки эксперимента.

Такие фразы, как «в условиях, подходящих для получения» или «в условиях, достаточных для получения» и подобные, в контексте способов синтеза, используемые в настоящем документе, относятся к условиям реакции, таким как время, температура, растворитель, концентрации реагентов и подобные, которые экспериментатор может варьировать в пределах своей квалификации, которые дают полезное количество или выход продукта реакции. Необязательно, чтобы желаемый продукт реакции был единственным продуктом реакции, или чтобы исходные материалы были полностью израсходованы, при условии, что желаемый продукт реакции может быть выделен или иным образом использован в дальнейшем.

Под «химически возможным» подразумевается расположение связей или соединение, в котором не нарушаются общепонятные правила органической структуры; например, структуру в рамках определения пункта формулы изобретения, которая в определенных ситуациях может содержать пентавалентный атом углерода, которого не существует в природе, следует понимать как не подпадающую под формулу изобретения. Структуры, описанные в настоящем документе, во всех их вариантах осуществления, предназначены для включения только «химически возможных» структур и любых перечисленных структур, которые не являются химически возможными, например, в структуре, показанной с переменными атомами или группами не предназначены для описания или заявки в настоящем документе.

«Аналог» химической структуры, как этот термин используется в настоящем документе, относится к химической структуре, которая сохраняет существенное сходство с исходной структурой, хотя ее нельзя легко получить синтетическим путем из исходной структуры. В некоторых вариантах осуществления, аналог нуклеотида представляет собой нуклеотид не природного происхождения. В некоторых вариантах осуществления, аналог нуклеозида представляет собой нуклеозид не природного происхождения. Родственная химическая структура, которая легко получается синтетическим путем из исходной химической структуры, называется «производной».

Соответственно, термин полинуклеотид, используемый в настоящем документе, относится к ДНК, РНК, ДНК- или РНК-подобным полимерам, таким как пептидные нуклеиновые кислоты (ПНК), закрытые нуклеиновые кислоты (ЗНК), фосфоротиоаты, основания не природного происхождения и подобные, которые хорошо известны в данной области техники. Полинуклеотиды могут быть синтезированы в автоматических синтезаторах, например, с использованием химии фосфорамидитов или других химических подходов, адаптированных для использования в синтезаторах.

ДНК включает, но не ограничен ими, кДНК и геномную ДНК. ДНК может быть присоединена ковалентными или не ковалентными способами к другой биомолекуле, включая, но не ограничиваясь ими, РНК и пептид. РНК включает кодирующую РНК, например информационную РНК. (иРНК). В некоторых вариантах осуществления, РНК представляет собой рРНК, РНКи, мякРНК, микроРНК, киРНК, мяРНК, эксРНК, риРНК, длинную нкРНК или любую их комбинацию или гибрид. В некоторых случаях РНК представляет собой компонент рибозима. ДНК и РНК может быть в любой форме, включая, помимо прочего, линейную, кольцевую, сверхспиральную, одноцепочечную и двухцепочечную.

Пептидная нуклеиновая кислота (ПНК) представляет собой синтетический аналог ДНК/РНК, в котором пептидоподобный остов заменяет сахаро-фосфатный остов ДНК или РНК. Олигомеры ПНК демонстрируют более высокую силу связывания и большую специфичность в связывании с комплементарными ДНК, где несоответствие оснований ПНК/ДНК является более дестабилизирующим, чем аналогичное несоответствие в дуплексе ДНК/ДНК. Эта сила связывания и специфичность также применимы к дуплексам ПНК/РНК. ПНК не легко распознаются ни нуклеазами, ни протеазами, что делает их устойчивыми к расщеплению ферментами. ПНК также стабильны в широком диапазоне pH. См. также Nielsen PE, Egholm M, Berg RH, Buchardt O (December 1991). "Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide", *Science* **254** (5037): 1497-500. doi:10.1126/science.1962210. PMID 1962210; и Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Nordén B, and Nielsen PE (1993), "PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen Bonding Rules". *Nature* **365** (6446): 566-8. doi:10.1038/365566a0. PMID 7692304

Запертая нуклеиновая кислота (ЗНК) представляет собой модифицированный нуклеотид РНК, в котором рибозная группа нуклеотида ЗНК модифицирована

дополнительным мостиком, соединяющим 2'-кислород и 4'-углерод. Мост «запирает» рибозу в конформации 3'-*эндо* (North), которая часто встречается в дуплексах А-формы. Нуклеотиды ЗНК можно смешивать с остатками ДНК или РНК в олигонуклеотиде, когда это необходимо. Такие олигомеры можно синтезировать химическим путем, и они коммерчески доступны. Заблокированная конформация рибозы улучшает укладку оснований и предварительную организацию остова. См., например, Kaur, H; Arora, A; Wengel, J; Maiti, S (2006), "Thermodynamic, Counterion, and Hydration Effects for the Incorporation of Locked Nucleic Acid Nucleotides into DNA Duplexes", *Biochemistry* **45** (23): 7347-55. doi:10.1021/bi060307w. PMID 16752924; Owczarzy R.; You Y., Groth C.L., Tataurov A.V. (2011), "Stability and mismatch discrimination of locked nucleic acid-DNA duplexes.", *Biochem.* **50** (43): 9352-9367. doi:10.1021/bi200904e. PMC 3201676. PMID 21928795; Alexei A. Koshkin; Sanjay K. Singh, Poul Nielsen, Vivek K. Rajwanshi, Ravindra Kumar, Michael Meldgaard, Carl Erik Olsen, Jesper Wengel (1998), "LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition", *Tetrahedron* **54** (14): 3607-30. doi:10.1016/S0040-4020(98)00094-5; и Satoshi Obika; Daishu Nanbu, Yoshiyuki Hari, Ken-ichiro Morio, Yasuko In, Toshimasa Ishida, Takeshi Imanishi (1997), "Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3'-*endo* sugar puckering", *Tetrahedron Lett.* **38** (50): 8735-8. doi:10.1016/S0040-4039(97)10322-7.

Молекулярный маяк или зонд молекулярного маяка представляет собой зонд гибридизации олигонуклеотидов, который может обнаруживать присутствие определенной последовательности нуклеиновой кислоты в гомогенном растворе. Молекулярные маяки представляют собой молекулы в форме шпильки с внутренне погашенным флуорофором, флуоресценция которых восстанавливается, когда они связываются с последовательностью нуклеиновой кислоты-мишени. См., например, Tyagi S, Kramer FR (1996), "Molecular beacons: probes that fluoresce upon hybridization", *Nat Biotechnol.* **14** (3): 303-8. PMID 9630890; Tapp I, Malmberg L, Rennel E, Wik M, Syvänen AC (2000 Apr), "Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan assay and Molecular Beacon probes", *Biotechniques* **28** (4): 732-8. PMID 10769752; и Akimitsu Okamoto (2011), "ECHO probes: a concept of fluorescence control for practical nucleic acid sensing", *Chem. Soc. Rev.* **40**: 5815-5828.

В некоторых вариантах осуществления, нуклеотидное основание обычно представляет собой часть гетероциклического основания нуклеозида. Нуклеотидные основания могут встречаться в природе, могут быть модифицированы, могут не иметь сходства с природными основаниями и могут быть синтезированы, например, с помощью органического синтеза. В определенных вариантах осуществления, нуклеотидное основание содержит любой атом или группу атомов, способных взаимодействовать с основанием другой нуклеиновой кислоты с или без использования водородных связей. В некоторых вариантах осуществления, нуклеотидное основание не природного происхождения не происходит из природного нуклеотидного основания. Необходимо

отметить, что нуклеотидные основания не природного происхождения не обязательно обладают основными свойствами, однако для простоты называются нуклеотидными основаниями. В некоторых вариантах осуществления, когда речь идет о нуклеотидных основаниях, «(d)» указывает, что нуклеотидные основания могут быть присоединены к дезоксирибозе или рибозе.

В некоторых вариантах осуществления, нуклеозид представляет собой соединение, содержащее группу нуклеотидного основания и группу сахара. Нуклеозиды включают, но не ограничены ими, встречающиеся в природе нуклеозиды (такие, как обнаружены в ДНК и РНК), нуклеозиды с удаленным азотистым основанием, модифицированные нуклеозиды и нуклеозиды, имеющие миметические основания и/или сахарные группы. Нуклеозиды включают нуклеозиды, содержащие любое разнообразие заместителей. Нуклеозид может быть гликозидным соединением, образованным посредством гликозидной связи между нуклеиновым основанием и восстанавливающей группой сахара.

Заголовки разделов, используемые в настоящем документе, предназначены только для организационных целей и не должны толковаться как ограничивающие описываемый объект.

Способы, Системы и Композиции, Содержащие Пары Оснований Не Природного Происхождения в Эукариотических Клетках

В некоторых вариантах осуществления, в настоящем документе описаны *in vivo* способы и композиции для получения нуклеиновой кислоты с расширенным генетическим алфавитом (**ФИГ. 1А-3В**) в эукариотической клетке. В некоторых случаях, нуклеиновая кислота кодирует белок не природного происхождения, где белок не природного происхождения содержит, по меньшей мере, одну аминокислоту не природного происхождения. В некоторых случаях, *in vivo* способ или композиция, описанные в настоящем документе, используют или содержат полусинтетический организм. В некоторых случаях, способ включает, по меньшей мере, одну пару оснований не природного происхождения (UBP) в одном или двух нуклеозидах. В типовом технологическом процессе, представленном на **ФИГ.1В**, ДНК **101**, кодирующая белок **102**, и тРНК **103**, каждая из которых содержит комплементарные нуклеотидные основания не природного происхождения (X, Y), транскрибируется **104** с образованием тРНК **106** и иРНК **107**. После зарядки тРНК аминокислотой не природного происхождения **105**, иРНК **107** транслируется **108** с образованием белка **110**, содержащего одну или несколько аминокислот не природного происхождения **109**. Способы и композиции, описанные в настоящем документе, в некоторых случаях позволяют встраивать аминокислоты не природного происхождения в сайты с высокой точностью и выходом. В настоящем документе также описаны полусинтетические организмы, включающие расширенный генетический алфавит, способы использования организмов для продуцирования белковых продуктов, включая такие, которые содержат, по меньшей мере, один аминокислотный остаток не природного происхождения.

Выбор нуклеотидных оснований не природного происхождения позволяет

оптимизировать одну или несколько стадий в описанных в настоящем документе способах. Например, нуклеотидные основания выбирают для высокоэффективной репликации, транскрипции и/или трансляции. В некоторых случаях, более одной пары нуклеотидных оснований не природного происхождения используют для способов, описанных в настоящем документе. Например, первый набор нуклеотидных оснований, содержащих группу дезоксирибо, используют для репликации ДНК (такой как первое нуклеотидное основание и второе нуклеотидное основание, сконфигурированные для образования первой пары оснований), и второй набор нуклеотидных оснований (такой как третье нуклеотидное основание и четвертое нуклеотидное основание, где третье и четвертое нуклеотидные основания присоединены к рибозе, сконфигурированы с образованием второй пары оснований) используют для транскрипции/трансляции. Комплементарное спаривание между нуклеотидным основанием первого набора и нуклеотидным основанием второго набора в некоторых случаях позволяет транскрипцию генов для создания тРНК или белков из матрицы ДНК, содержащей нуклеотидные основания из первого набора. Комплементарное спаривание между нуклеотидными основаниями второго набора (вторая пара оснований) в некоторых случаях позволяет трансляцию путем сопоставления тРНК, содержащих нуклеиновые кислоты не природного происхождения и иРНК. В некоторых случаях, нуклеотидные основания в первом наборе присоединены к группе дезоксирибозы. В некоторых случаях, нуклеотидные основания в первом наборе присоединены к группе рибозы. В некоторых случаях, нуклеотидные основания обоих наборов уникальны. В некоторых случаях, по меньшей мере, одно нуклеотидное основание одинаково в обоих наборах. В некоторых случаях, первое нуклеотидное основание и третье нуклеотидное основание являются одинаковыми. В некоторых вариантах осуществления, первая пара оснований и вторая пара оснований не являются одинаковыми. В некоторых случаях, первая пара оснований, вторая пара оснований и третья пара оснований не являются одинаковыми.

Эукариотические Сконструированные Организмы

В некоторых вариантах осуществления, способы и плазмиды, описанные в настоящем документе, дополнительно используются для создания эукариотических сконструированных организмов, например сконструированного организма, который включает и реплицирует нуклеотид не природного происхождения или пару кислотных оснований не природного происхождения (UBP), и может также использовать нуклеиновую кислоту, содержащую нуклеотид не природного происхождения, для транскрипции иРНК и тРНК, которые используются для трансляции белков, содержащих аминокислотный остаток не природного происхождения. В некоторых случаях, организм является полусинтетическим организмом (SSO). В некоторых случаях, SSO не является прокариотическим. В некоторых случаях SSO, является SSO млекопитающего. В некоторых случаях, SSO млекопитающего является SSO человека. В некоторых случаях, SSO млекопитающего является SSO хомяка. В некоторых случаях, SSO человека получают из клетки HEK293T. В некоторых случаях, SSO человека получают из клеток яичника

китайского хомяка (СНО).

В некоторых случаях, используемая клетка генетически трансформируется кассетой экспрессии, кодирующей гетерологичный белок, например, тРНК-синтетазу. В некоторых вариантах осуществления, тРНК-синтетаза предпочтительно аминоацилирует тРНК, содержащую антикодон, содержащий основание не природного происхождения, с аминокислотой не природного происхождения. В некоторых вариантах осуществления, клетка содержит тРНК-синтетазу, которая преимущественно аминоацилирует тРНК, содержащую антикодон, содержащий основание не природного происхождения, с аминокислотой не природного происхождения.

Клетка может быть эукариотической клеткой, и пара взаимно спаривающихся нуклеотидов не природного происхождения может быть ТРТЗ и NaМ или КНМО.

В настоящем документе описаны композиции и способы, включающие использование двух или нескольких нуклеотидов не природного происхождения, образующих пары оснований. Такие нуклеотиды, образующие пары оснований, в некоторых случаях попадают в клетку с помощью стандартных способов трансформации нуклеиновых кислот, известных в данной области техники (например, электропорация, химическая трансформация или другие способы, в которых нуклеиновые кислоты, содержащие нуклеотиды не природного происхождения, могут быть введены в клетку). В некоторых случаях, используются три или несколько нуклеотидов не природного происхождения, спаривающих основания. В некоторых случаях, нуклеотид не природного происхождения, спаривающий основания, входит в клетку как часть полинуклеотида, например иРНК и/или тРНК. Один или несколько нуклеотидов не природного происхождения, спаривающих основания, которые входят в клетку как часть полинуклеотида (РНК), сами по себе не нуждаются в репликации *in vivo*.

В некоторых случаях генетически сконструированные клетки получают путем введения в клетки нуклеиновых кислот, *например*, гетерологичных нуклеиновых кислот. Любая клетка, описанная в настоящем документе, может быть клеткой-хозяином и может содержать вектор экспрессии. В некоторых вариантах осуществления, клетка представляет собой клетку млекопитающего. В некоторых вариантах осуществления, клетка млекопитающего представляет собой клетку человека (*например*, клетку НЕК293Т). В некоторых вариантах осуществления, клетка млекопитающего представляет собой клетку хомяка (*например*, клетку СНО). В некоторых вариантах осуществления, клетка содержит один или несколько гетерологичных полинуклеотидов. Реагенты на основе нуклеиновой кислоты могут быть введены в микроорганизмы с использованием различных методов. Неограничивающие примеры способов, используемых для введения гетерологичных нуклеиновых кислот в различные организмы, включают трансформацию, трансфекцию, трансдукцию, электропорацию, ультразвук-медиированную трансформацию, конъюгацию, бомбардировку частицами и подобные. В некоторых случаях добавление молекул-носителей (*например*, бис-бензоимидазолильных соединений, *например*, см. патент США № 5,595,899) может увеличить поглощение ДНК в клетках, обычно тех, которые трудно

трансформировать обычными способами. Обычные способы трансформации легко доступны специалистам и могут быть найдены в Maniatis, T., E. F. Fritsch and J. Sambrook (1982) *Molecular Cloning: a Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

В некоторых случаях генетическую трансформацию получают путем прямого переноса кассеты экспрессии, в частности, в плазмиды, вирусные векторы, вирусные нуклеиновые кислоты, фаговые нуклеиновые кислоты, фаги, космиды и искусственные хромосомы, или путем переноса генетического материала в клетки или носители, такие как катионные липосомы. Такие способы доступны в данной области техники и легко адаптируются для использования в способе, описанном в настоящем документе. Векторы переноса могут представлять собой любую нуклеотидную конструкцию, используемую для доставки генов в клетки (*например*, плазмиду), или как часть общей стратегии доставки генов, *например*, в составе рекомбинантного ретровируса или аденовируса (Ram et al. *Cancer Res.* 53:83-88, (1993)). Подходящие средства для трансфекции, включая вирусные векторы, химические трансфектанты или физико-механические способы, такие как электропорация и прямая диффузия ДНК, описаны, например, в Wolff, J. A., et al., *Science*, 247, 1465-1468, (1990); and Wolff, J. A. *Nature*, 352, 815-818, (1991).

Молекулы Нуклеиновой Кислоты

В некоторых вариантах осуществления, нуклеиновая кислота (*например*, также называемая в настоящем документе представляющей интерес молекулой нуклеиновой кислоты) происходит из любого источника или композиции, такой как РНК, кРНК (короткая ингибиторная РНК), РНКи, тРНК, иРНК или рРНК (рибосомная РНК), например, и находится в любой форме (*например*, линейной, кольцевой, сверхспиральной, одноцепочечной, двухцепочечной и подобной). В некоторых вариантах осуществления, нуклеиновые кислоты содержат нуклеотиды, нуклеозиды или полинуклеотиды. В некоторых случаях, нуклеиновые кислоты также содержат нуклеиновые кислоты не природного происхождения, такие как аналоги РНК (*например*, содержащие аналоги оснований, аналоги сахаров и/или не нативный остов и подобные). Понятно, что термин «нуклеиновая кислота» не относится к конкретной длине полинуклеотидной цепи или не подразумевает ее, поэтому в определение также включены полинуклеотиды и олигонуклеотиды. Примеры природных нуклеотидов включают, без ограничения, АТФ, УТФ, СТФ, ГТФ, АДП, УДП, СДП, ГДП, АМФ, УМФ, СМФ, ГМФ, дАТФ, дТТФ, дСТФ, дГТФ, дАДП, дТДП, дСДП, дГДП, дАМФ, дТМФ, дСМФ и дГМФ. Примеры природных рибонуклеотидов включают дАТФ, дТТФ, дСТФ, дГТФ, дАДП, дТДП, дСДП, дГДП, дАМФ, дТМФ, дСМФ и дГМФ. Для природной РНК урациловым основанием является уридин. Иногда нуклеиновая кислота представляет собой вектор, плазмиду, фагмид, автономно реплицирующуюся последовательность (ARS), центромер, искусственную хромосому, искусственную хромосому дрожжей (*например*, YAC) или другую нуклеиновую кислоту, способную реплицироваться или быть реплицированной в клетке-хозяине. В некоторых случаях, нуклеиновая кислота не природного происхождения представляет собой аналог

нуклеиновой кислоты. В дополнительных случаях, нуклеиновая кислота не природного происхождения происходит из внеклеточного источника. В других случаях, нуклеиновая кислота не природного происхождения доступна для внутриклеточного пространства организма, представленного в настоящем документе, *например*, генетически модифицированного организма. В некоторых вариантах осуществления, нуклеотид не природного происхождения не является природным нуклеотидом. В некоторых вариантах осуществления, нуклеотид, который не содержит природное основание, содержит нуклеотидное основание не природного происхождения.

Нуклеиновые Кислоты Не Природного Происхождения

Нуклеотидный аналог или нуклеотид не природного происхождения включает нуклеотид, который содержит некоторый тип модификации основания, сахарной или фосфатной группы. В некоторых вариантах осуществления, модификация содержит химическую модификацию. В некоторых случаях, модификации происходят в 3'ОН или 5'ОН группе в основной цепи, в сахарном компоненте или в нуклеотидном основании. Модификации, в некоторых случаях, необязательно включают линкерные молекулы не природного происхождения и/или межцепочечные или внутрицепочечные поперечные связи. В одном аспекте, модифицированная нуклеиновая кислота содержит модификацию одной или нескольких групп 3'ОН или 5'ОН, остова, сахарного компонента или нуклеотидного основания, и/или добавление линкерных молекул не природного происхождения. В одном аспекте, модифицированный остов включает остов, отличный от фосфодиэфирного остова. В одном аспекте, модифицированный сахар содержит сахар, отличный от дезоксирибозы (в модифицированной ДНК) или отличный от рибозы (модифицированная РНК). В одном аспекте, модифицированное основание содержит аденин, гуанин, цитозин или тимин (в модифицированной ДНК) или основание, отличное от аденина, гуанина, цитозина или урацила (в модифицированной РНК).

В некоторых вариантах осуществления, нуклеиновая кислота содержит, по меньшей мере, одно модифицированное основание. В некоторых случаях, нуклеиновая кислота содержит 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 или более модифицированных оснований. В некоторых случаях, модификации основной группы включают природные и синтетические модификации А, С, G и T/U, а также различные пуриновые или пиримидиновые основания. В некоторых вариантах осуществления, модификация представляет собой модифицированную форму аденина, гуанина, цитозина или тимина (в модифицированной ДНК) или модифицированную форму аденина, гуанина, цитозина или урацила (в модифицированной РНК).

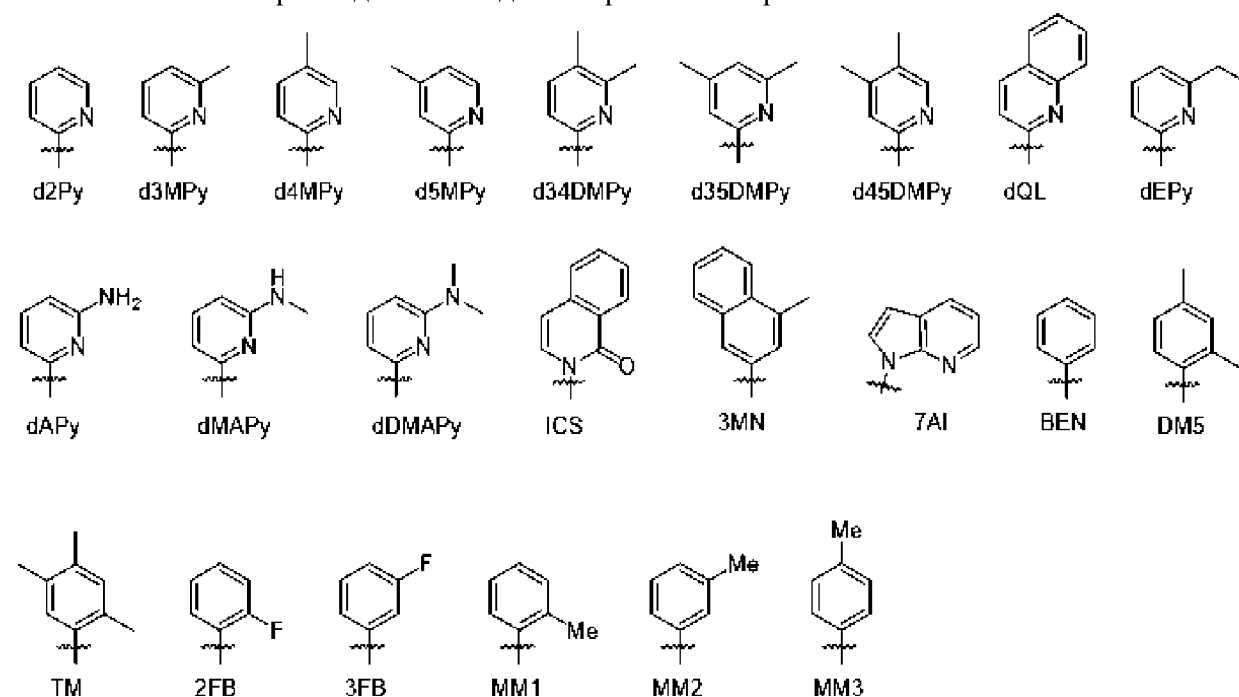
Модифицированное основание не встречающейся в природе нуклеиновой кислоты включает, но не ограничено ими, урацил-5-ил, гипоксантин-9-ил (I), 2-аминоаденин-9-ил, 5-метилцитозин (5-me-C), 5-гидроксиметилцитозин, ксантин, гипоксантин, 2-аминоаденин, 6-метил и другие алкильные производные аденина и гуанина, 2-пропил и другие алкильные производные аденина и гуанина, 2-тиоурацил, 2-тиотимин и 2-тиоцитозин, 5-галоурацил и -цитозин, 5-пропинилурацил и -цитозин, 6-азоурацил, -цитозин и -тимин, 5-урацил

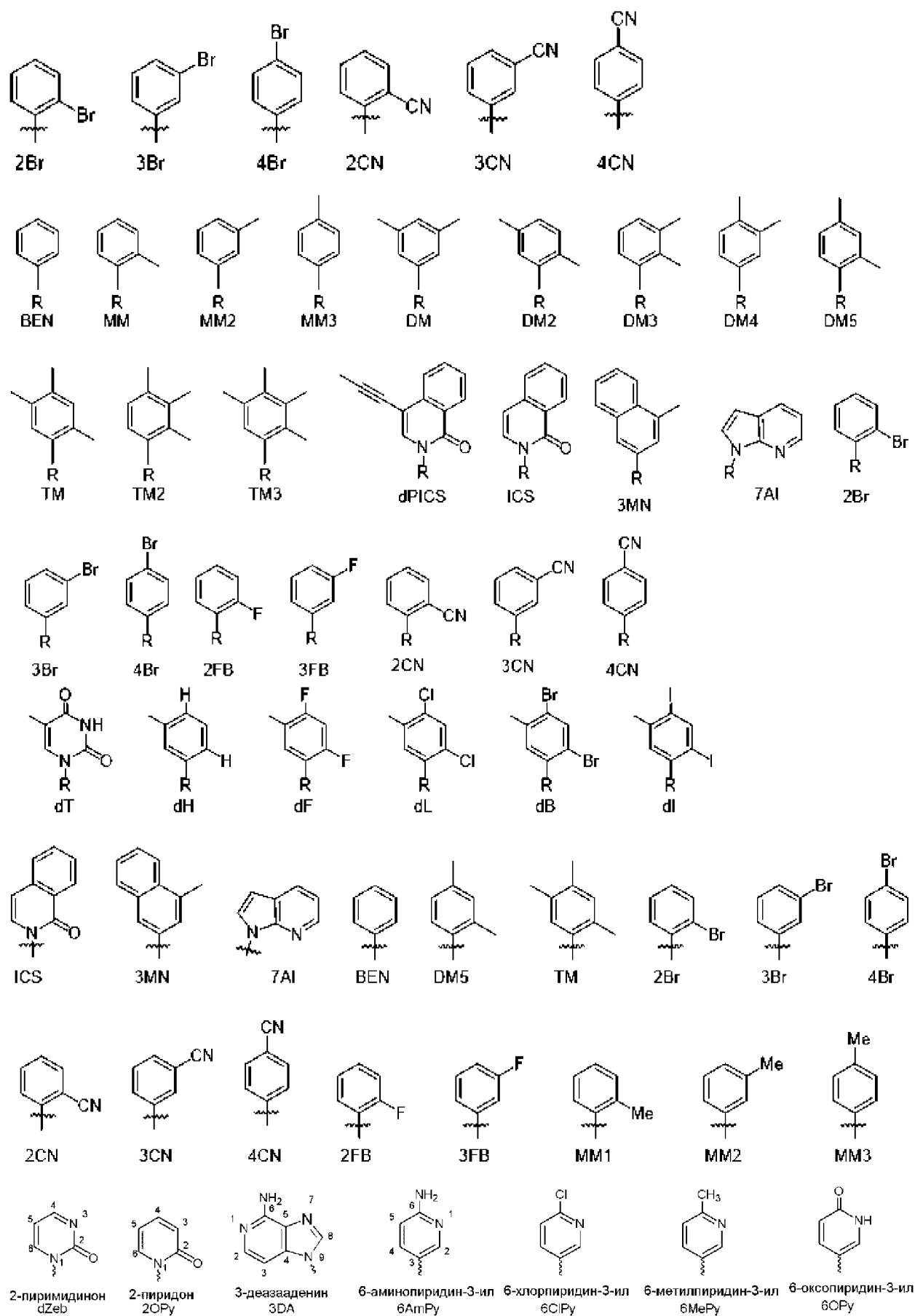
(псевдоурацил), 4-тиоурацил, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил-, 8- гидроксил- и другие 8-замещенные аденины и гуанины, 5-галоген-, особенно 5-бром-, 5-трифторметил- и другие 5-замещенные урацилы и цитозины, 7-метилгуанин и 7-метиладенин, 8-азагуанин и 8-азааденин, 7-деазагуанин и 7-деазааденин и 3-деазагуанин и 3-деазааденин. Некоторые нуклеиновые кислоты не природного происхождения, такие как 5-замещенные пиримидины, 6-азапиримидины и N-2 замещенные пурины, N-6 замещенные пурины, O-6 замещенные пурины, 2-аминопропиладенин, 5-пропинилурацил, 5-пропинилцитозин, 5-метилцитозин, которые повышают стабильность образования дуплексов, универсальные нуклеиновые кислоты, гидрофобные нуклеиновые кислоты, смешанные нуклеиновые кислоты, нуклеиновые кислоты увеличенного размера, фторированные нуклеиновые кислоты, 5-замещенные пиримидины, 6-азапиримидины и N-2, N-6 и O-6 замещенные пурины, включая 2-аминопропиладенин, 5-пропинилурацил и 5-пропинилцитозин, 5-метилцитозин (5-me-C), 5-гидроксиметилцитозин, ксантин, гипоксантин, 2-аминоаденин, 6-метил, другие алкильные производные аденина и гуанина, 2-пропил и другие алкильные производные аденина и гуанина, 2-тиоурацил, 2-тиотимин и 2-тиоцитозин, 5-галоурацил, 5-галоцитозин, 5-пропинил ($-C\equiv C-CH_3$)урацил, 5-пропинилцитозин, другие алкильные производные пиримидиновых нуклеиновых кислот, 6-азоурацил, 6-азоцитозин, 6-азотимин, 5 -урацил (псевдоурацил), 4-тиоурацил, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил-, 8-гидроксил- и другие 8-замещенные аденины и гуанины, 5-галоген-, особенно 5-бром-, 5-трифторметил-, другие 5-замещенные урацилы и цитозины, 7-метилгуанин, 7-метиладенин, 2-F-аденин, 2-аминоаденин, 8-азагуанин, 8-азааденин, 7-деазагуанин, 7-деазааденин, 3-деазагуанин, 3-деазааденин, трициклические пиримидины, феноксазин цитидин ([5,4-b][1,4]бензоксазин-2(3H)-он), фенотиазин цитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-он), G-фиксирующие основания, феноксазин цитидин (*например*, 9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-он), карбазол цитидин (2H-пиримидо[4,5-b]индол-2-он), пиридоиндолцитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-он), в которых пуриновое или пиримидиновое основание заменено другими гетероциклами, 7-деазааденин, 7-деазагуанозин, 2-аминопиридин, 2-пиридон, азацитозин, 5-бромцитозин, бромурацил, 5-хлорцитозин, хлорированный цитозин, циклоцитозин, цитозин арабинозид, 5-фторцитозин, фторпиримидин, фторурацил, 5,6-дигидроцитозин, 5-йодоцитозин, гидроксимочевина, йодоурацил, 5-нитроцитозин, 5-бромурацил, 5-хлорурацил, 5-фторурацил и 5-йодоурацил, 2-аминоаденин, 6-тиогуанин, 2-тиотимин, 4-тиотимин, 5-пропинилурацил, 4-тиоурацил, N4-этилцитозин, 7-деазагуанин, 7-деаза-8-азагуанин, 5-гидроксицитозин, 2'-дезоксинуридин, 2-амино-2'-дезоксиаденозин и те, которые описаны в патентах США №№ 3,687,808; 4,845,205; 4,910,300; 4,948,882; 5,093,232; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121; 5,596,091; 5,614,617; 5,645,985; 5,681,941; 5,750,692; 5,763,588; 5,830,653 и 6,005,096; Kandimalla et al., (2001) Bioorg. Med. Chem. 9:807-813; The Concise Encyclopedia of Polymer Science and Engineering, Kroschwitz, J.I., Ed., John Wiley & Sons, 1990, 858- 859; Englisch et al., Angewandte Chemie, International Edition,

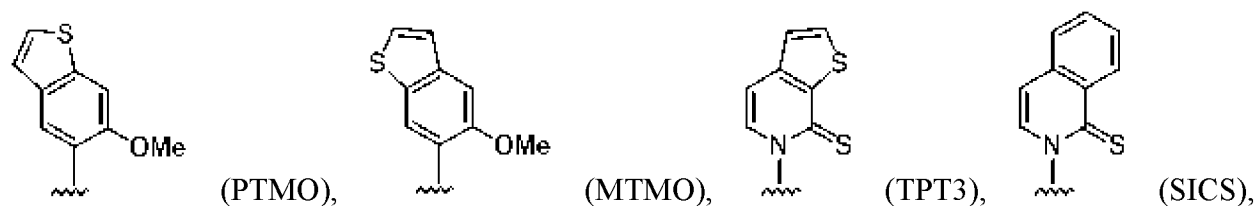
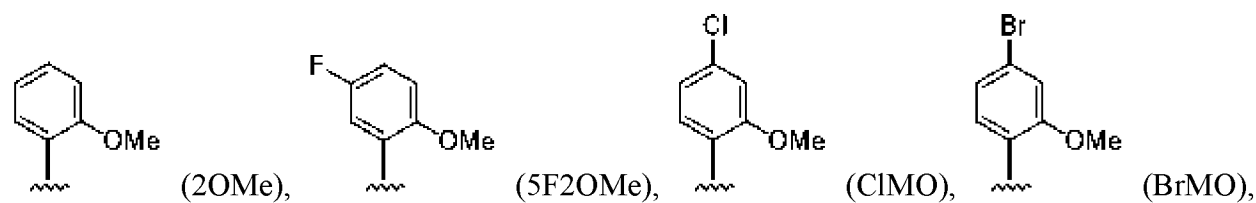
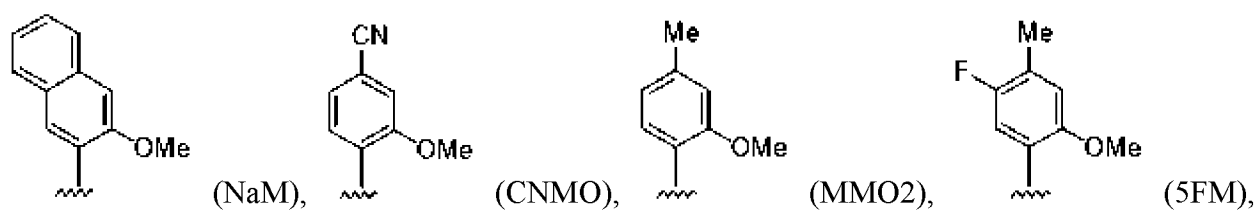
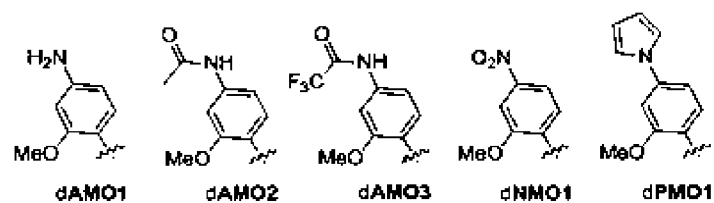
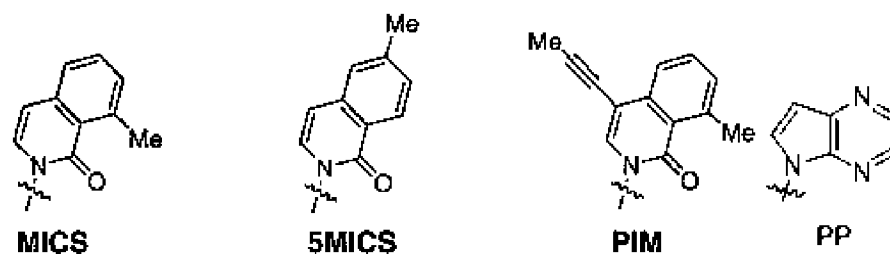
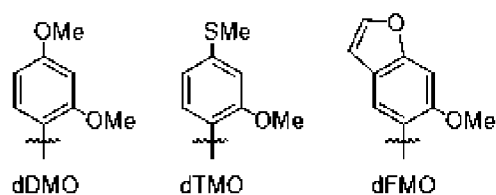
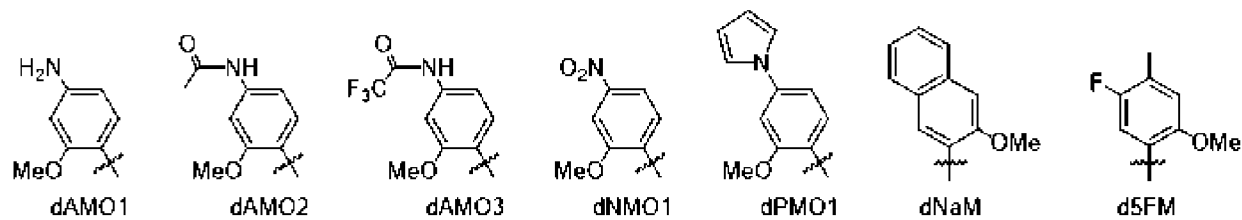
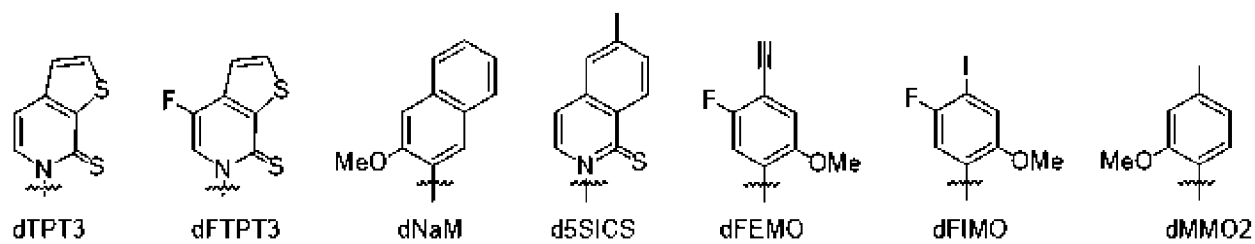
1991, 30, 613; и Sanghvi, Chapter 15, Antisense Research and Applications, Crooke and Lebleu Eds., CRC Press, 1993, 273-288. Дополнительные модификации оснований можно найти, например, в патенте США № 3,687,808, Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613. В некоторых случаях, нуклеиновая кислота не природного происхождения содержит нуклеотидное основание с **ФИГ. 2**. В некоторых случаях, нуклеиновая кислота не природного происхождения содержит нуклеотидное основание с **ФИГ. 3А**. В некоторых случаях, нуклеиновая кислота не природного происхождения содержит нуклеотидное основание с **ФИГ. 3В**.

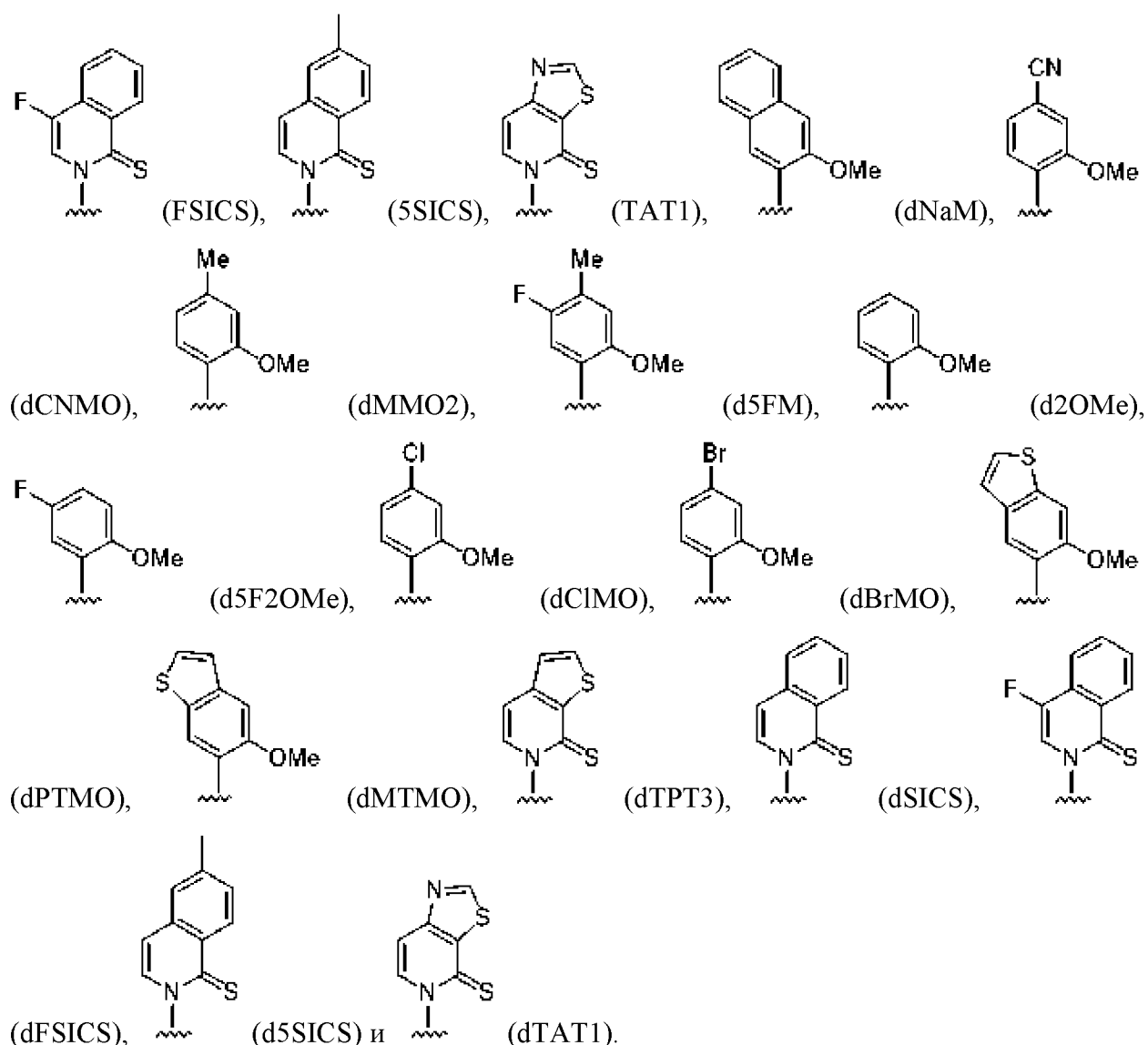
Нуклеиновые кислоты не природного происхождения, содержащие различные гетероциклические основания и различные группы сахаров (и аналоги сахаров), доступны в данной области техники, и нуклеиновые кислоты в некоторых случаях включают одно или несколько гетероциклических оснований, отличных от пяти основных компонентов природных нуклеиновых кислот. Например, гетероциклическое основание включает в некоторых случаях группы урацил-5-ил, цитозин-5-ил, аденин-7-ил, аденин-8-ил, гуанин-7-ил, гуанин-8-ил, 4-аминопирроло[2,3-d]пиримидин-5-ил, 2-амино-4-оксопирроло[2,3-d]пиримидин-5-ил, 2-амино-4-оксопирроло[2,3-d]пиримидин-3-ил, где пурины присоединены к сахарной группе нуклеиновой кислоты через положение 9, пиримидины через положение 1, пирролопиримидины через положение 7 и пиразолопиримидины через положение 1.

В некоторых вариантах осуществления, модифицированное основание не встречающейся в природе нуклеиновой кислоты изображено ниже, где волнистой линией обозначена точка присоединения к дезоксирибозе или рибозе.









В некоторых вариантах осуществления, аналоги нуклеотидов также модифицированы на фосфатной группе. Модифицированные фосфатные группы включают, но не ограничены ими, соединения с модификацией на связи между двумя нуклеотидами и содержат, например, фосфоротиоат, хиральный фосфоротиоат, фосфородитиоат, фосфотриэфир, аминоксилфосфотриэфир, метил и другие алкилфосфонаты, включая 3'-алкиленфосфонат и хиральные фосфонаты, фосфинаты, фосфорамидаты, включая 3'-аминофосфорамидат и аминоксилфосфорамидаты, тионофосфорамидаты, тионоалкилфосфонаты, тионоалкилфосфотриэфиры и боронофосфаты. Понятно, что эти фосфатные или модифицированные фосфатные связи между двумя нуклеотидами проходят через связь 3'-5' или связь 2'-5', и связь имеет обратную полярность, такую как 3'-5' к 5'-3' или 2'-5' к 5'-2'. Также включены различные соли, смешанные соли и формы свободных кислот. Многочисленные патенты США учат тому, как получать и использовать нуклеотиды, содержащие модифицированные фосфаты, и включают, но не ограничены ими, 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111;

5,563,253; 5,571,799; 5,587,361; и 5,625,050.

В некоторых вариантах осуществления, нуклеиновые кислоты не природного происхождения включают 2',3'-дидезокси-2',3'-дидегидронуклеозиды (PCT/US2002/006460), 5'-замещенные производные ДНК и РНК (PCT/US2011/033961; Saha et al., *J. Org Chem.*, 1995, 60, 788-789; Wang et al., *Bioorganic & Medicinal Chemistry Letters*, 1999, 9, 885-890; и Mikhailov et al., *Nucleosides & Nucleotides*, 1991, 10(1-3), 339-343; Leonid et al., 1995, 14(3-5), 901-905; и Eppacher et al., *Helvetica Chimica Acta*, 2004, 87, 3004-3020; PCT/JP2000/004720; PCT/JP2003/002342; PCT/JP2004/013216; PCT/JP2005/020435; PCT/JP2006/315479; PCT/JP2006/324484; PCT/JP2009/056718; PCT/JP2010/067560), или 5'-замещенные мономеры, сделанные в виде монофосфата с модифицированными основаниями (Wang et al., *Nucleosides Nucleotides & Nucleic Acids*, 2004, 23 (1 & 2), 317-337).

В некоторых вариантах осуществления, нуклеиновые кислоты не природного происхождения включают модификации в 5'-положении и 2'-положении сахарного кольца (PCT/US94/02993), такие как 5'-CH₂-замещенные 2'-О-защищенные нуклеозиды (Wu et al., *Helvetica Chimica Acta*, 2000, 83, 1127-1143 and Wu et al., *Bioconjugate Chem.* 1999, 10, 921-924). В некоторых случаях, нуклеиновые кислоты не природного происхождения включают амид-связанные нуклеозидные димеры, которые были получены для введения в олигонуклеотиды, где 3'-связанный нуклеозид в димере, (5'-3') содержит 2'-OCH₃ и 5'-(S)-CH₃ (Mesmaeker et al., *Synlett*, 1997, 1287-1290). Нуклеиновые кислоты не природного происхождения могут включать 2'-замещенные 5'-CH₂ (или О) модифицированные нуклеозиды (PCT/US92/01020). нуклеиновые кислоты не природного происхождения могут включать 5'-метиленфосфонатные ДНК и РНК мономеры и димеры (Bohringer et al., *Tet. Lett.*, 1993, 34, 2723-2726; Collingwood et al., *Synlett*, 1995, 7, 703-705; and Hutter et al., *Helvetica Chimica Acta*, 2002, 85, 2777-2806). Нуклеиновые кислоты не природного происхождения могут включать 5'-фосфонатные мономеры с 2'-заменой (US 2006/0074035) и другие модифицированные 5'-фосфонатные мономеры (WO 1997/35869). Нуклеиновые кислоты не природного происхождения могут включать 5'-модифицированные метиленфосфонатные мономеры (EP614907 и EP629633). Нуклеиновые кислоты не природного происхождения могут включать аналоги 5'- или 6'-фосфонатных рибонуклеозидов, содержащие гидроксильную группу в 5'- и/или 6'-положении (Chen et al., *Phosphorus, Sulfur and Silicon*, 2002, 777, 1783-1786; Jung et al., *Bioorg. Med. Chem.*, 2000, 8, 2501-2509; Gallier et al., *Eur. J. Org. Chem.*, 2007, 925-933; and Hampton et al., *J. Med. Chem.*, 1976, 19(8), 1029-1033). Нуклеиновые кислоты не природного происхождения могут включать 5'-фосфонатдезоксирибонуклеозидные мономеры и димеры, имеющие 5'-фосфатную группу (Nawrot et al., *Oligonucleotides*, 2006, 16(1), 68-82). Нуклеиновые кислоты не природного происхождения могут включать нуклеозиды, имеющие 6'-фосфонатную группу, где 5'- и/или 6'-положение не замещено или замещено тио-трет-бутильной группой (SC(CH₃)₃) (и их аналоги); метиленаминогруппой (CH₂NH₂) (и их аналоги) или цианогруппой (CN) (и их аналоги) (Fairhurst et al., *Synlett*, 2001, 4, 467-472;

Kappler et al., J. Med. Chem., 1986, 29, 1030-1038; Kappler et al., J. Med. Chem., 1982, 25, 1179-1184; Vrudhula et al., J. Med. Chem., 1987, 30, 888-894; Hampton et al., J. Med. Chem., 1976, 19, 1371-1377; Geze et al., J. Am. Chem. Soc, 1983, 105(26), 7638-7640; и Hampton et al., J. Am. Chem. Soc, 1973, 95(13), 4404-4414).

В некоторых вариантах осуществления, нуклеиновые кислоты не природного происхождения также включают модификации сахарной группы. В некоторых случаях, нуклеиновые кислоты содержат один или несколько нуклеозидов, в которых сахарная группа была модифицирована. Такие сахар-модифицированные нуклеозиды могут придавать повышенную нуклеазную стабильность, повышенную аффинность связывания или некоторые другие полезные биологические свойства. В некоторых вариантах осуществления, нуклеиновые кислоты содержат химически модифицированную рибофуранозную кольцевую группу. Примеры химически модифицированных рибофуранозных колец включают, без ограничений, добавление замещающих групп (включая 5'- и/или 2'-замещающие группы; связывание мостиком двух кольцевых атомов с образованием бициклических нуклеиновых кислот (БНК); замену атома кислорода рибозильного кольца с S, N(R) или C(R₁)(R₂) (R=H, C₁-C₁₂ алкил или защитная группа); и их комбинации. Примеры химически модифицированных сахаров могут быть найдены в WO 2008/101157, US 2005/0130923 и WO 2007/134181.

В некоторых случаях, модифицированная нуклеиновая кислота включает модифицированные сахара или аналоги сахаров. Таким образом, в дополнение к рибозе и дезоксирибозе, сахарная группа может представлять собой пентозу, дезоксипентозу, гексозу, дезоксигексозу, глюкозу, арабинозу, ксилозу, ликсозу или сахарную «аналог» цикlopентильную группу. Сахар может быть в пиранозильной или фуранозильной форме. Сахарная группа может представлять собой фуранозид рибозы, дезоксирибозы, арабинозы или 2'-О-алкилрибозы, и сахар может быть присоединен к соответствующим гетероциклическим основаниям либо в [альфа], либо в [бета] аномерной конфигурации. Сахарные модификации включают, но не ограничены ими, аналоги 2'-алкокси-РНК, аналоги 2'-амино-РНК, 2'-фтор-ДНК и химеры 2'-алкокси- или амино-РНК/ДНК. Например, модификация сахара может включать 2'-О-метилуридин или 2'-О-метилцитидин. Сахарные модификации включают 2'-О-алкилзамещенные дезоксирибонуклеозиды и 2'-О-этиленгликоль-подобные рибонуклеозиды. Получение этих сахаров или аналоги сахаров и соответствующие «нуклеозиды», где такие сахара или аналоги присоединены к гетероциклическому основанию (основанию нуклеиновой кислоты) известны. Можно также производить сахарные модификации и комбинировать их с другими модификациями.

Модификации группы сахара включают природные модификации рибозы и дезоксирибозы, а также модификации не природного происхождения. Сахарные модификации включают, но не ограничены ими, следующие модификации в положении 2': OH, F, O-, S- или N-алкил; O-, S- или N-алкенил; O-, S- или N-алкинил; или O-алкил-O-алкил, где алкил, алкенил и алкинил могут быть замещенным или незамещенным C₁-C₁₀ алкилом или C₂-C₁₀ алкенилом и алкинилом. 2' сахарные модификации также включают, но

не ограничены ими, $-O[(CH_2)_nO]_mCH_3$, $-O(CH_2)_nOCH_3$, $-O(CH_2)_nNH_2$, $-O(CH_2)_nCH_3$, $-O(CH_2)_nONH_2$ и $-O(CH_2)_nON[(CH_2)_nCH_3]_2$, где n и m составляют от 1 до примерно 10.

Другие модификации в положении 2' включают, но не ограничены ими: C_1 - C_{10} низший алкил, замещенный низший алкил, алкарил, аралкил, О-алкарил, О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, гетероциклоалкил, гетероциклоалкарил, аминоалкиламино, полиалкиламино, замещенный силил, расщепляющую РНК группу, репортерную группу, интеркалятор, группу для улучшения фармакокинетических свойств олигонуклеотида или группу для улучшения фармакодинамических свойств олигонуклеотида и другие заместители, обладающие аналогичными свойствами. Аналогичные модификации могут быть также сделаны в других положениях сахара, в частности в положении 3' сахара на 3' концевом нуклеотиде или в 2'-5'-связанных олигонуклеотидах, и в положении 5' на 5' концевом нуклеотиде. Модифицированные сахара также включают такие, которые содержат модификации в кислороде мостикового кольца, такие как CH₂ и S. Нуклеотидные аналоги сахара могут также иметь миметики сахара, такие как циклобутильные группы вместо пентофуранозилового сахара. Существует множество патентов США, которые описывают получение таких модифицированных сахарных структур, и которые подробно описывают ряд модификаций оснований, такие как патенты США №№ 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121; 5,596,091; 5,614,617; 5,681,941; и 5,700,920, каждый из которых включен в настоящий документ посредством ссылки.

Примеры нуклеиновых кислот, содержащих модифицированные группы сахара, включают, без ограничения, нуклеиновые кислоты, содержащие 5'-винил, 5'-метил (R или S), 4'-S, 2'-F, 2'-OCH₃ и 2'-O(CH₂)₂OCH₃ группы заместителей. Заместители в положении 2' также могут быть выбраны из аллила, амино, азидо, тио, О-аллила, О-(C₁-C₁₀ алкила), OCF₃, O(CH₂)₂SCH₃, O(CH₂)₂-O-N(R_m)(R_n) и O-CH₂-C(=O)-N(R_m)(R_n), где каждый R_m и R_n независимо представляют собой H или замещенный или незамещенный C₁-C₁₀ алкил.

В некоторых вариантах осуществления, нуклеиновые кислоты, описанные в настоящем документе, включают одну или несколько бициклических нуклеиновых кислот. В некоторых таких вариантах осуществления, бициклические нуклеиновые кислоты содержат мостик между 4' и 2' атомами рибозильного кольца. В некоторых вариантах осуществления, нуклеиновые кислоты, представленные в настоящем документе, включают одно или несколько бициклических нуклеиновых кислот, где мостик содержит 4'-2' бициклическую нуклеиновую кислоту. Примеры таких 4'-2' бициклических нуклеиновых кислот включают, но не ограничены ими, одну из формул: 4'-(CH₂)-O-2' (LNA); 4'-(CH₂)-S-2'; 4'-(CH₂)₂-O-2' (ENA); 4'-CH(CH₃)-O-2' и 4'-CH(CH₂OCH₃)-O-2' и их аналоги (см. патент США №7,399,845), 4'-C(CH₃)(CH₃)-O-2' и ее аналоги (см. WO2009/006478,

WO2008/150729, US 2004/0171570, патент США № 7,427,672, Chattopadhyaya et al., J. Org. Chem., 209, 74, 118-134, и WO 2008/154401). См. также, например: 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; Singh et al., J. Org. Chem., 1998, 63, 10035-10039; Srivastava et al., J. Am. Chem. Soc., 2007, 129(26) 8362-8379; Elayadi et al., Curr. Opinion Invens. Drugs, 2001, 2, 558-561; Braasch et al., Chem. Biol, 2001, 8, 1-7; Oram et al., Curr. Opinion Mol. Ther., 2001, 3, 239-243; патенты США №№ 4,849,513; 5,015,733; 5,118,800; 5,118,802; 7,053,207; 6,268,490; 6,770,748; 6,794,499; 7,034,133; 6,525,191; 6,670,461; и 7,399,845; международные публикации №№ WO2004/106356, WO1994/14226, WO2005/021570, WO2007/090071, и WO2007/134181; заявки на патент США №№ US2004/0171570, US2007/0287831 и US2008/0039618; предварительные заявки США №№ 60/989,574, 61/026,995, 61/026,998, 61/056,564, 61/086,231, 61/097,787 и 61/099,844; и международные заявки №№ PCT/US2008/064591, PCT US2008/066154, PCT US2008/068922 и PCT/DK98/00393.

В некоторых вариантах осуществления, нуклеиновые кислоты содержат связанные нуклеиновые кислоты. Нуклеиновые кислоты могут быть связаны друг с другом с помощью любой связи между нуклеиновыми кислотами. Два основных класса связывающих групп между нуклеиновыми кислотами определяются присутствием или отсутствием атома фосфора. Типовые фосфорсодержащие связи между нуклеиновыми кислотами включают, но не ограничены ими, фосфодиэфиры, фосфотриэфиры, метилфосфонаты, фосфорамидаты и фосфоротиоаты (P=S). Типовые не содержащие фосфор связывающие группы между нуклеиновыми кислотами включают, но не ограничены ими, метиленметилямино (-CH₂-N(CH₃)-O-CH₂-), тиодиэфир (-OC(O)-S-), тионокарбамат (-OC(O)(NH)-S-), силоксан (-O-Si(H)₂-O-); и N, N*-диметилгидразин (-CH₂-N(CH₃)-N(CH₃)). В некоторых вариантах осуществления, связи между нуклеиновыми кислотами, имеющие хиральный атом, могут быть получены в виде рацемической смеси, в виде отдельных энантиомеров, *например*, алкилфосфонатов и фосфоротиоатов. Нуклеиновые кислоты не природного происхождения могут содержать одну модификацию. Нуклеиновые кислоты не природного происхождения могут содержать несколько модификаций в пределах одной из групп или между различными группами.

Фосфатные модификации остова нуклеиновой кислоты включают, но не ограничены ими, метилфосфонат, фосфоротиоат, фосфорамидат (мостиковый или не мостиковый), фосфотриэфир, фосфородитиоат, фосфодитиоат и боранофосфат, и могут использоваться в любой комбинации. Другие не фосфатные связи также могут использоваться.

В некоторых вариантах осуществления, модификации остова (*например*, метилфосфонатные, фосфоротиоатные, фосфорамидатные и фосфородитиоатные межнуклеотидные связи) могут придавать модифицированной нуклеиновой кислоте иммуномодулирующую активность и/или повышать ее стабильность *in vivo*.

В некоторых случаях, производное фосфора (или модифицированная фосфатная группа) присоединено к группе сахара или аналога сахара, и может представлять собой

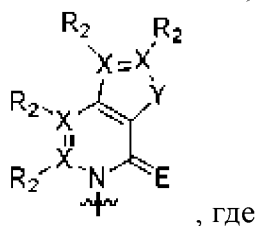
монофосфат, дифосфат, трифосфат, алкилфосфонат, фосфоротиоат, фосфородитиоат, фосфорамидат и подобные. Примеры полинуклеотидов, содержащих модифицированные фосфатные связи или не фосфатные связи можно найти у Peyrottes et al., 1996, *Nucleic Acids Res.* 24: 1841-1848; Chaturvedi et al., 1996, *Nucleic Acids Res.* 24:2318-2323; и Schultz et al., (1996) *Nucleic Acids Res.* 24:2966-2973; Matteucci, 1997, "Oligonucleotide Analogs: an Overview" in *Oligonucleotides as Therapeutic Agents*, (Chadwick and Cardew, ed.) John Wiley and Sons, New York, NY; Zon, 1993, "Oligonucleoside Phosphorothioates" in *Protocols for Oligonucleotides and Analogs, Synthesis and Properties*, Humana Press, pp. 165-190; Miller et al., 1971, *JACS* 93:6657-6665; Jager et al., 1988, *Biochem.* 27:7247-7246; Nelson et al., 1997, *JOC* 62:7278-7287; патент США № 5,453,496; и Micklefield, 2001, *Curr. Med. Chem.* 8: 1157-1179.

В некоторых случаях, модификация остова включает замену фосфодиэфирной связи альтернативной группой, такой как анионная, нейтральная или катионная группа. Примеры таких модификаций включают: анионную межнуклеозидную связь, N3'-P5' фосфорамидатную модификацию; боранофосфатную ДНК; проолигонуклеотиды; нейтральные межнуклеозидные связи такие как метилфосфонаты; ДНК, связанную амидом; метиленовые (метилямино) связи; формацетальные и тиоформацетальные связи; остовы, содержащие сульфонильные группы; олигонуклеотиды морфолино; пептидные нуклеиновые кислоты (ПНК); и положительно заряженные олигонуклеотиды дезоксирибонуклеинового гуанидина (DNG) (Micklefield, 2001, *Current Medicinal Chemistry* 8: 1157-1179). Модифицированная нуклеиновая кислота может содержать химерный или смешанный остов, содержащий одну или несколько модификаций, *например*, комбинацию фосфатных связей, такую как комбинация фосфодиэфирных и тиофосфатных связей.

Заменители фосфата включают, например, короткоцепочечные алкильные или циклоалкильные межнуклеозидные связи, смешанные гетероатомные и алкильные или циклоалкильные межнуклеозидные связи или одну или несколько короткоцепочечных гетероатомных или гетероциклических межнуклеозидных связей. Они включают такие, которые имеют морфолино связи (образованные частично из сахарной части нуклеозида); силоксановые остовы; сульфидные, сульфоксидные и сульфоновые остовы; формацетильные и тиоформацетильные остовы; метиленформацетильные и тиоформацетильные остовы; алкенсодержащие остовы; сульфаматные остовы; метиленимино и метиленигидразино остовы; сульфонатные и сульфонамидные остовы; амидные остовы; и другие, имеющие смешанные N, O, S и CH₂ части компонентов. Многочисленные патенты США описывают как производить и использовать эти типы заменителей фосфатов, и включают, но не ограничены ими, патенты США 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; и 5,677,439. Также необходимо понимать, что при замещении нуклеотида и сахар, и фосфатные группы нуклеотида может быть замещены, например, связью амидного типа (аминоэтилглицин) (ПНК). В патентах США №№ 5,539,082; 5,714,331 и 5,719,262

описано, как получать и использовать молекулы ПНК, каждый из которых включен в настоящий документ в качестве ссылки. См. также Nielsen et al., Science, 1991, 254, 1497-1500. Также можно связать другие типы молекул (конъюгаты) с нуклеотидами или аналогами нуклеотидов для усиления, например, клеточного поглощения. Конъюгаты могут быть химически связаны с нуклеотидом или аналогами нуклеотидов. Такие конъюгаты включают, но не ограничены ими, липидные группы, такие как холестериновая группа (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556), холевую кислоту (Manoharan et al., Bioorg. Med. Chem. Lett., 1994, 4, 1053-1060), тиозфир, *например*, гексил-S-тримитилтиол (Manoharan et al., Ann. KY. Acad. Sci., 1992, 660, 306-309; Manoharan et al., Bioorg. Med. Chem. Lett., 1993, 3, 2765-2770), тиохолестерин (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533-538), алифатическую цепь, *например*, додекандиольные или ундецильные остатки (Saison-Behmoaras et al., EMBOJ, 1991, 10, 1111-1118; Kabanov et al., FEBS Lett., 1990, 259, 327-330; Svinarchuk et al., Biochimie, 1993, 75, 49-54), фосфолипид, *например*, ди-гексадецил-рац-глицерин или триэтиламмоний 1-ди-О-гексадецил-рац-глицеро-SH-фосфонат (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777-3783), цепь полиамина или полиэтиленгликоля (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969-973), или адамантануксусную кислоту (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654), пальмитиновую группу (Mishra et al., Biochem. Biophys. Acta, 1995, 1264, 229-237) или октадециламин- или гексиламинокарбонил-оксихолестериновую группу (Crooke et al., J. Pharmacol. Exp. Ther., 1996, 277, 923-937). Многочисленные патенты США описывают получение таких конъюгатов и включают, но не ограничены ими, патенты США №№ 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717; 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241; 5,391,723; 5,416,203; 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 и 5,688,941.

В настоящем документе описаны нуклеотидные основания, используемые в композициях и способах репликации, транскрипции, трансляции и включения аминокислот не природного происхождения в белки. В некоторых вариантах осуществления, нуклеотидное основание, описанное в настоящем документе, имеет структуру:



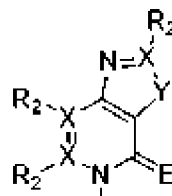
каждый X независимо представляет собой углерод или азот;

R_2 является необязательным, и когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселено, галоген, циано или азидную группу;

где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин;

где каждый E независимо представляет собой кислород, серу или селен; и

где волнистая линия указывает на точку связывания с рибозильной, дезоксирибозильной или дидезоксирибозильной группой или их аналогом, где рибозильная, дезоксирибозильная или дидезоксирибозильная группа или их аналог находятся в свободной форме, соединены с монофосфатной, дифосфатной или трифосфатной группой, необязательно содержащей α -тиотрифосфатную, β -тиотрифосфатную или γ -тиотрифосфатную группу, или включенную в РНК или ДНК, или в аналог РНК, или аналог ДНК. В некоторых вариантах осуществления, R_2 представляет собой низший алкил (например, C_1 - C_6), водород или галоген. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, R_2 представляет собой фтор. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, X представляет собой углерод. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, E представляет собой серу. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, Y представляет собой серу. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем



документе, нуклеотидное основание имеет структуру: . В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, E представляет собой серу, и Y представляет собой серу. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, волнистая линия указывает точку связывания с рибозильной или дезоксирибозильной группой. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, волнистая линия указывает точку связывания с рибозильной или дезоксирибозильной группой, соединенной с трифосфатной группой. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание представляет собой компонент тРНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе,

нуклеотидное основание представляет собой компонент антикодона в тРНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание представляет собой компонент иРНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание является компонентом кодона иРНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание представляет собой компонент РНК или ДНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание представляет собой компонент кодона ДНК. В некоторых вариантах осуществления нуклеотидного основания, описанного в настоящем документе, нуклеотидное основание образует пару нуклеотидных оснований с другим комплементарным нуклеотидным основанием.

Дезоксирибонуклеиновая кислота не природного происхождения (ДНК), в некоторых случаях, транскрибируется в информационную РНК (иРНК), содержащую основания не природного происхождения, описанные в настоящем документе (*например*, d5SICS, dNaM, dTPT3, dMTMO, dCNMO, dTAT1). Типовые кодоны иРНК кодируются типовыми областями ДНК не природного происхождения, содержащей три последовательных дезоксирибонуклеотида (NNN), включая TTX, TGX, CGX, AGX, GAX, CAX, GXT, CXT, GXG, AXG, GXC, AXC, GXA, CXC, TXC, ATX, CTX, TTX, GTX, TAX или GGX, где X представляет собой основание не природного происхождения, присоединенное к 2' дезоксирибозильной группе. Типовые кодоны иРНК, полученные в результате транскрипции ДНК не природного происхождения, содержат три непрерывных рибонуклеотида (NNN), включая UUX, UGX, CGX, AGX, GAX, CAX, GXU, CXU, GXG, AXG, GXC, AXC, GXA, CXC, UXC, AUX, CUX, UUX, GUX, UAX или GGX, соответственно, где X представляет собой основание не природного происхождения, присоединенное с рибозильной группой. В некоторых вариантах осуществления, основание не природного происхождения находится в первом положении в последовательности кодона (X-N-N). В некоторых вариантах осуществления, основание не природного происхождения находится во втором (или среднем) положении в последовательности кодона (N-X-N). В некоторых вариантах осуществления, основание не природного происхождения находится в третьем (последнем) положении в последовательности кодона (N-N-X).

иРНК, содержащая кодоны, описанные в настоящем документе, в некоторых случаях транслируется *in vivo* в клетке (*например*, в эукариотической клетке). Трансляция иРНК, содержащая основание не природного происхождения, описанное в настоящем документе, опосредуется транспортной РНК (тРНК), содержащей последовательность антикодона, которая является обратным комплементом последовательности кодона иРНК, описанной в настоящем документе. В некоторых вариантах осуществления, антикодон

тРНК содержит основание не природного происхождения, содержащее YAA, XAA, YCA, XCA, YCG, XCG, YCU, XCU, YUC, XUC, YUG, XUG, AYC, AYG, CYC, CYU, GYC, GYU, UYC, GYG, GYA, YAU, XAU, XAG, YAG, XAC, YAC, XUA, YUA, XCC или YCC, где X и Y, каждый представляет собой основание не природного происхождения, где X и Y не являются одинаковыми. В некоторых вариантах осуществления, основание не природного происхождения находится в первом положении в последовательности антикодона (X/YN-N). В некоторых вариантах осуществления, основание не природного происхождения находится во втором (или среднем) положении в последовательности антикодона. (NX/YN). В некоторых вариантах осуществления, основание не природного происхождения находится в третьем (последнем) положении в последовательности антикодона (N-NX/Y).

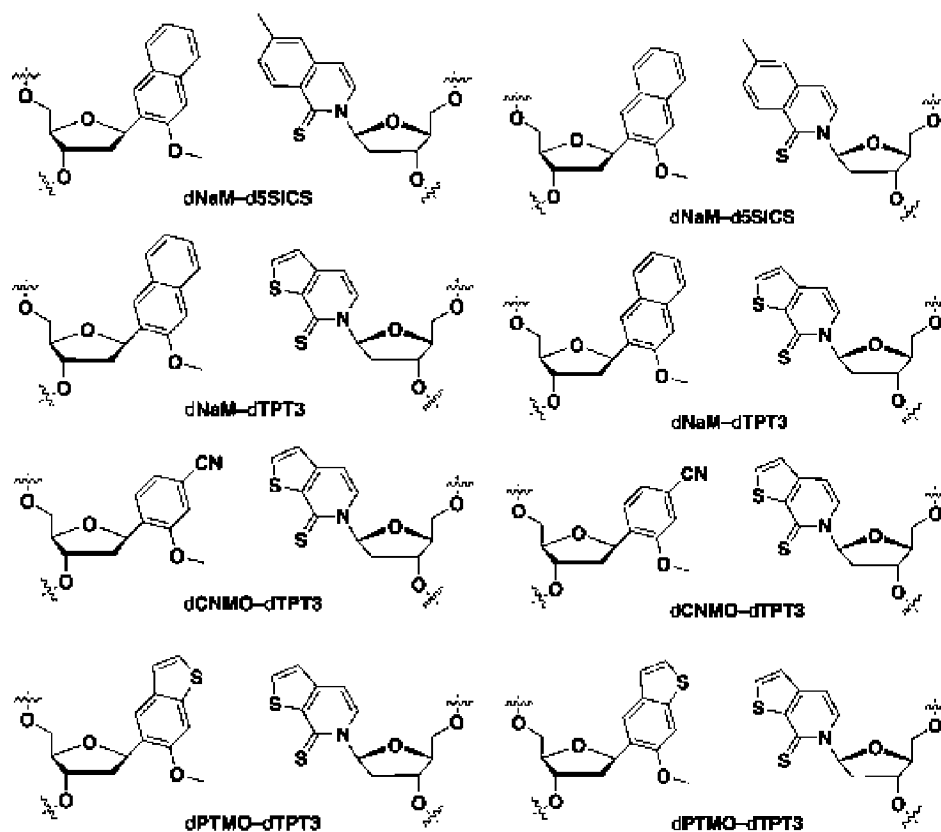
Свойства Спаривания Оснований Нуклеиновой Кислоты

В некоторых вариантах осуществления, нуклеотид не природного происхождения образует пару оснований (пару оснований не природного происхождения; UBP) с другим нуклеотидом не природного происхождения, например, во время трансляции. Например, первая нуклеиновая кислота не природного происхождения может образовывать пару оснований со второй нуклеиновой кислотой не природного происхождения. Например, одна пара нуклеозидтрифосфатов не природного происхождения, которые могут образовывать пары оснований, например, во время трансляции, включает нуклеотид, содержащий (d)5SICS, и нуклеотид, содержащий (d)NaM. Другие примеры включают, но не ограничиваются ими: нуклеотид, содержащий (d)CNMO и нуклеотид, содержащий (d)TPT3. Такие нуклеотиды не природного происхождения могут иметь группу сахара рибозы или дезоксирибозы (обозначается «(d)»). Например, одна пара нуклеозидтрифосфатов не природного происхождения, которые могут образовывать пары оснований при включении в нуклеиновые кислоты, включает нуклеотид, содержащий TAT1, и нуклеотид, содержащий NaM. В некоторых вариантах осуществления, одна пара нуклеозидтрифосфатов не природного происхождения, которые могут образовывать пары оснований при включении в нуклеиновые кислоты, включает нуклеотид dCNMO и нуклеотид, содержащий TAT1. В некоторых вариантах осуществления, одна пара нуклеозидтрифосфатов не природного происхождения, которые могут образовывать пары оснований при включении в нуклеиновые кислоты, включает нуклеотид dTPT3 и нуклеотид, содержащий NaM. В некоторых вариантах осуществления, нуклеиновая кислота не природного происхождения по существу не образует пару оснований с природной нуклеиновой кислотой (A, T, G, C). В некоторых вариантах осуществления, нуклеиновая кислота не природного происхождения может образовывать пару оснований с природной нуклеиновой кислотой.

В некоторых вариантах осуществления, (дезокси)рибонуклеотид не природного происхождения представляет собой (дезокси)рибонуклеотид не природного происхождения, который может образовывать UBP, но по существу не образует пару оснований с каждым из природных (дезокси)рибонуклеотидов. В некоторых вариантах осуществления, (дезокси)рибонуклеотид не природного происхождения представляет

собой (дезокси)рибонуклеотид не природного происхождения, который может образовывать UBP, но по существу не образует пару оснований с одной или несколькими природными нуклеиновыми кислотами. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А, Т и С, но может образовывать пару оснований с G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А, Т и G, но может образовывать пару оснований с С. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с С, G и А, но может образовывать пару оснований с Т. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с С, G и Т, но может образовывать пару оснований с А. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А и Т, но может образовывать пару оснований с С и G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А и С, но может образовывать пару оснований с Т и G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А и G, но может образовывать пару оснований с С и Т. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с С и Т, но может образовывать пару оснований с А и G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с С и G, но может образовывать пару оснований с Т и G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с Т и G, но может образовывать пару оснований с А и G. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с G, но может образовывать пару оснований с А, Т и С. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с А, но может образовывать пару оснований с G, Т и С. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с Т, но может образовывать пару оснований с G, А и С. Например, нуклеиновая кислота не природного происхождения может по существу не образовывать пару оснований с С, но может образовывать пара оснований с G, Т и А.

Примеры нуклеотидов не природного происхождения, способных образовывать пару оснований не природного происхождения (UBP) (например, в РНК, например, между тРНК и иРНК) в условиях *in vivo* включают, но не ограничены ими, 5SICS, d5SICS, NaM, dNaM, dTPT3, dMTMO, dCNMO, TAT1 и их комбинации. В некоторых вариантах осуществления, пары нуклеотидных оснований не природного происхождения включают, но не ограничены ими:



и их соответствующие рибо (РНК) формы.

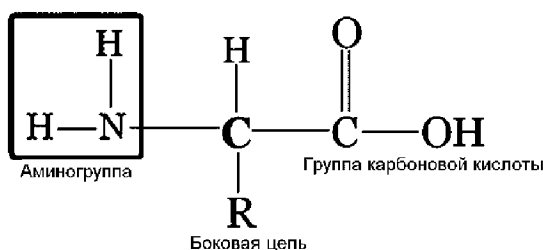
Пары оснований не природного происхождения (УВР) образуются между последовательностью кодона иРНК и последовательностью антикодона тРНК, чтобы способствовать трансляции иРНК в полипептид не природного происхождения. УВР кодон-антикодон содержат, в некоторых случаях, последовательность кодона, содержащую три готовых непрерывных аминокислоты 5'-3' иРНК (*например*, UUX), и последовательность антикодона, содержащую три готовых непрерывных аминокислоты 5'-3' тРНК (*например*, YAA или XAA). В некоторых вариантах осуществления, когда кодон иРНК представляет собой UUX, антикодон тРНК представляет собой YAA или XAA. В некоторых вариантах осуществления, когда кодон иРНК представляет собой UGX, антикодон тРНК представляет собой YCA или XCA. В некоторых вариантах осуществления, когда кодон иРНК представляет собой CGX, антикодон тРНК представляет собой YCG или XCG. В некоторых вариантах осуществления, когда кодон иРНК представляет собой AGX, антикодон тРНК представляет собой YCU или XCU. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GAX, антикодон тРНК представляет собой YUC или XUC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой CAX, антикодон тРНК представляет собой YUG или XUG. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GXU, антикодон тРНК представляет собой AYC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой CXU, антикодон тРНК представляет собой AYG. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GXG, антикодон тРНК представляет собой CYC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой AXG,

антикодон тРНК представляет собой CYU. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GXC, антикодон тРНК представляет собой GYC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой AXC, антикодон тРНК представляет собой GYU. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GXA, антикодон тРНК представляет собой UYC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой CXC, антикодон тРНК представляет собой GYG. В некоторых вариантах осуществления, когда кодон иРНК представляет собой UXC, антикодон тРНК представляет собой GYA. В некоторых вариантах осуществления, когда кодон иРНК представляет собой AUX, антикодон тРНК представляет собой YAU или XAU. В некоторых вариантах осуществления, когда кодон иРНК представляет собой CUX, антикодон тРНК представляет собой XAG или YAG. В некоторых вариантах осуществления, когда кодон иРНК представляет собой UUX, антикодон тРНК представляет собой XAA или YAA. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GUX, антикодон тРНК представляет собой XAC или YAC. В некоторых вариантах осуществления, когда кодон иРНК представляет собой UAX, антикодон тРНК представляет собой XUA или YUA. В некоторых вариантах осуществления, когда кодон иРНК представляет собой GGX, антикодон тРНК представляет собой XCC или YCC.

Аминокислоты Природного и Не Природного Происхождения

В настоящем документе, аминокислотный остаток может относиться к молекуле, содержащей как аминогруппу, так и карбоксильную группу. Подходящие аминокислоты включают, без ограничения, как D-, так и L-изомеры встречающихся в природе аминокислот, а также аминокислот не природного происхождения, полученных органическим синтезом или любым другим способом. Термин аминокислота, используемый в настоящем документе, включает, без ограничения, α -аминокислоты, природные аминокислоты, аминокислоты не природного происхождения и аналоги аминокислот.

Термин « α -аминокислота» может относиться к молекуле, содержащей как аминогруппу, так и карбоксильную группу, связанную с углеродом, который обозначается как α -углерод. Например:



Термин « β -аминокислота» может относиться к молекуле, содержащей как аминогруппу, так и карбоксильную группу в β -конфигурации.

«Встречающаяся в природе аминокислота» может относиться к любой из двадцати аминокислот, обычно встречающихся в пептидах, синтезированных в природе и известных

под однобуквенными сокращениями A, R, N, C, D, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y и V.

В следующей таблице приведены краткие сведения о свойствах природных аминокислот:

Аминокислота	3- буквенный код	1- буквенный код	Полярнос ть боковой цепи	Заряд боковой цепи (рН 7,4)	Индекс гидропатии
Аланин	Ala	A	неполярна я	нейтральный	1,8
Аргинин	Arg	R	полярная	положительный	-4,5
Аспарагин	Asn	N	полярная	нейтральный	-3,5
Аспарагиновая кислота	Asp	D	полярная	отрицательный	-3,5
Цистеин	Cys	C	полярная	нейтральный	2,5
Глутаминовая кислота	Glu	E	полярная	отрицательный	-3,5
Глутамин	Gln	Q	полярная	нейтральный	-3,5
Глицин	Gly	G	неполярна я	нейтральный	-0,4
Гистидин	His	H	полярная	положительный (10%) нейтральный (90%)	-3,2
Изолейцин	Ile	I	неполярна я	нейтральный	4,5
Лейцин	Leu	L	неполярна я	нейтральный	3,8
Лизин	Lys	K	полярная	положительный	-3,9
Метионин	Met	M	неполярна я	нейтральный	1,9
Фенилаланин	Phe	F	неполярна я	нейтральный	2,8

Пролин	Pro	P	неполярная	нейтральный	-1,6
Серин	Ser	S	полярная	нейтральный	-0,8
Треонин	Thr	T	полярная	нейтральный	-0,7
Триптофан	Trp	W	неполярная	нейтральный	-0,9
Тирозин	Tyr	Y	полярная	нейтральный	-1,3
Валин	Val	V	неполярная	нейтральный	4,2

«Гидрофобные аминокислоты» включают малые гидрофобные аминокислоты и большие гидрофобные аминокислоты. «Малые гидрофобные аминокислоты» могут представлять собой глицин, аланин, пролин и их аналоги. «Большие гидрофобные аминокислоты» могут представлять собой валин, лейцин, изолейцин, фенилаланин, метионин, триптофан и их аналоги. «Полярными аминокислотами» могут быть серин, треонин, аспарагин, глутамин, цистеин, тирозин и их аналоги. «Заряженными аминокислотами» могут быть лизин, аргинин, гистидин, аспартат, глутамат и их аналоги.

«Аналог аминокислоты» может представлять собой молекулу, которая структурно подобна аминокислоте и которая может быть заменена аминокислотой при образовании пептидомиметического макроцикла. Аналоги аминокислот включают, без ограничения, β-аминокислоты и аминокислоты, где амино- или карбоксигруппа замещена аналогично реакционноспособной группой (например, замещение первичного амина вторичным или третичным амином или замещение карбоксигруппы сложным эфиром).

Не каноническая аминокислота (ncAA) или «аминокислота не природного происхождения» может быть аминокислотой, которая не является одной из двадцати аминокислот, обычно встречающихся в пептидах, синтезированных в природе, и известна под однобуквенными сокращениями A, R, N, C, D, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y и V. В некоторых случаях, аминокислоты не природного происхождения представляют собой подмножество не канонических аминокислот.

Аналоги аминокислот могут включать аналоги β-аминокислот. Примеры аналогов β-аминокислот включают, но не ограничены ими, следующие: циклические аналоги β-аминокислот; β-аланин; (R)-β-фенилаланин; (R)-1,2,3,4-тетрагидроизохинолин-3-уксусную кислоту; (R)-3-амино-4-(1-нафтил)масляную кислоту; (R)-3-амино-4-(2,4-дихлорфенил)масляную кислоту; (R)-3-амино-4-(2-хлорфенил)масляную кислоту; (R)-3-амино-4-(2-цианфенил)масляную кислоту; (R)-3-амино-4-(2-фторфенил)масляную кислоту; (R)-3-амино-4-(2-фурил)масляную кислоту; (R)-3-амино-4-(2-метилфенил)масляную кислоту; (R)-3-амино-4-(2-нафтил)масляную кислоту; (R)-3-амино-4-(2-тиенил)масляную кислоту; (R)-3-амино-4-(2-трифторметилфенил)масляную кислоту, (R)-3-амино-4-(3,4-дихлорфенил)масляную кислоту; (R)-3-амино-4-(3,4-

дифторфенил)масляную кислоту; (R)-3-амино-4-(3-бензотиенил)масляную кислоту; (R)-3-амино-4-(3-хлорфенил)масляную кислоту; (R)-3-амино-4-(3-цианфенил)масляную кислоту; (R)-3-амино-4-(3-фторфенил)масляную кислоту; (R)-3-амино-4-(3-метилфенил)масляную кислоту; (R)-3-амино-4-(3-пиридил)масляную кислоту; (R)-3-амино-4-(3-тиенил)масляную кислоту; (R)-3-амино-4-(3-трифторметилфенил)масляную кислоту; (R)-3-амино-4-(4-бромфенил)масляную кислоту; (R)-3-амино-4-(4-хлорфенил)масляную кислоту; (R)-3-амино-4-(4-цианфенил)масляную кислоту; (R)-3-амино-4-(4-фторфенил)масляную кислоту; (R)-3-амино-4-(4-йодофенил)масляную кислоту; (R)-3-амино-4-(4-метилфенил)масляную кислоту; (R)-3-амино-4-(4-нитрофенил)масляную кислоту; (R)-3-амино-4-(4-пиридил)масляную кислоту; (R)-3-амино-4-(4-трифторметилфенил)масляную кислоту; (R)-3-амино-4-пентафторфенилмасляную кислоту; (R)-3-амино-5-гексеновую кислоту; (R)-3-амино-5-гексиновую кислоту; (R)-3-амино-5-фенилпентановую кислоту; (R)-3-амино-6-фенил-5-гексеновую кислоту; (S)-1,2,3,4-тетрагидроизохинолин-3-уксусную кислоту; (S)-3-амино-4-(1-нафтил)масляную кислоту; (S)-3-амино-4-(2,4-дихлорфенил)масляную кислоту; (S)-3-амино-4-(2-хлорфенил)масляную кислоту; (S)-3-амино-4-(2-цианфенил)масляную кислоту; (S)-3-амино-4-(2-фторфенил)масляную кислоту; (S)-3-амино-4-(2-фурил)масляную кислоту; (S)-3-амино-4-(2-метилфенил)масляную кислоту; (S)-3-амино-4-(2-нафтил)масляную кислоту; (S)-3-амино-4-(2-тиенил)масляную кислоту; (S)-3-амино-4-(2-трифторметилфенил)масляную кислоту; (S)-3-амино-4-(3,4-дихлорфенил)масляную кислоту; (S)-3-амино-4-(3,4-дифторфенил)масляную кислоту; (S)-3-амино-4-(3-бензотиенил)масляную кислоту; (S)-3-амино-4-(3-хлорфенил)масляную кислоту; (S)-3-амино-4-(3-цианфенил)масляную кислоту; (S)-3-амино-4-(3-фторфенил)масляную кислоту; (S)-3-амино-4-(3-метилфенил)масляную кислоту; (S)-3-амино-4-(3-пиридил)масляную кислоту; (S)-3-амино-4-(3-тиенил)масляную кислоту; (S)-3-амино-4-(3-трифторметилфенил)масляную кислоту; (S)-3-амино-4-(4-бромфенил)масляную кислоту; (S)-3-амино-4-(4-хлорфенил)масляную кислоту; (S)-3-амино-4-(4-цианфенил)масляную кислоту; (S)-3-амино-4-(4-фторфенил)масляную кислоту; (S)-3-амино-4-(4-йодофенил)масляную кислоту; (S)-3-амино-4-(4-метилфенил)масляную кислоту; (S)-3-амино-4-(4-нитрофенил)масляную кислоту; (S)-3-амино-4-(4-пиридил)масляную кислоту; (S)-3-амино-4-(4-трифторметилфенил)масляную кислоту; (S)-3-амино-4-пентафторфенилмасляную кислоту; (S)-3-амино-5-гексеновую кислоту; (S)-3-амино-5-гексиновую кислоту; (S)-3-амино-5-фенилпентановую кислоту; (S)-3-амино-6-фенил-5-гексеновую кислоту; 1,2,5,6-тетрагидропиридин-3-карбоновую кислоту, 1,2,5,6-тетрагидропиридин-4-карбоновую кислоту, 3-амино-3-(2-хлорфенил)пропионовую кислоту, 3-амино-3-(2-тиенил)пропионовую кислоту; 3-амино-3-(3-бромфенил)пропионовую кислоту; 3-амино-3-(4-хлорфенил)пропионовую кислоту; 3-амино-3-(4-метоксифенил)пропионовую кислоту; 3-амино-4,4,4-трифтормасляную кислоту; 3-аминоадипиновую кислоту; D-β-фенилаланин; β-лейцин; L-β-гомоаланин; γ-бензиловый эфир L-β-гомоаспарагиновой кислоты; δ-бензиловый эфир L-β-гомоглутаминовой кислоты;

L-β-гомоизолейцин; L-β-гомолейцин; L-β-гомометионин; L-β-гомофенилаланин; L-β-гомопролин; L-β-гомотриптофан; L-β-гомовалин; L-Nω-бензилоксикарбонил-β-гомолизин; Nω-L-β-гомоаргинин; O-бензил-L-β-гомогидроксипролин; O-бензил-L-β-гомосерин; O-бензил-L-β-гомотреонин; O-бензил-L-β-гомотириозин; γ-третил-L-β-гомоаспарагин; (R)-β-фенилаланин; γ-трет-бутиловый эфир L-β-гомоаспарагиновой кислоты; δ-трет-бутиловый эфир L-β-гомоглутаминовой кислоты; L-Nω-β-гомолизин; Nδ-третил-L-β-гомоглутамин; Nω-2,2,4,6,7-пентаметил-дигидробензофуран-5-сульфонил-L-β-гомоаргинин; O-т-бутил-L-β-гомогидроксипролин; O-т-бутил-L-β-гомосерин; O-т-бутил-L-β-гомотреонин; O-т-бутил-L-β-гомотириозин; 2-аминоциклопентанкарбоновую кислоту и 2-аминоциклогексанкарбоновую кислоту.

Аналоги аминокислот могут включать аналоги аланина, валина, глицина или лейцина. Примеры аминокислотных аналогов аланина, валина, глицина и лейцина включают, но не ограничены ими, следующие: α-метоксиглицин; α-аллил-L-аланин; α-аминоизомасляную кислоту; α-метиллейцин; β-(1-нафтил)-D-аланин; β-(1-нафтил)-L-аланин; β-(2-нафтил)-D-аланин; β-(2-нафтил)-L-аланин; β-(2-пиридил)-D-аланин; β-(2-пиридил)-L-аланин; β-(2-тиенил)-D-аланин; β-(2-тиенил)-L-аланин; β-(3-бензотиенил)-D-аланин; β-(3-бензотиенил)-L-аланин; β-(3-пиридил)-D-аланин; β-(3-пиридил)-L-аланин; β-(4-пиридил)-D-аланин; β-(4-пиридил)-L-аланин; β-хлор-L-аланин; β-циано-L-аланин; β-циклогексил-D-аланин, β-циклогексил-L-аланин, β-циклопентен-1-илаланин, β-циклопентилаланин, β-циклопропил-L-Ala-OH.дициклогексиламмониевая соль, β-трет-бутил-D-аланин; β-трет-бутил-L-аланин; γ-аминомасляную кислоту; L-α,β-диаминопропионовую кислоту; 2,4-динитрофенилглицин; 2,5-дигидро-D-фенилглицин; 2-амино-4,4,4-трифтормасляную кислоту; 2-фторфенилглицин; 3-амино-4,4,4-трифтормасляную кислоту; 3-фторвалин; 4,4,4-трифторвалин; 4,5-дегидро-L-leu-OH.дициклогексиламмониевую соль, 4-фтор-D-фенилглицин, 4-фтор-L-фенилглицин, 4-гидрокси-D-фенилглицин, 5,5,5-трифторлейцин, 6-аминогексановую кислоту, циклопентил-D-Gly-OH.дициклогексиламмониевую соль; циклопентил-Gly-OH.дициклогексиламмониевую соль; D-α,β-диаминопропионовую кислоту; D-α-аминомасляную кислоту; D-α-трет-бутилглицин; D-(2-тиенил)глицин; D-(3-тиенил)глицин; D-2-аминокапроновую кислоту; D-2-инданилглицин; D-аллилглицин-дициклогексиламмониевую соль; D-циклогексилглицин; D-норвалин; D-фенилглицин; β-аминомасляную кислоту; β-аминоизомасляную кислоту; (2-бромфенил)глицин; (2-метоксифенил)глицин; (2-метилфенил)глицин; (2-тиазол)глицин; (2-тиенил)глицин; 2-амино-3-(диметиламино)пропионовую кислоту; L-α,β-диаминопропионовую кислоту; L-α-аминомасляную кислоту; L-α-трет-бутилглицин; L-(3-тиенил)глицин; L-2-амино-3-(диметиламино)пропионовую кислоту; дициклогексиламмониевую соль L-2-аминокапроновой кислоты; L-2-инданилглицин; дициклогексиламмониевую соль L-аллилглицина; L-циклогексилглицин; L-фенилглицин; L-пропаргилглицин; L-норвалин; N-α-аминометил-L-аланин; D-α,γ-диаминомасляную кислоту; L-α,γ-диаминомасляную кислоту; β-циклопропил-L-аланин; (N-β-(2,4-динитрофенил))-L-α,β-диаминопропионовую

кислоту; (N-β-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-D-α,β-диаминопропионовую кислоту; (N-β-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-L-α,β-диаминопропионовую кислоту; (N-β-4-метилтритил)-L-α,β-диаминопропионовую кислоту; (N-β-аллилоксикарбонил)-L-α,β-диаминопропионовую кислоту; (N-γ-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-D-α,γ-диаминомасляную кислоту; (N-γ-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-L-α,γ-диаминомасляную кислоту; (N-γ-4-метилтритил)-D-α,γ-диаминомасляную кислоту; (N-γ-4-метилтритил)-L-α,γ-диаминомасляную кислоту; (N-γ-аллилоксикарбонил)-L-α,γ-диаминомасляную кислоту; D-α,γ-диаминомасляную кислоту; 4,5-дегидро-L-лейцин; циклопентил-D-Gly-OH; циклопентил-Gly-OH; D-аллилглицин, D-гомоциклогексилаланин, L-1-пиренилаланин, L-2-аминокапроновую кислоту, L-аллилглицин, L-гомоциклогексилаланин и N-(2-гидрокси-4-метокси-Bzl)-Gly-OH.

Аналоги аминокислот могут включать аналоги аргинина или лизина. Примеры аналогов аминокислот аргинина и лизина включают, но не ограничены ими, следующие: цитруллин; L-2-амино-3-гуанидинопропионовую кислоту; L-2-амино-3-уреидопропионовую кислоту; L-цитруллин; Lys(Me)₂-OH; Lys(N₃)-OH; Nδ-бензилоксикарбонил-L-орнитин; Nω-нитро-D-аргинин; Nω-нитро-L-аргинин; α-метилорнитин; 2,6-диаминогептандионовую кислоту; L-орнитин; (Nδ-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-D-орнитин; (Nδ-1-(4,4-диметил-2,6-диоксоциклогекс-1-илиден)этил)-L-орнитин; (Nδ-4-метилтритил)-D-орнитин; (Nδ-4-метилтритил)-L-орнитин; D-орнитин; L-орнитин; Arg(Me)(Pbf)-OH; Arg(Me)₂-OH (асимметричный); Arg(Me)₂-OH (симметричный); Lys(ivDde)-OH; Lys(Me)₂-OH.HCl, Lys(Me₃)-OH хлорид, Nω-нитро-D-аргинин и Nω-нитро-L-аргинин.

Аналоги аминокислот могут включать аналоги аспарагиновой или глутаминовой кислот. Примеры аминокислотных аналогов аспарагиновой и глутаминовой кислот включают, но не ограничены ими, следующие: α-метил-D-аспарагиновую кислоту; α-метилглутаминовую кислоту; α-метил-L-аспарагиновую кислоту; γ-метиленглутаминовую кислоту; (N-γ-этил)-L-глутамин; [N-α-(4-аминобензоил)]-L-глутаминовую кислоту; 2,6-диаминопимелиновую кислоту; L-α-аминомасляную кислоту; D-2-аминоадипиновую кислоту; D-α-аминосубериновую кислоту; α-аминопимелиновую кислоту; иминодиуксусную кислоту; L-2-аминоадипиновую кислоту; трео-β-метиласпарагиновую кислоту; γ,γ-ди-трет-бутиловый эфир γ-карбокси-D-глутаминовой кислоты; γ,γ-ди-трет-бутиловый эфир γ-карбокси-L-глутаминовой кислоты; Glu(OAll)-OH; L-Asu(OtBu)-OH; и пироглутаминовую кислоту.

Аналоги аминокислот могут включать аналоги цистеина и метионина. Примеры аналогов аминокислот цистеина и метионина включают, но не ограничены ими, Цис(фарнезил)-OH, Цис(фарнезил)-OMe, α-метилметионин, Cys(2-гидроксиэтил)-OH, Cys(3-аминопропил)-OH, 2-амино-4-(этилтио)масляную кислоту, бутионин, бутионинсульфоксимин, этионин, хлорид метионин метилсульфония, селенометионин, цистеиновую кислоту, [2-(4-пиридил)этил]-DL-пеницилламин, [2-(4-пиридил)этил]-L-

цистеин, 4-метоксибензил-D-пеницилламин, 4-метоксибензил-L-пеницилламин, 4-метилбензил-D-пеницилламин, 4-метилбензил-L-пеницилламин, бензил-D-цистеин, бензил-L-цистеин, бензил-DL-гомоцистеин, карбамоил-L-цистеин, карбоксиэтил-L-цистеин, карбоксиметил-L-цистеин, дифенилметил-L-цистеин, этил-L-цистеин, метил-L-цистеин, трет-бутил-D-цистеин, тритил-L-гомоцистеин, тритил-D-пеницилламин, цистатионин, гомоцистеин, L-гомоцистеин, (2-аминоэтил)-L-цистеин, селено-L-цистин, цистатионин, Cys(StBu)-ОН и ацетамидометил-D-пеницилламин.

Аналоги аминокислот могут включать аналоги фенилаланина и тирозина. Примеры аналогов аминокислот фенилаланина и тирозина включают β -метилфенилаланин, β -гидроксифенилаланин, α -метил-3-метокси-DL-фенилаланин, α -метил-D-фенилаланин, α -метил-L-фенилаланин, 1,2,3,4-тетрагидроизохинолин-3-карбоновую кислоту, 2,4-дихлорфенилаланин, 2-(трифторметил)-D-фенилаланин, 2-(трифторметил)-L-фенилаланин, 2-бром-D-фенилаланин, 2-бром-L-фенилаланин, 2-хлор-D-фенилаланин, 2-хлор-L-фенилаланин, 2-циано-D-фенилаланин, 2-циано-L-фенилаланин, 2-фтор-D-фенилаланин, 2-фтор-L-фенилаланин, 2-метил-D-фенилаланин, 2-метил-L-фенилаланин, 2-нитро-D-фенилаланин, 2-нитро-L-фенилаланин, 2,4,5-тригидроксифенилаланин, 3,4,5-трифтор-D-фенилаланин, 3,4,5-трифтор-L-фенилаланин, 3,4-дихлор-D-фенилаланин, 3,4-дихлор-L-фенилаланин, 3,4-дифтор-D-фенилаланин, 3,4-дифтор-L-фенилаланин, 3,4-дигидрокси-L-фенилаланин, 3,4-диметокси-L-фенилаланин, 3,5,3'-трийод-L-тиронин, 3,5-дийод-D-тирозин, 3,5-дийод-L-тирозин, 3,5-дийод-L-тиронин, 3-(трифторметил)-D-фенилаланин, 3-(трифторметил)-L-фенилаланин, 3-амино-L-тирозин, 3-бром-D-фенилаланин, 3-бром-L-фенилаланин, 3-хлор-D-фенилаланин, 3-хлор-L-фенилаланин, 3-хлор-L-тирозин, 3-циано-D-фенилаланин, 3-циано-L-фенилаланин, 3-фтор-D-фенилаланин, 3-фтор-L-фенилаланин, 3-фтортирозин, 3-йод-D-фенилаланин, 3-йод-L-фенилаланин, 3-йод-L-тирозин, 3-метокси-L-тирозин, 3-метил-D-фенилаланин, 3-метил-L-фенилаланин, 3-нитро-D-фенилаланин, 3-нитро-L-фенилаланин, 3-нитро-L-тирозин, 4-(трифторметил)-D-фенилаланин, 4-(трифторметил)-L-фенилаланин, 4-амино-D-фенилаланин, 4-амино-L-фенилаланин, 4-бензоил-D-фенилаланин, 4-бензоил-L-фенилаланин, 4-бис(2-хлорэтил)амино-L-фенилаланин, 4-бром-D-фенилаланин, 4-бром-L-фенилаланин, 4-хлор-D-фенилаланин, 4-хлор-L-фенилаланин, 4-циано-D-фенилаланин, 4-циано-L-фенилаланин, 4-фтор-D-фенилаланин, 4-фтор-L-фенилаланин, 4-йод-D-фенилаланин, 4-йод-L-фенилаланин, гомофенилаланин, тироксин, 3,3-дифенилаланин, тиронин, этилтирозин и метилтирозин.

Аналоги аминокислот могут включать аналоги пролина. Примеры аналогов аминокислот пролина включают, но не ограничены ими, 3,4-дегидропролин, 4-фторпролин, цис-4-гидроксипролин, тиазолидин-2-карбоновую кислоту и транс-4-фторпролин.

Аналоги аминокислот могут включать аналоги серина и треонина. Примеры аналогов аминокислот серина и треонина включают, но не ограничены ими, 3-амино-2-гидрокси-5-метилгексановую кислоту, 2-амино-3-гидрокси-4-метилпентановую кислоту, 2-амино-3-этоксипентановую кислоту, 2-амино-3-метоксибутановую кислоту, 4-амино-3-гидрокси-6-метилгептановую кислоту, 2-амино-3-бензилоксипропионовую кислоту, 2-амино-3-

бензилоксипропионовую кислоту, 2-амино-3-этоксипропионовую кислоту, 4-амино-3-гидроксипропановую кислоту и α -метилсерин.

Аналоги аминокислот могут включать аналоги триптофана. Примеры аналогов аминокислот триптофана включают, но не ограничены ими, следующие: α -метилтриптофан; β -(3-бензотиенил)-D-аланин; β -(3-бензотиенил)-L-аланин, 1-метилтриптофан, 4-метилтриптофан, 5-бензилокситриптофан, 5-бромтриптофан, 5-хлортриптофан, 5-фтортриптофан, 5-гидрокситриптофан, 5-гидрокси-L-триптофан, 5-метокситриптофан, 5-метокси-L-триптофан, 5-метилтриптофан, 6-бромтриптофан, 6-хлор-D-триптофан, 6-хлортриптофан, 6-фтортриптофан; 6-метилтриптофан; 7-бензилокситриптофан; 7-бромтриптофан; 7-метилтриптофан; D-1,2,3,4-тетрагидроноргарман-3-карбоновую кислоту; 6-метокси-1,2,3,4-тетрагидроноргарман-1-карбоновая кислота, 7-азатриптофан, L-1,2,3,4-тетрагидроноргарман-3-карбоновую кислоту, 5-метокси-2-метилтриптофан и 6-хлор-L-триптофан.

Аналоги аминокислот могут быть рацемическими. В некоторых случаях, используется D-изомер аналога аминокислоты. В некоторых случаях, используется L-изомер аналога аминокислоты. В некоторых случаях, аналог аминокислоты содержит хиральные центры, имеющие конфигурацию R или S. Иногда аминокислотная группа(ы) аналога β -аминокислоты замещена защитной группой, например, трет-бутилоксикарбонил (группа BOC), 9-флуоренилметилоксикарбонил (FMOC), тозил и подобными. Иногда функциональную группу карбоновой кислоты аналога β -аминокислоты защищают, например, в виде ее сложного эфира. В некоторых случаях, используют соль аналога аминокислоты.

В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой аминокислоту не природного происхождения, описанную в Liu C.C., Schultz, P.G. *Annu. Rev. Biochem.* 2010, 79, 413. В некоторых вариантах осуществления, аминокислота не природного происхождения содержит N⁶((2-азидоэтокси)карбонил)-L-лизин.

В некоторых вариантах осуществления, аминокислотный остаток, описанный в настоящем документе (например, в белке), мутирован до аминокислоты не природного происхождения до связывания с конъюгирующей группой. В некоторых случаях, мутация до аминокислоты не природного происхождения предотвращает или сводит к минимуму аутоантигенный ответ иммунной системы. Используемый в настоящем документе термин «аминокислота не природного происхождения» относится к аминокислоте, отличной от 20 аминокислот, которые встречаются в белке в природе. Неограничивающие примеры аминокислот не природного происхождения включают: п-ацетил-L-фенилаланин, п-йод-L-фенилаланин, п-метоксифенилаланин, O-метил-L-тирозин, п-пропаргилоксифенилаланин, п-пропаргилфенилаланин, L-3-(2-нафтил)аланин, 3-метилфенилаланин, O-4-аллил-L-тирозин, 4-пропил-L-тирозин, три-O-ацетил-GlcNAc-серин, L-Дора, фторированный фенилаланин, изопропил-L-фенилаланин, п-азидо-L-фенилаланин, п-ацил-L-фенилаланин, п-бензоил-L-фенилаланин, п-боронофенилаланин, O-пропаргилтирозин, L-фосфосерин,

фосфоносерин, фосфонотирозин, п-бромфенилаланин, селеноцистеин, п-амино-L-фенилаланин, изопропил-L-фенилаланин, N6-((азидозтокси)-карбонил)-L-лизин, AzK), N6-(((2-азидобензил)окси)карбонил)-L-лизин, N6-(((3-азидобензил)окси)карбонил)-L-лизин или N6-(((4-азидобензил)окси)карбонил)-L-лизин, аналог аминокислоты тирозина не природного происхождения; аналог глутаминовой аминокислоты не природного происхождения; аналог фенилаланиновой аминокислоты не природного происхождения; аналог сериновой аминокислоты не природного происхождения; аналог треониновой аминокислоты не природного происхождения; алкил, арил, ацил, азидо, циано, гало, гидразин, гидразид, гидроксил, алкенил, алкинил, простой эфир, тиол, сульфонил, селено, сложный эфир, тиокислоту, борат, боронат, фосфо, фосфоно, фосфин, гетероцикл, енон, имин, альдегид, гидроксилламин, кето или аминозамещенную аминокислоту или их комбинацию; аминокислоту с фотоактивируемым сшивающим агентом, спин-меченую аминокислоту; флуоресцентную аминокислоту; металл-связывающую аминокислоту; металлсодержащую аминокислоту; радиоактивную аминокислоту; аминокислоту с фотолabileйной временной защитой и/или фотоизомеризуемую аминокислоту; аминокислоту, содержащую биотин или аналог биотина; кетосодержащую аминокислоту; аминокислоту, содержащую полиэтиленгликоль или простой полиэфир; аминокислоту, замещенную тяжелым атомом; химически расщепляемую или фоторасщепляемую аминокислоту; аминокислоту с удлиненной боковой цепью; аминокислоту, содержащую токсическую группу; замещенную сахаром аминокислоту; углерод-связанную аминокислоту, содержащую сахар; редокс-активную аминокислоту; α-гидроксисодержащую кислоту; аминотиокислоту; α,α-дизамещенную аминокислоту; β-аминокислоту; циклическую аминокислоту, отличную от пролина или гистидина, и ароматическую аминокислоту, отличную от фенилаланина, тирозина или триптофана.

В некоторых вариантах осуществления, аминокислота не природного происхождения содержит селективную реакционноспособную группу или реакционноспособную группу для сайт-селективного мечения белка-мишени или полипептида. В некоторых случаях, химия представляет собой биоортогональную реакцию (например, биосовместимые и селективные реакции). В некоторых случаях, химия представляет собой реакцию образования алкин-азидтриазола, катализируемую Cu (I) или «без меди», лигирование Штаудингера, реакцию Дильса-Альдера с обратными электронными требованиями (IEDDA), химию «фото-клик», или процесс, опосредованный металлами, такой как метатезис олефинов и перекрестное сочетание Сузуки-Мияура или Соногашира. В некоторых вариантах осуществления, аминокислота не природного происхождения содержит фотореактивную группу, которая поперечно сшивается при облучении, например, УФ. В некоторых вариантах осуществления, аминокислота не природного происхождения включает аминокислоту с фотолabileйной временной защитой. В некоторых случаях, аминокислота не природного происхождения представляет собой *пара*-замещенное, *мета*-замещенное или *орто*-замещенное производное аминокислоты.

В некоторых случаях, аминокислота не природного происхождения включает п-

ацетил-L-фенилаланин, п-азидометил-L-фенилаланин (pAMF), п-йодо-L-фенилаланин, О-метил-L-тирозин, п-метоксифенилаланин, п-пропаргилоксифенилаланин, п-пропаргилфенилаланин, L-3-(2-нафтил)аланин, 3-метилфенилаланин, О-4-аллил-L-тирозин, 4-пропил-L-тирозин, три-О-ацетил-GlcNAc α -серин, L-Dopa, фторированный фенилаланин, изопропил-L-фенилаланин, п-азидо-L-фенилаланин, п-ацил-L-фенилаланин, п-бензоил-L-фенилаланин, L-фосфосерин, фосфоносерин, фосфонотирозин, п-бромфенилаланин, п-амино-L-фенилаланин или изопропил-L-фенилаланин.

В некоторых случаях, аминокислота не природного происхождения представляет собой 3-аминотирозин, 3-нитротирозин, 3,4-дигидроксифенилаланин или 3-йодтирозин. В некоторых случаях, аминокислота не природного происхождения представляет собой фенилселеноцистеин. В некоторых случаях, аминокислота не природного происхождения представляет собой производное фенилаланина, содержащее бензофенон, кетон, йодид, метокси, ацетил, бензоил или азид. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, содержащее бензофенон, кетон, йодид, метокси, ацетил, бензоил или азид. В некоторых случаях, аминокислота не природного происхождения содержит ароматическую боковую цепь. В некоторых случаях, аминокислота не природного происхождения не содержит ароматическую боковую цепь. В некоторых случаях, аминокислота не природного происхождения содержит азидогруппу. В некоторых случаях, аминокислота не природного происхождения содержит акцепторную группу Михаэля. В некоторых случаях, акцепторные группы Михаэля содержат ненасыщенную группу, способную образовывать ковалентную связь через реакцию 1,2-присоединения. В некоторых случаях, акцепторные группы Михаэля содержат электрон-дефицитные алкины или алкины. В некоторых случаях, акцепторные группы Михаэля включают, но не ограничены ими, альфа-, бета-ненасыщенные: кетоны, альдегиды, сульфоксиды, сульфоны, нитрилы, имины или ароматические соединения. В некоторых случаях, аминокислота не природного происхождения представляет собой дегидроаланин. В некоторых случаях, аминокислота не природного происхождения содержит альдегидную или кетоновую группу. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, содержащее альдегидную или кетоновую группу. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, содержащее одну или несколько атомов O, N, Se или S в бета-, гамма- или дельта-положении. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, содержащее атомы O, N, Se или S в гамма-положении. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, в котором атом эпсилон N заменен атомом кислорода. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, которое не является посттрансляционно модифицированным не природного происхождения.

В некоторых случаях, аминокислота не природного происхождения представляет собой аминокислоту, содержащую боковую цепь, в которой шестой атом от альфа-

положения содержит карбонильную группу. В некоторых случаях, аминокислота не природного происхождения представляет собой аминокислоту, содержащую боковую цепь, в которой шестой атом от альфа-положения содержит карбонильную группу, и пятый атом от альфа-положения представляет собой азот. В некоторых случаях, аминокислота не природного происхождения представляет собой аминокислоту, содержащую боковую цепь, в которой седьмой атом от альфа-положения представляет собой атом кислорода.

В некоторых случаях, аминокислота не природного происхождения представляет собой производное серина, содержащее селен. В некоторых случаях, аминокислота не природного происхождения представляет собой селеносерин (2-амино-3-гидроселенопропановую кислоту). В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-3-((2-((3-(бензилокси)-3-оксипропил)амино)этил)селанил)пропановую кислоту. В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-3-(фенилселанил)пропановую кислоту. В некоторых случаях, аминокислота не природного происхождения содержит селен, где окисление селена приводит к образованию аминокислоты не природного происхождения, содержащей алкен.

В некоторых случаях, аминокислота не природного происхождения содержит циклооктинильную группу. В некоторых случаях, аминокислота не природного происхождения содержит трансциклоктенильную группу. В некоторых случаях, аминокислота не природного происхождения содержит норборненильную группу. В некоторых случаях, аминокислота не природного происхождения содержит циклопропенил. В некоторых случаях, аминокислота не природного происхождения содержит диазириновую группу. В некоторых случаях, аминокислота не природного происхождения содержит тетразиновую группу.

В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, где азот боковой цепи карбамилирован. В некоторых случаях, аминокислота не природного происхождения представляет собой производное лизина, в котором азот боковой цепи ацилирован. В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-6-[[[(трет-бутокси)карбонил]амино}гексановую кислоту. В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-6-[[[(трет-бутокси)карбонил]амино}гексановую кислоту. В некоторых случаях, аминокислота не природного происхождения представляет собой N⁶-Вос-N⁶-метиллизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N⁶-ацетиллизин. В некоторых случаях, аминокислота не природного происхождения представляет собой пирролизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N⁶-трифторацетиллизин. В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-6-[[[(бензилокси)карбонил]амино}гексановую кислоту. В некоторых случаях, аминокислота не природного происхождения представляет собой 2-амино-6-[[[(п-

[illegible]

В некоторых случаях, аминокислота не природного происхождения представляет собой N6-пропионилизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-бутириллизин. В некоторых случаях, аминокислота

не природного происхождения представляет собой N6-(бут-2-еноил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((бицикло[2.2.1]гепт-5-ен-2-илокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((спиро[2.3]гекс-1-ен-5-илметокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-(((4-(1-(трифторметил)циклопроп-2-ен-1-ил)бензил)окси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((бицикло[2.2.1]гепт-5-ен-2-илметокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой цистеиниллизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((1-(6-нитробензо[d][1,3]диоксол-5-ил)этокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((2-(3-метил-3Н-диазиридин-3-ил)этокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((3-(3-метил-3Н-диазиридин-3-ил)пропокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((метанитробенилокси)N6-метилкарбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((бицикло[6.1.0]нон-4-ин-9-илметокси)карбонил)лизин. В некоторых случаях, аминокислота не природного происхождения представляет собой N6-((циклогепт-3-ен-1-илокси)карбонил)-L-лизин.

В некоторых вариантах осуществления, аминокислота не природного происхождения включена в белок с помощью кодона не природного происхождения, содержащего нуклеотид не природного происхождения.

В некоторых случаях, включение не встречающейся в природе аминокислоты в белок опосредуется ортогональной модифицированной парой синтетаза/тРНК. Такие ортогональные пары содержат природную или мутированную синтетазы, которые способны заряжать не встречающуюся в природе тРНК специфической аминокислотой не природного происхождения, часто при этом минимизируя зарядку а) других эндогенных аминокислот или альтернативных аминокислот не природного происхождения в тРНК не природного происхождения и б) любые другие (включая эндогенные) тРНК. Такие ортогональные пары включают тРНК, которые способны заряжаться синтетазой, при этом избегая зарядки другими эндогенными аминокислотами через эндогенные синтетазы. В некоторых вариантах осуществления, такие пары идентифицируют из различных организмов, таких как бактерии, дрожжи, археи или человеческие источники. В некоторых вариантах осуществления, пара ортогональная синтетаза/тРНК содержит компоненты из одного организма. В некоторых вариантах осуществления, пара ортогональная синтетаза/тРНК содержит компоненты из двух разных организмов. В некоторых вариантах осуществления, пара ортогональная синтетаза/тРНК содержит компоненты, которые перед модификацией способствуют трансляции различных аминокислот. В некоторых вариантах

[illegible]

тРНК глутамин. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК глутаминовой кислоты. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированный аланин глицин. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК гистидина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК лейцина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК изолейцина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК лизина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК метионина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК фенилаланина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК пролина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК серина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК треонина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК триптофана. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК тирозина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК валина. В некоторых вариантах осуществления, ортогональная тРНК представляет собой модифицированную тРНК фосфосерина.

В некоторых вариантах осуществления, аминокислота не природного происхождения включена в белок с помощью пары аминоацил (aaRS или RS)-тРНК синтетаза-тРНК. Примеры пар aaRS-тРНК включают, но не ограничены ими, пары *Methanococcus jannaschii* (Mj-Tyr) aaRS/тРНК, пары *E. coli* TyrRS (*Ec-Tyr*)/*B. stearothermophilus* tRNA_{CUA}, пары *E. coli* LeuRS (*Ec-Leu*)/*B. stearothermophilus* tRNA_{CUA} и пары пирролизил-тРНК. В некоторых случаях, кислота не природного происхождения включается в белок с помощью пары *Mj-TyrRS*/тРНК. Примеры аминокислот не природного происхождения (UAA), которые могут быть включены с помощью пары *Mj-TyrRS*/тРНК, включают, но не ограничены ими, пара-замещенные производные фенилаланина, такие как как *n*-аминофенилаланин и *n*-метойфенилаланин; мета-замещенные производные тирозина, такие как 3-аминотирозин, 3-нитротирозин, 3,4-дигидроксифенилаланин и 3-йодтирозин; фенилселенцистеин; *n*-боронофенилаланин; и *o*-нитробензилтирозин.

В некоторых вариантах осуществления, аминокислота не природного происхождения включена в белок с помощью пары *Ec-Tyr*/tRNA_{CUA} или *Ec-Leu*/tRNA_{CUA}. Примеры аминокислот не природного происхождения (UAA), которые могут быть включены с помощью пары *Ec-Tyr*/tRNA_{CUA} или *Ec-Leu*/tRNA_{CUA}, включают, но не ограничены ими, производные фенилаланина, содержащие бензофеноновые, кетоновые,

йодидные или азидные заместители; *O*-пропаргилтирозин; α -аминокаприловую кислоту, *O*-метилтирозин, *O*-нитробензилцистеин; и 3-(нафталин-2-иламино)-2-аминопропановую кислоту.

В некоторых вариантах осуществления, аминокислота не природного происхождения включена в белок с помощью пары пирролизил-тРНК. В некоторых случаях, PylRS получают из архебактериальных видов, например, из метаногенных архебактерий. В некоторых случаях, PylRS получают из *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*. Примеры UAA, которые могут быть включены в пару пирролизил-тРНК, включают, но не ограничены ими, амид- и карбамат-замещенные лизины, такие как 2-амино-6-((R)-тетрагидрофуран-2-карбоксамидо)гексановая кислота, *N*- ϵ -D-пролил-L-лизин и *N*- ϵ -циклопентилокскарбонил-L-лизин, *N*- ϵ -акрилоил-L-лизин, *N*- ϵ -[(1-(6-нитробензо[d][1,3]диоксол-5-ил)этокси)карбонил]-L-лизин и *N*- ϵ -(1-метилциклопро-2-енкарбоксамидо)лизин.

В некоторых случаях, аминокислота не природного происхождения включена в белок, описанный в настоящем документе, синтетазой, описанной в US 9,988,619 и US 9,938,516. Примеры UAA, которые могут быть включены такими синтетазами, включают пара-метилазидо-L-фенилаланин, аралкил, гетероцикл, гетероаралкильные аминокислоты не природного происхождения и другие. В некоторых вариантах осуществления, такие UAA включают пиридил, пиразинил, пиразолил, триазолил, оксазолил, тиазолил, тиофенил или другой гетероцикл. Такие аминокислоты в некоторых примерах включают азиды, тетразины или другую химическую группу, способную конъюгироваться с партнером по связыванию, таким как водорастворимая группа. В некоторых вариантах осуществления, такие синтетазы экспрессируются и используются для включения UAA в белки *in vivo*. В некоторых вариантах осуществления, такие синтетазы используются для включения UAA в белки с использованием бесклеточной системы трансляции, такой как лизат клеток или восстановленная система очищенных компонентов. тРНК может быть заряжена аминокислотой не природного происхождения в бесклеточной системе или в отдельной реакции заранее (таким образом, чтобы заряженная тРНК добавлялась непосредственно в систему, содержащую рибосомы, иРНК и другие компоненты, без необходимости добавления в систему синтетазы или конструкции, кодирующей синтетазу).

Системы для трансляции *in vitro* описаны, например, в Zeenko et al., RNA 14:593-602 (2008); Spirin, Trends Biotechnol. 2004:538-545 (2004); и Endo et al., Curr. Opin. Biotechnol. 17:373-380 (2006). Системы могут быть приготовлены из клеточных лизатов (например, экстрактов) или восстановлены из очищенных компонентов. Системы могут содержать, помимо рибосом, тРНК и других компонентов, описанных в настоящем документе, один или несколько факторов инициации трансляции; АТФ; и один или несколько факторов терминации трансляции. В некоторых вариантах осуществления, система дополнительно содержит один или несколько молекулярных шаперонов, которые могут способствовать сворачиванию формирующегося полипептида во время и/или после трансляции.

В некоторых случаях, аминокислота не природного происхождения включена в белок, описанный в настоящем документе, с помощью природной синтетазы. В некоторых вариантах осуществления, аминокислота не природного происхождения включена в белок организмом, который является ауксотрофным для одной или нескольких аминокислот. В некоторых вариантах осуществления, синтетазы, соответствующие ауксотрофной аминокислоте, способны заряжать соответствующую тРНК аминокислотой не природного происхождения. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой селеноцистеин или его производное. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой селенометионин или его производное. В некоторых вариантах осуществления, аминокислота не природного происхождения представляет собой ароматическую аминокислоту, где ароматическая аминокислота содержит арилгалогенид, такой как йодид. В некоторых вариантах осуществления, аминокислота не природного происхождения структурно подобна ауксотрофной аминокислоте.

В некоторых случаях, аминокислота не природного происхождения содержит аминокислоту не природного происхождения, показанную на **фиг. 4А**.

В некоторых случаях, аминокислота не природного происхождения содержит производное или аналог лизина или фенилаланина. В некоторых случаях, аминокислота не природного происхождения содержит производное лизина или аналог лизина. В некоторых случаях, аминокислота не природного происхождения содержит пирролизин (Pyl). В некоторых случаях, аминокислота не природного происхождения содержит производное фенилаланина или аналог фенилаланина. В некоторых случаях, аминокислота не природного происхождения представляет собой аминокислоту не природного происхождения, описанную в Wan, et al., "Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool," *Biochem Biophys Acta* 1844(6): 1059-4070 (2014). В некоторых случаях, аминокислота не природного происхождения содержит аминокислоту не природного происхождения, показанную на **фиг.4В** и **фиг.4С**.

В некоторых вариантах осуществления, аминокислота не природного происхождения содержит аминокислоту не природного происхождения, показанную на **фиг. 4D - фиг. 4G** (взято из таблицы 1 из Dumas *et al.*, *Chemical Science* 2015, **6**, 50-69).

В некоторых вариантах осуществления, аминокислота не природного происхождения, включенная в белок, описанный в настоящем документе, описана в US 9,840,493; US 9,682,934; US 2017/0260137; US 9,938,516; или US 2018/0086734. Типовые UAA, которые могут быть включены такими синтетазами, включают пара-метилазидо-L-фенилаланин, аралкил, гетероциклил и гетероаралкил и производные лизина. В некоторых вариантах осуществления, такие UAA включают пиридил, пиразинил, пиразолил, триазолил, оксазолил, тиазолил, тиофенил или другой гетероцикл. Такие аминокислоты в некоторых вариантах осуществления, включают азиды, тетразины или другую химическую группу, способную конъюгироваться с партнером по связыванию, таким как водорастворимая группа. В некоторых вариантах осуществления, UAA содержит азид,

присоединенный к ароматической группе через алкильный линкер. В некоторых вариантах осуществления, алкильный линкер представляет собой C₁-C₁₀ линкер. В некоторых вариантах осуществления, UAA содержит тетразин, присоединенный к ароматической группе через алкильный линкер. В некоторых вариантах осуществления, UAA содержит тетразин, присоединенный к ароматической группе через аминогруппу. В некоторых вариантах осуществления, UAA содержит тетразин, присоединенный к ароматической группе через алкиламиногруппу. В некоторых вариантах осуществления, UAA содержит азид, присоединенный к концевому атому азота (например, N6 производного лизина, или N5, N4 или N3 производного, содержащего более короткую алкильную боковую цепь) боковой цепи аминокислоты через алкильную цепь. В некоторых вариантах осуществления, UAA содержит тетразин, присоединенный к концевому атому азота боковой цепи аминокислоты через алкильную цепь. В некоторых вариантах осуществления, UAA содержит азид или тетразин, присоединенный к амиду через алкильный линкер. В некоторых вариантах осуществления, UAA представляет собой азид- или тетразин-содержащий карбамат или амид 3-аминоаланина, серина, лизина или их производного. В некоторых вариантах осуществления, такие UAA включаются в белки *in vivo*. В некоторых вариантах осуществления, такие UAA включаются в белки в бесклеточной системе.

Типы Клеток

В некоторых вариантах осуществления, используются многие типы клеток/микроорганизмов, например, для трансформации или генной инженерии. В некоторых вариантах осуществления, клетка представляет собой эукариотическую клетку. В некоторых случаях, клетка представляет собой эукариотическую клетку, такую как культивируемое животное, растение или клетка человека. В дополнительных случаях, клетка присутствует в организме, таком как растение или животное.

В некоторых вариантах осуществления, сконструированный микроорганизм представляет собой одноклеточный организм, часто способный к делению и пролиферации. Микроорганизм может иметь один или несколько из следующих признаков: аэробный, анаэробный, нитчатый, не нитчатый, моноплоидный, диплоидный, ауксотрофный и/или не ауксотрофный. В некоторых вариантах осуществления, сконструированный микроорганизм представляет собой не прокариотический микроорганизм. В некоторых вариантах осуществления, сконструированный микроорганизм представляет собой эукариотический микроорганизм (*например*, дрожжи, грибки, амёбу). В некоторых вариантах осуществления, сконструированный микроорганизм представляет собой гриб. В некоторых вариантах осуществления, сконструированный организм представляет собой дрожжи.

Любые подходящие дрожжи могут быть выбраны в качестве микроорганизма-хозяина, сконструированного микроорганизма, генетически модифицированного организма или источника гетерологичного или модифицированного полинуклеотида. Дрожжи включают, но не ограничены ими, дрожжи *Yarrowia* (*например*, *Y. lipolytica* (ранее классифицированный как *Candida lipolytica*)), дрожжи *Candida* (*например*, *C. revkaufi*, *C.*

viswanathii, *C. pulcherrima*, *C. tropicalis*, *C. utilis*), дрожжи *Rhodotorula* (*например*, *R. glutinus*, *R. graminis*), дрожжи *Rhodosporidium* (*например*, *R. toruloides*), дрожжи *Saccharomyces* (*например*, *S. cerevisiae*, *S. bayanus*, *S. pastorianus*, *S. carlsbergensis*), дрожжи *Cryptococcus*, дрожжи *Trichosporon* (*например*, *T. pullans*, *T. cutaneum*), дрожжи *Pichia* (*например*, *P. pastoris*) и дрожжи *Lipomyces* (*например*, *L. starkeyii*, *L. lipoferus*). В некоторых вариантах осуществления, подходящие дрожжи относятся к роду *Arachniotus*, *Aspergillus*, *Aureobasidium*, *Auxarthron*, *Blastomyces*, *Candida*, *Chrysosporium*, *Chrysosporium*, *Debaryomyces*, *Coccidiodes*, *Cryptococcus*, *Gymnoascus*, *Hansenula*, *Histoplasma*, *Issatchenkia*, *Kluyveromyces*, *Lipomyces*, *Issatchenkia*, *Microsporum*, *Myxotrichum*, *Myxozyma*, *Oidiodendron*, *Pachysolen*, *Penicillium*, *Pichia*, *Rhodosporidium*, *Rhodotorula*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, *Scopulariopsis*, *Sepedonium*, *Trichosporon* или *Yarrowia*. В некоторых вариантах осуществления, подходящие дрожжи принадлежат к видам *Arachniotus flavoluteus*, *Aspergillus flavus*, *Aspergillus fumigatus*, *Aspergillus niger*, *Aureobasidium pullulans*, *Auxarthron thaxteri*, *Blastomyces dermatitidis*, *Candida albicans*, *Candida dubliniensis*, *Candida famata*, *Candida glabrata*, *Candida guilliermondii*, *Candida kefyr*, *Candida krusei*, *Candida lambica*, *Candida lipolytica*, *Candida lusitanae*, *Candida parapsilosis*, *Candida pulcherrima*, *Candida revkaufi*, *Candida rugosa*, *Candida tropicalis*, *Candida utilis*, *Candida viswanathii*, *Candida xestobii*, *Chrysosporium keratinophilum*, *Coccidiodes immitis*, *Cryptococcus albidus* var. *diffluens*, *Cryptococcus laurentii*, *Cryptococcus neoformans*, *Debaryomyces hansenii*, *Gymnoascus dugwayensis*, *Hansenula anomala*, *Histoplasma capsulatum*, *Issatchenkia occidentalis*, *Issatchenkia orientalis*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kluyveromyces thermotolerans*, *Kluyveromyces waltii*, *Lipomyces lipoferus*, *Lipomyces starkeyii*, *Microsporum gypseum*, *Myxotrichum deflexum*, *Oidiodendron echinulatum*, *Pachysolen tannophilis*, *Penicillium notatum*, *Pichia anomala*, *Pichia pastoris*, *Pichia stipitis*, *Rhodosporidium toruloides*, *Rhodotorula glutinus*, *Rhodotorula graminis*, *Saccharomyces cerevisiae*, *Saccharomyces kluyveri*, *Schizosaccharomyces pombe*, *Scopulariopsis acremonium*, *Sepedonium chrysospermum*, *Trichosporon cutaneum*, *Trichosporon pullans*, *Yarrowia lipolytica*, or *Yarrowia lipolytica* (ранее классифицирован как *Candida lipolytica*). В некоторых вариантах осуществления дрожжами является штамм *Y. lipolytica*, который включает, но не ограничен ими, штаммы ATCC20362, ATCC8862, ATCC18944, ATCC20228, ATCC76982 и LGAM S(7)1 (Papanikolaou S., and Aggelis G., *Bioresour. Technol.* 82(1):43-9 (2002)). В некоторых вариантах осуществления, дрожжи представляют собой дрожжи *Candida* (т.е. *Candida* spp.). Любые подходящие виды *Candida* могут быть использованы и/или генетически модифицированы для получения жирной дикарбоновой кислоты (*например*, октандиикислоты, декандиикислоты, додекандиикислоты, тетрадекандиикислоты, гексадекандиикислоты, октадекандиикислоты, эйкозандиикислоты). В некоторых вариантах осуществления, подходящие виды *Candida* включают, но не ограничены ими, *Candida albicans*, *Candida dubliniensis*, *Candida famata*, *Candida glabrata*, *Candida guilliermondii*, *Candida kefyr*, *Candida krusei*, *Candida lambica*, *Candida lipolytica*, *Candida lusitanae*, *Candida parapsilosis*, *Candida pulcherrima*, *Candida revkaufi*, *Candida rugosa*, *Candida tropicalis*, *Candida*

utilis, *Candida viswanathii*, *Candida xestobii* и любые другие дрожжи *Candida spp.*, описанные в настоящем документе. Неограничивающие примеры штаммов *Candida spp.* включают, но не ограничены ими, штаммы sAA001 (ATCC20336), sAA002 (ATCC20913), sAA003 (ATCC20962), sAA496 (US2012/0077252), sAA106 (US2012/0077252), SU-2 (*ura3-/ura3-*), H5343 (блокирован бета-окислением; патент США № 5,648,247). Любые подходящие штаммы дрожжей *Candida spp.* можно использовать в качестве родительских штаммов для генетической модификации.

Роды, виды и штаммы дрожжей часто настолько тесно связаны по генетическому содержанию, что их бывает трудно различить, классифицировать и/или назвать. В некоторых случаях, штаммы *C. lipolytica* и *Y. lipolytica* трудно различить, классифицировать и/или назвать, и в некоторых случаях их можно рассматривать как один и тот же организм. В некоторых случаях, различные штаммы *C. tropicalis* и *C. viswanathii* могут быть трудно различимы, классифицированы и/или названы (например, см. Arie et.al., J. Gen. Appl. Microbiol., 46, 257-262 (2000)). Некоторые штаммы *C. tropicalis* и *C. viswanathii*, полученные от ATCC, а также из других коммерческих или академических источников, можно считать эквивалентными и в равной степени подходящими для вариантов осуществления, описанных в настоящем документе. В некоторых вариантах осуществления, считается, что некоторые родительские штаммы *C. tropicalis* и *C. viswanathii* отличаются только названием.

Любой подходящий гриб может быть выбран в качестве микроорганизма-хозяина, сконструированного микроорганизма или источника гетерологичного полинуклеотида. Неограничивающие примеры грибов включают, но не ограничены ими, грибы *Aspergillus* (например, *A. parasiticus*, *A. nidulans*), грибы *Thraustochytrium*, грибы *Schizochytrium* и грибы *Rhizopus* (например, *R. arrhizus*, *R. oryzae*, *R. nigricans*). В некоторых вариантах осуществления, гриб представляет собой штамм *A. parasiticus*, который включает, но не ограничен им, штамм ATCC24690, и в некоторых примерах, гриб представляет собой штамм *A. nidulans*, который включает, но не ограничен им, штамм ATCC38163.

Клетки не микробных организмов можно использовать в качестве микроорганизма-хозяина, сконструированного микроорганизма или источника гетерологичного полинуклеотида. Примеры таких клеток включают, но не ограничены ими, клетки насекомых (например, *Drosophila* (например, *D. melanogaster*), *Spodoptera* (например, клетки Sf9 или Sf21 *S. frugiperda*) и *Trichoplusia* (например, клетки High-Five); клетки нематод (например, клетки *C. elegans*); клетки птиц; клетки земноводных (например, клетки *Xenopus laevis*); клетки рептилий; клетки млекопитающих (например, клетки NIH3T3, 293, CHO, COS, VERO, C127, BHK, Per-C6, клетки меланомы Боуэса и HeLa); и клетки растений (например, *Arabidopsis thaliana*, *Nicotiana tabacum*, *Cuphea acinifolia*, *Cuphea aequipetala*, *Cuphea angustifolia*, *Cuphea appendiculata*, *Cuphea avigera*, *Cuphea avigera* var. *pulcherrima*, *Cuphea axilliflora*, *Cuphea bahiensis*, *Cuphea baillonis*, *Cuphea brachypoda*, *Cuphea bustamanta*, *Cuphea calcarata*, *Cuphea calophylla*, *Cuphea calophylla* subsp. *mesostemon*, *Cuphea carthagenensis*, *Cuphea circaeoides*, *Cuphea confertiflora*, *Cuphea cordata*, *Cuphea crassiflora*,

Cuphea cyanea, *Cuphea decandra*, *Cuphea denticulata*, *Cuphea disperma*, *Cuphea epilobiifolia*, *Cuphea ericoides*, *Cuphea flava*, *Cuphea flavisetula*, *Cuphea fuchsiifolia*, *Cuphea gaumeri*, *Cuphea glutinosa*, *Cuphea heterophylla*, *Cuphea hookeriana*, *Cuphea hyssopifolia* (вереск Мексиканский), *Cuphea hyssopoides*, *Cuphea ignea*, *Cuphea ingrata*, *Cuphea jorullensis*, *Cuphea lanceolata*, *Cuphea linarioides*, *Cuphea llavea*, *Cuphea lophostoma*, *Cuphea lutea*, *Cuphea lutescens*, *Cuphea melanium*, *Cuphea melvilla*, *Cuphea micrantha*, *Cuphea micropetala*, *Cuphea mimuloides*, *Cuphea nitidula*, *Cuphea palustris*, *Cuphea parsonsia*, *Cuphea pascuorum*, *Cuphea paucipetala*, *Cuphea procumbens*, *Cuphea pseudosilene*, *Cuphea pseudovaccinium*, *Cuphea pulchra*, *Cuphea racemosa*, *Cuphea repens*, *Cuphea salicifolia*, *Cuphea salvadorensis*, *Cuphea schumannii*, *Cuphea sessiliflora*, *Cuphea sessilifolia*, *Cuphea setosa*, *Cuphea spectabilis*, *Cuphea spermacoce*, *Cuphea splendida*, *Cuphea splendida* var. *viridiflava*, *Cuphea strigulosa*, *Cuphea subuligera*, *Cuphea teleandra*, *Cuphea thymoides*, *Cuphea tolucana*, *Cuphea urens*, *Cuphea utriculosa*, *Cuphea viscosissima*, *Cuphea watsoniana*, *Cuphea wrightii*, *Cuphea lanceolata*).

Микроорганизмы или клетки, используемые в качестве организмов-хозяев или источника гетерологичного полинуклеотида, являются коммерчески доступными. Описанные в настоящем документе микроорганизмы и клетки, и другие подходящие микроорганизмы и клетки доступны, например, от Invitrogen Corporation, (Carlsbad, CA), American Type Culture Collection (Manassas, Virginia) и Agricultural Research Culture Collection (NRRL; Peoria, Illinois). Микроорганизмы-хозяева и сконструированные микроорганизмы могут быть предоставлены в любой подходящей форме. Например, такие микроорганизмы могут быть предоставлены в виде жидкой культуры или твердой культуры (*например*, среды на основе агара), которая может представлять собой первичную культуру или может быть пассирована (*например*, разведена и культивирована) один или несколько раз. Микроорганизмы также могут быть предоставлены в замороженной или сухой форме (*например*, лиофилизированными). Микроорганизмы могут быть предоставлены в любой подходящей концентрации.

Реагенты и Инструменты Для Нуклеиновых Кислот

Нуклеотид и/или реагент нуклеиновой кислоты (или полинуклеотид) для применения в способе, клетке или сконструированном микроорганизме, описанных в настоящем документе, содержит одну или несколько ORF с или без нуклеотида не природного происхождения. ORF может быть из любого подходящего источника, иногда из геномной ДНК, иРНК, обратно транскрибируемой РНК или комплементарной ДНК (кДНК) или библиотеки нуклеиновых кислот, содержащей одну или несколько из вышеперечисленных, и из любого вида организма, который содержит представляющую интерес последовательность нуклеиновой кислоты, представляющий интерес белок или представляющую интерес активность. Неограничивающие примеры организмов, из которых можно получить ORF, включают, например, бактерии, дрожжи, грибы, человека, насекомых, нематод, крупный рогатый скот, лошадей, собак, кошек, крыс или мышей. В некоторых вариантах осуществления, нуклеотид и/или реагент нуклеиновой кислоты или другой реагент, описанный в настоящем документе, выделяют или очищают. Могут быть

созданы ORF, которые включают нуклеотиды не природного происхождения, с помощью опубликованных способов *in vitro*. В некоторых случаях, нуклеотидный или нуклеиновый реагент содержит нуклеиновое основание не природного происхождения.

Реагент нуклеиновой кислоты иногда содержит нуклеотидную последовательность, примыкающую к ORF, которая транслируется вместе с ORF и кодирует аминокислотную метку. Последовательность нуклеотидов, кодирующая метку, расположена на 3' и/или 5' ORF реагента нуклеиновой кислоты, тем самым кодируя метку на С-конце или N-конце белка или пептида, кодируемого ORF. Можно использовать любую метку, которая не аннулирует транскрипцию и/или трансляцию *in vitro*, и она может быть надлежащим образом выбрана специалистом в данной области техники. Метки могут облегчить выделение и/или очистку желаемого продукта ORF из культуральной или ферментационной среды. В некоторых случаях, библиотеки реагентов нуклеиновых кислот используют с описанными в настоящем документе способами и композициями. Например, библиотека, по меньшей мере, 100, 1000, 2000, 5000, 10000 или более 50000 уникальных полинуклеотидов присутствуют в библиотеке, где каждый полинуклеотид содержит, по меньшей мере, одно нуклеотидное основание не природного происхождения.

Нуклеиновая кислота или реагент нуклеиновой кислоты, с или без нуклеотидом не природного происхождения, могут содержать определенные элементы, *например*, регуляторные элементы, часто выбранные в соответствии с предполагаемым использованием нуклеиновой кислоты. Любой из следующих элементов может быть включен или исключен из реагента нуклеиновой кислоты. Реагент нуклеиновой кислоты, *например*, может включать один или несколько или все из следующих нуклеотидных элементов: один или несколько промоторных элементов, одну или несколько 5' нетранслируемых областей (5'UTR), одну или несколько областей в которую может быть вставлена последовательность нуклеотида-мишени («элемент вставки»), одна или несколько последовательностей нуклеотидов-мишеней, одна или несколько 3' нетранслируемых областей (3'UTR) и один или несколько элементов селекции. Реагент нуклеиновой кислоты может быть представлен с одним или несколькими такими элементами, и другие элементы могут быть вставлены в нуклеиновую кислоту до того, как нуклеиновая кислота будет введена в желаемый организм. В некоторых вариантах осуществления, представленный реагент нуклеиновой кислоты содержит промотор, 5'UTR, необязательно, 3'UTR и элемент(ы) вставки, с помощью которых нуклеотидная последовательность-мишень встраивается (т.е. клонируется) в реагент нуклеотидной кислоты. В некоторых вариантах осуществления, представленный реагент нуклеиновой кислоты содержит промотор, элемент(ы) вставки и, необязательно, 3'UTR, и 5'UTR/нуклеотидная последовательность-мишень вставляется с необязательной 3'UTR. Элементы могут быть расположены в любом порядке, подходящем для экспрессии в выбранной системе экспрессии (*например*, экспрессии в выбранном организме или экспрессии в бесклеточной системе, *например*), и в некоторых вариантах осуществления, реагент нуклеиновой кислоты, содержащий следующие элементы в направлении 5'-3': (1)

промоторный элемент, 5'UTR, и элемент(ы) вставки; (2) промоторный элемент, 5'UTR и нуклеотидную последовательность-мишень; (3) промоторный элемент, 5'UTR, элемент(ы) вставки и 3'UTR; и (4) промоторный элемент, 5'UTR, нуклеотидную последовательность-мишень и 3'UTR. В некоторых вариантах осуществления, UTR можно оптимизировать для изменения или увеличения транскрипции или трансляции ORF, которые либо полностью являются существующими в природе, либо содержат нуклеотиды не природного происхождения.

Нуклеиновая кислота (*например*, иРНК), содержащая нуклеотидное основание, описанное в настоящем документе, в некоторых случаях, содержит 5' UTR и/или 3' UTR, которые повышают стабильность иРНК *in vivo* (*например*, в эукариотической клетке или эукариотическом SSO). В некоторых случаях, 5' или 3' UTR, или и оба, сконструированы для снижения деградации или распада иРНК *in vivo*. Неограничивающим примером 5' и 3' UTR, повышающим стабильность иРНК в описанных в настоящем документе эукариотических системах, является CS2 3' и 5' UTR. В некоторых вариантах осуществления, иРНК модифицируют для снижения степени удаления поли(А) хвоста иРНК по сравнению с иРНК, содержащей нуклеотидные основания, описанные в настоящем документе, которые не модифицированы иным образом. В некоторых вариантах осуществления, *цис*-действующие AU-богатые элементы (ARE) блокируются от внутри- и внеклеточной передачи сигналов, которая способствует распаду иРНК. В некоторых вариантах осуществления, преждевременные стоп-кодоны в иРНК удаляют из иРНК для уменьшения не смыслового опосредованного распада (NMD) иРНК.

В некоторых случаях, 5' и/или 3' UTR прямо или косвенно увеличивают трансляцию иРНК в полипептид. Неограничивающие примеры того, как 5' UTR или 3' UTR напрямую влияют на трансляцию иРНК в полипептид, включают рекрутирование РНК-связывающих белков, которые связываются с 5' или 3' *цис*-элементами и влияют на рекрутирование рибосом или эффекторных белков (*например*, иРНК-деаденилазы, декапирующих ферментов). Неограничивающие примеры того, как 5'UTR или 3'UTR влияет на трансляцию иРНК в полипептид, косвенно включают образование вторичных структур 5' и 3' UTR, которые блокируют или усиливают связывание РНК-связывающих белков с областями 5' или 3' UTR, и также субклеточную локализацию иРНК.

В некоторых вариантах осуществления, 5'UTR и/или 3'UTR увеличивает эффективность трансляции иРНК *in vitro* или *in vivo* по сравнению с эффективностью трансляции иРНК, содержащей не сконструированное нуклеотидное основание. В некоторых вариантах осуществления, эффективность трансляции увеличивается путем конструирования иРНК для уменьшения проскока выбранных AUG (стартовых кодонов) рибосомой во время сканирования. В некоторых вариантах осуществления, иРНК содержит элементы последовательности, которые улучшают распознавание стартового кодона, такие как *последовательности Козака* или их варианты. В некоторых вариантах осуществления, 5' UTR иРНК сконструирована для снижения общего содержания гуанина-цитозина (GC).

В некоторых вариантах осуществления, снижается образование вторичных структур

в иРНК (*например*, структур G-квадруплекса РНК, RG4), включающих стартовый кодон AUG в 5' UTR, что повышает эффективность трансляции из этой AUG. В некоторых вариантах осуществления, 5'UTR сконструирована так, чтобы иметь отрицательную свободную энергию фолдинга (ΔG) по сравнению с иРНК, которая не была сконструирована. В некоторых вариантах осуществления, ΔG составляет, самое большее -40, -41, -42, -43, -44, -45, -46, -47, -48, -49, -50, -51, -52, -53, -54, -55, -56, -57, -58, -59 или -60. В некоторых вариантах осуществления, мРНК химически модифицирована в 5' UTR или 3' UTR для повышения эффективности трансляции. В некоторых вариантах осуществления, химическая модификация представляет собой N⁶-метиладенозин. В системе *in vitro* (*например*, сконструированной эукариотической клетке или полусинтетическом организме), сверхэкспрессия eIF4A, субъединица комплекса eIF4F, которая способствует раскручиванию РНК вторичных структур в кооперации с eIF3B и eIF4H, повышает эффективность трансляции иРНК. В некоторых вариантах осуществления, нокаут или нокдаун стабилизирующих белков (*например*, белка синдрома ломкой X-хромосомы (FMRP)), которые способствуют образованию вторичной структуры иРНК, снижает образование вторичных структур, тем самым повышая эффективность трансляции иРНК. В некоторых вариантах осуществления, *транс*-действующие агенты (*например*, РНК, малые молекулы, белки) вводят в клетку (*например*, эукариотическую клетку) для стимуляции трансляции иРНК.

В некоторых случаях, 5' UTR и/или 3' UTR способствуют субклеточной локализации иРНК, тем самым способствуя трансляции иРНК *in vivo*. В некоторых вариантах осуществления, 3' или 5' UTR *цис*-действующие элементы, такие как зип-коды иРНК, модифицированы таким образом, что связывание зип-кодов иРНК белками, связывающими зип-коды (*например*, Staufen), подавляется или усиливается, тем самым повышая эффективность трансляции иРНК.

Реагенты нуклеиновой кислоты, *например*, кассеты экспрессии и/или векторы экспрессии (*например*, для экспрессии гетерологичной тРНК-синтетазы), могут включать различные регуляторные элементы, включая промоторы, энхансеры, последовательности инициации трансляции, последовательности терминации транскрипции и другие элементы. «Промотор» обычно представляет собой последовательность или последовательности ДНК, которые функционируют, когда находятся в относительно фиксированной локализации по отношению к сайту начала транскрипции. Например, промотор может располагаться выше сегмента нуклеиновой кислоты транспортера нуклеозидтрифосфата. «Промотор» содержит коровые элементы, необходимые для основного взаимодействия РНК полимеразы и факторов транскрипции, и могут содержать вышерасположенные элементы и элементы ответа. «Энхансер» обычно относится к последовательности ДНК, которая функционирует на не фиксированном расстоянии от сайта начала транскрипции, и может быть либо 5', либо 3' к единице транскрипции. Кроме того, энхансеры могут находиться как внутри интрона, так и внутри самой кодирующей последовательности. Они обычно имеют длину от 10 до 300 и функционируют в *цис*-положении. Энхансеры усиливают

транскрипцию из близлежащих промоторов. Энхансеры, как и промоторы, также часто содержат элементы ответа, которые опосредуют регуляцию транскрипции. Энхансеры часто определяют регуляцию экспрессии и могут использоваться для изменения или оптимизации экспрессии ORF, включая ORF, которые являются полностью природными или содержат нуклеотиды не природного происхождения.

Как отмечалось выше, реагенты нуклеиновой кислоты могут также содержать одну или несколько 5'UTR и одну или несколько 3'UTR. Например, векторы экспрессии, используемые в эукариотических клетках-хозяевах (*например*, дрожжах, грибах, насекомых, растениях, животных, человеческом или ядерных клетках) и прокариотических клетках-хозяевах (*например*, вирусе, бактерии) могут содержать последовательности, которые сигнализируют об окончании транскрипции, что может влиять на экспрессию иРНК. Эти области могут быть транскрибированы как полиаденилированные сегменты в нетранслируемой части иРНК, кодирующей белок фактора ткани. 3' нетранслируемые области также включают сайты терминации транскрипции. В некоторых предпочтительных вариантах осуществления, единица транскрипции содержит область полиаденилирования. Одно из преимуществ этой области заключается в том, что она увеличивает вероятность того, что транскрибируемая единица будет процессироваться и транспортироваться подобно иРНК. Идентификация и использование сигналов полиаденилирования в конструкциях экспрессии хорошо известны. В некоторых предпочтительных вариантах осуществления, гомологичные сигналы полиаденилирования можно использовать в трансгенных конструкциях.

5' UTR может содержать один или несколько элементов, эндогенных по отношению к нуклеотидной последовательности, из которой она происходит, и иногда содержит один или несколько экзогенных элементов. 5'UTR может происходить из любой подходящей нуклеиновой кислоты, такой как геномная ДНК, плазмидная ДНК, РНК или иРНК, например, из любого подходящего организма (*например*, вируса, бактерии, дрожжей, гриба, растения, насекомого или млекопитающего). Специалист в данной области техники может выбрать подходящие элементы для 5' UTR на основе выбранной системы экспрессии (*например*, экспрессии в выбранный организм или экспрессии в бесклеточной системе, например). 5'UTR иногда содержит один или несколько из следующих элементов, известных специалисту в данной области техники: энхансерные последовательности (*например*, транскрипционные или трансляционные), сайт инициации транскрипции, сайт связывания фактора транскрипции, сайт регуляции трансляции, сайт инициации трансляции, сайт связывания фактора трансляции, сайт связывания акцессорного белка, сайты связывания агента регуляции с обратной связью, бокс Прибнова, бокс ТАТА, -35 элемент, Е-бокс (спираль-петля-спираль связывающий элемент), сайт связывания рибосомы, репликон, сайт внутренней посадки рибосомы (IRES), ослабляющий элемент и подобные. В некоторых вариантах осуществления, промоторный элемент может быть выделен таким образом, чтобы все 5' UTR элементы, необходимые для надлежащей условной регуляции, содержались в фрагменте промоторного элемента или внутри

функциональной субпоследовательности фрагмента промоторного элемента.

5'UTR в реагенте нуклеиновой кислоты может содержать нуклеотидную последовательность энхансера трансляции. Нуклеотидная последовательность энхансера трансляции часто расположена между промотором и последовательностью нуклеотидов-мишеней в реагенте нуклеиновой кислоты. Последовательность энхансера трансляции часто связывается с рибосомой, иногда представляет собой рибонуклеотидную последовательность, связывающую 18S рРНК (т. е. последовательность связывания рибосомы 40S), и иногда представляет собой последовательность внутренней посадки рибосомы (IRES). IRES обычно образует каркас РНК с точно расположенными третичными структурами РНК, которые контактируют с субъединицей 40S рибосомы через ряд специфических межмолекулярных взаимодействий. Примеры последовательностей рибосомных энхансеров известны и могут быть идентифицированы специалистом в данной области техники (*например*, Mignone et al., Nucleic Acids Research 33: D141-D146 (2005); Paulous et al., Nucleic Acids Research 31: 722-733 (2003); Akbergenov et al., Nucleic Acids Research 32: 239-247 (2004); Mignone et al., Genome Biology 3(3): reviews0004.1-0001.10 (2002); Gallie, Nucleic Acids Research 30: 3401-3411 (2002); Shaloiko et al., DOI: 10.1002/bit.20267; и Gallie et al., Nucleic Acids Research 15: 3257-3273 (1987)).

Последовательность энхансера трансляции иногда представляет собой эукариотическую последовательность, такую как консенсусная последовательность Козака, или другую последовательность (*например*, последовательность гидроидного полипа, номер доступа в GenBank U07128). Последовательность энхансера трансляции иногда является прокариотической последовательностью, такой как консенсусная последовательность Шайна-Далгарно. В некоторых вариантах осуществления, последовательность энхансера трансляции представляет собой вирусную нуклеотидную последовательность. Последовательность энхансера трансляции иногда происходит из 5' UTR вируса растений, такого как вирус табачной мозаики (TMV), вирус мозаики люцерны (AMV); вирус гравировки табака (ETV); вирус Y картофеля (PVY); вирус мозаики репы (poty) и вирус мозаики, переносимый семенами гороха, например. В некоторых вариантах осуществления, омега последовательность длиной около 67 оснований из TMV включена в реагент нуклеиновой кислоты в качестве последовательности энхансера трансляции (*например*, лишенной гуанозиновых нуклеотидов и включающей поли (CAA) центральную область длиной 25 нуклеотидов).

3' UTR может содержать один или несколько элементов, эндогенных по отношению к нуклеотидной последовательности, из которой она происходит, и иногда включает один или несколько экзогенных элементов. 3' UTR может происходить из любой подходящей нуклеиновой кислоты, такой как геномная ДНК, плазмидная ДНК, РНК или иРНК, например, из любого подходящего организма (*например*, вируса, бактерии, дрожжей, гриба, растения, насекомого или млекопитающего). Специалист в данной области техники может выбрать подходящие элементы для 3' UTR на основе выбранной системы экспрессии (*например*, экспрессии в выбранный организм, например). 3'UTR иногда содержит один или

несколько из следующих элементов, известных специалисту в данной области техники: сайт регуляции транскрипции, сайт инициации транскрипции, сайт терминации транскрипции, сайт связывания фактора транскрипции, сайт регуляции трансляции, сайт терминации трансляции, сайт инициации трансляции, сайт связывания фактора трансляции, сайт связывания рибосомы, репликон, энхансерный элемент, сайленсерный элемент и полиаденозиновый хвост. 3' UTR часто включает полиаденозиновый хвост, и иногда нет, и, если присутствует полиаденозиновый хвост, в него могут быть добавлены или удалены одна или несколько аденозиновых групп (*например*, примерно 5, примерно 10, примерно 15, примерно 20, примерно 25, примерно 30, примерно 35, примерно 40, примерно 45 или примерно 50 аденозиновых групп могут быть добавлены или удалены).

В некоторых вариантах осуществления, модификация 5' UTR и/или 3' UTR используется для изменения (*например*, увеличения, добавления, уменьшения или по существу устранения) активности промотора. Изменение активности промотора, в свою очередь, может изменить активность пептида, полипептида или белка (*например*, активность фермента) путем изменения транскрипции представляющей интерес нуклеотидной последовательности(ей) из функционально связанного промоторного элемента, содержащего модифицированную 5' или 3' UTR. Например, микроорганизм может быть сконструирован путем генетической модификации для экспрессии реагента нуклеиновой кислоты, содержащего модифицированную 5' или 3' UTR, которая может добавить новую активность (*например*, активность, обычно не обнаруживаемую в организме-хозяине) или увеличить экспрессию существующей активности путем повышения транскрипции из гомологичного или гетерологичного промотора, функционально связанного с представляющей интерес нуклеотидной последовательностью (*например*, гомологичной или гетерологичной представляющей интерес нуклеотидной последовательностью), в некоторых вариантах осуществления. А некоторых вариантах осуществления, микроорганизм может быть сконструирован генетической модификацией для экспрессии реагента нуклеиновой кислоты, содержащего модифицированную 5' или 3' UTR, который может снижать экспрессию активности за счет снижения или по существу устранения транскрипции из гомологичного или гетерологичного промотора, функционально связанного с представляющей интерес нуклеотидной последовательностью, в некоторых вариантах осуществления.

Экспрессия гетерологичного полипептида, такого как тРНК синтаза, из кассеты экспрессии или вектора экспрессии может контролироваться любым промотором, способным к экспрессии в прокариотических или эукариотических клетках. Промоторный элемент обычно необходим для синтеза ДНК и/или синтеза РНК. Промоторный элемент часто содержит участок ДНК, который может способствовать транскрипции определенного гена, обеспечивая сайт начала для синтеза РНК, соответствующей гену. Промоторы обычно расположены рядом с генами, которые они регулируют, расположены выше гена (*например*, 5' от гена), и находятся на той же цепи ДНК, что и смысловая цепь гена, в некоторых вариантах осуществления. В некоторых вариантах осуществления,

промоторный элемент может быть выделен из гена или организма и встроен в функциональную связь с полинуклеотидной последовательностью, обеспечивающей измененную и/или регулируемую экспрессию. Не нативный промотор (*например*, промотор, обычно не связанный с данной последовательностью нуклеиновой кислоты), используемый для экспрессии нуклеиновой кислоты, часто называют гетерологичным промотором. В некоторых вариантах осуществления, гетерологичный промотор и/или 5'UTR могут быть вставлены в функциональную связь с полинуклеотидом, который кодирует полипептид, обладающий желаемой активностью, как описано в настоящем документе. Термины «функционально связанный» и «в функциональной связи с», используемые в настоящем документе в отношении промоторов, относятся к взаимосвязи между кодирующей последовательностью и промоторным элементом. Промотор функционально связан или находится в функциональной связи с кодирующей последовательностью, когда экспрессия из кодирующей последовательности через транскрипцию регулируется или контролируется промоторным элементом. Термины «функционально связанный» и «в функциональной связи с» используются в настоящем документе взаимозаменяемо в отношении промоторных элементов.

Промотор часто взаимодействует с РНК-полимеразой. Полимераза представляет собой фермент, который катализирует синтез нуклеиновых кислот с использованием ранее существовавшего реагента нуклеиновой кислоты. Когда матрица представляет собой матрицу ДНК, молекула РНК транскрибируется до синтеза белка. Ферменты, обладающие полимеразной активностью, подходящей для применения в настоящих способах, включают любую полимеразу, активную в выбранной системе, с выбранной матрицей для синтеза белка. В некоторых вариантах осуществления, промотор (*например*, гетерологичный промотор), также называемый в настоящем документе промоторным элементом, может быть функционально связан с нуклеотидной последовательностью или открытой рамкой считывания (ORF). Транскрипция из промоторного элемента может катализировать синтез РНК, соответствующей нуклеотидной последовательности или ORF, функционально связанной с промотором, что, в свою очередь, приводит к синтезу желаемого пептида, полипептида или белка.

Промоторные элементы иногда демонстрируют реакцию на регуляторный контроль. Промоторные элементы также иногда могут регулироваться селективным агентом. То есть, транскрипция из промоторных элементов иногда может включаться, выключаться, усиливаться или подавляться в ответ на изменение в окружающей среде, питании или внутренних условиях или сигналах (*например*, промоторы, индуцируемые нагреванием, промоторы, регулируемые светом, промоторы, регулируемые с обратной связью, промоторы, влияющие на гормоны, тканеспецифические промоторы, промоторы, влияющие на кислород и pH, промоторы, которые реагируют на селективные агенты (*например*, канамицин) и подобные, например). Промоторы, находящиеся под влиянием сигналов окружающей среды, питания или внутренних сигналов, часто подвергаются влиянию сигнала (прямого или косвенному), который связывается на или рядом с

промотором, и увеличивает или уменьшает экспрессию последовательности-мишени при определенных условиях. Что касается всех описанных в настоящем документе способов, включение природных или модифицированных промоторов можно использовать для изменения или оптимизации экспрессии полностью природной ORF (например, aaRS) или ORF, содержащей нуклеотид не природного происхождения (например, иРНК или тРНК).

Неограничивающие примеры селективных или регуляторных агентов, которые влияют на транскрипцию из промоторного элемента, используемого в вариантах осуществления, описанных в настоящем документе, включают, без ограничения, (1) сегменты нуклеиновых кислот, которые кодируют продукты, которые представляют резистентность к другим токсичным соединениям (*например*, антибиотикам); (2) сегменты нуклеиновой кислоты, которые кодируют продукты, которые иначе отсутствуют в клетке-реципиенте (*например*, основные продукты, гены тРНК, ауксотрофные маркеры); (3) сегменты нуклеиновых кислот, которые кодируют продукты, подавляющие активность генного продукта; (4) сегменты нуклеиновой кислоты, которые кодируют продукты, которые можно легко идентифицировать (*например*, фенотипические маркеры, такие как антибиотики (*например*, β -лактамаза), β -галактозидаза, зеленый флуоресцентный белок (GFP).), желтый флуоресцентный белок (YFP), красный флуоресцентный белок (RFP), голубой флуоресцентный белок (CFP) и белки клеточной поверхности); (5) сегменты нуклеиновой кислоты, которые связывают продукты, которые в противном случае вредны для выживания и/или функции клеток; (6) сегменты нуклеиновой кислоты, которые иным образом ингибируют активность любого из сегментов нуклеиновой кислоты, описанных выше в пп. 1-5 (*например*, антисмысловые олигонуклеотиды); (7) сегменты нуклеиновой кислоты, которые связывают продукты, модифицирующие субстрат (*например*, рестрикционные эндонуклеазы); (8) сегменты нуклеиновой кислоты, которые можно использовать для выделения или идентификации желаемой молекулы (*например*, сайты связывания специфического белка); (9) сегменты нуклеиновой кислоты, которые кодируют специфическую нуклеотидную последовательность, которая в противном случае может быть не функциональной (*например*, для ПЦР амплификации субпопуляций молекул); (10) сегменты нуклеиновой кислоты, которые, при отсутствии, прямо или косвенно придают устойчивость или чувствительность к конкретным соединениям; (11) сегменты нуклеиновой кислоты, которые кодируют продукты, которые либо являются токсичными, либо превращают относительно нетоксичное соединение в токсичное соединение (*например*, тимидинкиназа Herpes simplex, цитозиндезаминаза) в клетках-реципиентах; (12) сегменты нуклеиновой кислоты, которые ингибируют репликацию, разделение или наследуемость молекул нуклеиновых кислот, которые их содержат; (13) сегменты нуклеиновой кислоты, которые кодируют условные функции репликации, *например*, репликацию в определенных хозяевах или штаммах клеток-хозяев или в определенных условиях окружающей среды (*например*, температура, условия питания и подобные); и/или (14) нуклеиновые кислоты, которые кодируют одну или несколько иРНК или тРНК, содержащие нуклеотиды не природного происхождения. В некоторых вариантах

осуществления, регуляторный или селективный агент может быть добавлен для изменения существующих условий роста, которым подвергается организм (*например*, роста в жидкой культуре, роста в ферментере, роста на твердых питательных чашках и подобное, *например*).

В некоторых вариантах осуществления, регулирование промоторного элемента можно использовать для изменения (*например*, увеличения, добавления, снижения или по существу устранения) активности пептида, полипептида или белка (*например*, активности фермента). Например, микроорганизм может быть сконструирован путем генетической модификации для экспрессии реагента нуклеиновой кислоты, который может добавить новую активность (*например*, активность, обычно не обнаруживаемую в организме-хозяине) или увеличить экспрессию существующей активности путем усиления транскрипции из гомологичного или гетерологичного промотора, функционально связанного с представляющей интерес последовательностью нуклеотидов (*например*, гомологичной или гетерологичной представляющей интерес последовательностью нуклеотидов) в некоторых вариантах осуществления. В некоторых вариантах осуществления, микроорганизм может быть сконструирован путем генетической модификации для экспрессии реагента нуклеиновой кислоты, который может снижать экспрессию активности путем уменьшения или по существу удаления транскрипции с гомологичного или гетерологичного промотора, функционально связанного с интересующей нуклеотидной последовательностью, в определенных вариантах осуществления.

Нуклеиновые кислоты, кодирующие гетерологичные белки, например тРНК-синтетазы, могут быть встроены или использованы с любой подходящей системой экспрессии. В некоторых вариантах осуществления, реагент нуклеиновой кислоты иногда стабильно интегрирован в хромосому организма-хозяина, или реагент нуклеиновой кислоты может быть делецией части хромосомы-хозяина, в некоторых вариантах осуществления, (*например*, генетически модифицированных организмов, где изменение генома-хозяина дает возможность селективно или предпочтительно поддерживать желаемый организм, несущий генетическую модификацию). Такие реагенты нуклеиновой кислоты (*например*, нуклеиновые кислоты или генетически модифицированные организмы, чей измененный геном придает организму селективный признак) могут быть отобраны по их способности направлять продуцирование желаемого белка или молекулы нуклеиновой кислоты. При желании, реагент нуклеиновой кислоты может быть изменен так, что кодоны кодируют (i) ту же аминокислоту, с использованием другой тРНК, отличной от той, которая указана в нативной последовательности, или (ii) отличающуюся от обычной аминокислоту, включающую нестандартные аминокислоты или аминокислоты не природного происхождения (включая определяемо меченые аминокислоты).

Рекомбинантную экспрессию полезно осуществлять с использованием кассеты экспрессии, которая может быть частью вектора, такого как плазида. Вектор может включать промотор, функционально связанный с нуклеиновой кислотой. Вектор также

может включать другие элементы, необходимые для транскрипции и трансляции, как описано в настоящем документе. Кассета экспрессии, вектор экспрессии и последовательности в кассете или векторе могут быть гетерологичны клетке, с которой контактируют нуклеотиды не природного происхождения.

Могут быть получены различные прокариотические и эукариотические векторы экспрессии, подходящие для переноса, кодирования и/или экспрессии гетерологичного белка, такого как тРНК синтаза. Такие векторы экспрессии включают, например, pET, pET3d, pCR2.1, pBAD, pUC и дрожжевые векторы. Векторы можно использовать, например, в различных ситуациях *in vivo* и *in vitro*. Неограничивающие примеры прокариотических промоторов, которые можно использовать, включают SP6, T7, T5, *tac*, *bla*, *trp*, *gal*, *lac* или мальтозные промоторы. Неограничивающие примеры эукариотических промоторов, которые можно использовать, включают конститутивные промоторы, *например*, вирусные промоторы, такие как промоторы CMV, SV40 и RSV, а также регулируемые промоторы, *например*, индуцируемый или репресслируемый промотор, такой как *tet* промотор, hsp70 промотор и синтетический промотор, регулируемый CRE. Векторы для бактериальной экспрессии включают pGEX-5X-3, и для эукариотической экспрессии включают pCIneo-CMV. Вирусные векторы, которые можно использовать, включают векторы, относящиеся к лентивирусам, аденовирусам, ассоциированный вирус, вирус герпеса, вирус коровьей оспы, вирус полиомиелита, вирус AIDS, нейрональный трофический вирус, вирус Синдбис и другие вирусы. Также пригодны любые вирусные семейства, обладающие свойствами этих вирусов, которые делают их подходящими для использования в качестве векторов. Ретровирусные векторы, которые могут применяться, включают описанные в Verma, American Society for Microbiology, pp. 229-232, Washington, (1985). Например, такие ретровирусные векторы могут включать вирус лейкоза Малони мыши, MMLV и другие ретровирусы, которые проявляют желательные свойства. Обычно, вирусные векторы содержат неструктурные ранние гены, структурные поздние гены, транскрипт РНК полимеразы III, инвертированные концевые повторы, необходимые для репликации и инкапсуляции, и промоторы для контроля транскрипции и репликации вирусного генома. При разработке в качестве векторов вирусы обычно имеют один или несколько удаленных ранних генов, и ген или кассету гена/промотора вставляют в вирусный геном вместо удаленной вирусной нуклеиновой кислоты.

Клонирование

Для включения элемента, такого как ORF, в реагент нуклеиновой кислоты можно использовать любую удобную стратегию клонирования, известную в данной области техники. Известные способы можно использовать для вставки элемента в матрицу независимо от элемента вставки, например (1) расщепление матрицы по одному или нескольким существующим сайтам рестрикционных ферментов и лигирование представляющего интерес элемента, и (2) добавление сайтов рестрикционных ферментов к матрице путем гибридизации олигонуклеотидных праймеров, которые включают один или несколько подходящих сайтов рестрикционных ферментов и амплификацию с помощью

полимеразной цепной реакции (описанной в более подробно в настоящем документе). В других стратегиях клонирования используется преимущество одного или нескольких сайтов вставки, присутствующих или встроенных в реагент нуклеиновой кислоты, например, сайта гибридизации олигонуклеотидного праймера для ПЦР, и других, описанных в настоящем документе. В некоторых вариантах осуществления, стратегия клонирования может быть объединена с генетическими манипуляциями, такими как рекомбинация (*например*, рекомбинация реагента нуклеиновой кислоты с последовательностью нуклеиновой кислоты, представляющей интерес, в геном модифицированного организма, как описано далее в настоящем документе). В некоторых вариантах осуществления, клонированная ORF может продуцировать (прямо или косвенно) модифицированные полимеразы или полимеразы дикого типа путем конструирования микроорганизма с одной или несколькими представляющими интерес ORF, где микроорганизмы содержат измененную активность полимеразной активности.

Нуклеиновая кислота может быть специфически расщеплена при контакте нуклеиновой кислоты с одним или несколькими специфическими расщепляющими агентами. Специфические расщепляющие агенты часто будут специфически расщеплять в соответствии с определенной последовательностью нуклеотидов на конкретном участке. Примеры фермент-специфических расщепляющих агентов включают, без ограничения, эндонуклеазы (*например*, ДНКазу (*например*, ДНКазу I, II); РНКазу (*например*, РНКазу E, F, H, P); фермент Cleavase™; Taq ДНК полимеразу; ДНК полимеразу I *E. coli* и эукариотические структурно-специфические эндонуклеазы; эндонуклеазы FEN-1 мыши; рестрикционные эндонуклеазы типа I, II или III, такие как Acc I, Afl III, Alu I, Alw44 I, Apa I, Asn I, Ava I, Ava II, BamH I, Ban II, Bcl I, Bgl I, Bgl II, Bln I, BsaI, Bsm I, BsmBI, BssH II, BstE II, Cfo I, Cla I, Dde I, Dpn I, Dra I, EcIX I, EcoR I, EcoR II, EcoR V, Hae II, Hae III, Hind II, Hind III, Hpa I, Hpa II, Kpn I, Ksp I, Mlu I, MluN I, Msp I, Nci I, Nco I, Nde I, Nde II, Nhe I, Not I, Nru I, Nsi I, Pst I, Pvu I, Pvu II, Rsa I, Sac I, Sal I, Sau3A I, Sca I, ScrF I, Sfi I, Sma I, Spe I, Sph I, Ssp I, Stu I, Sty I, Swa I, Taq I, Xba I, Xho I); гликозилазы (*например*, урацил-ДНК-гликозилазу (UDG), 3-метиладенин-ДНК-гликозилазу, 3-метиладенин-ДНК-гликозилазу II, пиримидингидрат-ДНК-гликозилазу, FaPy-ДНК-гликозилазу, тимин-несовместимую ДНК гликозилазу, гипоксантин-ДНК гликозилазу, 5-гидроксиметилурацил ДНК гликозилазу (HmUDG), 5-гидроксиметилцитозин ДНК гликозилазу или 1,N6-этенoadенин ДНК гликозилазу); экзонуклеазы (*например*, экзонуклеазу III); рибозимы и ДНКзимы. Образец нуклеиновой кислоты можно обработать химическим агентом или синтезировать с использованием модифицированных нуклеотидов, и модифицированная нуклеиновая кислота может быть расщеплена. В неограничивающих примерах, образец нуклеиновой кислоты можно обработать (i) алкилирующими агентами, такими как метилнитрозомочевина, которые образуют несколько алкилированных оснований, включая N3-метиладенин и N3-метилгуанин, которые распознаются и расщепляются алкилпуриновой ДНК-гликозилазой; (ii) бисульфитом натрия, который вызывает дезаминирование цитозиновых остатков в ДНК с образованием урацильных остатков,

которые могут быть расщеплены урацил N-гликозилазой, и (iii) химическим агентом, который превращает гуанин в его окисленную форму, 8-гидроксигуанином, который может быть расщеплен формамидопиримидин ДНК N-гликозилазой. Примеры процессов химического расщепления включают, без ограничения, алкилирование (*например*, алкилирование фосфоротиоат-модифицированной нуклеиновой кислоты), расщепление кислотнеустойчивой нуклеиновой кислоты, содержащей P3'-N5'-фосфорамидат; и обработку нуклеиновой кислоты тетроксидом осмия и пиперидином.

В некоторых вариантах осуществления, реагент нуклеиновой кислоты включает один или несколько сайтов встраивания рекомбиназы. Сайт встраивания рекомбиназы представляет собой последовательность распознавания на молекуле нуклеиновой кислоты, которая участвует в реакции интеграции/рекомбинации рекомбинационными белками. Например, сайт рекомбинации для рекомбиназы Cre представляет собой loxP, который представляет собой последовательность из 34 пар оснований, состоящую из двух инвертированных повторов из 13 пар оснований (служащих сайтами связывания рекомбиназы), фланкирующих коровую последовательность из 8 пар оснований (*например*, Sauer, Curr. Opin. Biotech. 5:521-527 (1994)). Другие примеры сайтов рекомбинации включают последовательности attB, attP, attL и attR, и их мутанты, фрагменты, варианты и производные, которые распознаются рекомбинационным белком λ Int и вспомогательными белками фактора интеграции хозяина (IHf), FIS и эксцизионазой (Xis) (*например*, патенты США №№ . 5,888,732; 6,143,557; 6,171,861; 6,270,969; 6,277,608; и 6,720,140; заявки на патент США №№ 09/517,466 и 09/732,914; публикацию патента США № 002/0007051; и Landy, Curr. Opin. Biotech., 3:699-707 (1993)).

Примерами рекомбиназного клонирования нуклеиновых кислот являются системы Gateway® (Invitrogen, California), которые содержат, по меньшей мере, один сайт рекомбинации для клонирования желаемых молекул нуклеиновых кислот *in vivo* или *in vitro*. В некоторых вариантах осуществления, в системе используются векторы, которые содержат, по меньшей мере, два разных сайт-специфических сайтов рекомбинации, часто основанных на лямбда системе бактериофага (*например*, att1 и att2), и мутируют из сайтов дикого типа (att0). Каждый мутированный сайт имеет уникальную специфичность для att сайта его родственного партнера (т.е. сайта рекомбинации его партнера по связыванию) того же типа (*например*, attB1 с attP1 или attL1 с attR1) и не будет перекрестно реагировать с сайтами рекомбинации другого мутантного типа или с сайтом att0 дикого типа. Разная сайт-специфичность позволяет направленное клонирование или связывание желаемых молекул, таким образом обеспечивая желаемую ориентацию клонированных молекул. Фрагменты нуклеиновой кислоты, фланкированные сайтами рекомбинации, клонируют и субклонируют с использованием системы Gateway® путем замены селективируемого маркера (*например*, ccdB), фланкированного att сайтами, на молекуле плазмиды-реципиента, иногда называемой принимающим вектором. Затем нужные клоны отбирают путем трансформации штамма-хозяина, чувствительного к ccdB, и положительной селекцией по маркеру на молекуле-реципиенте. Аналогичные стратегии отрицательной селекцией (*например*,

использование токсичных генов) можно использовать и в других организмах, таких как тимидинкиназа (ТК) у млекопитающих и насекомых.

Реагент нуклеиновой кислоты иногда содержит один или несколько элементов начала репликации (ORI). В некоторых вариантах осуществления, матрица содержит два или несколько ORI, где один эффективно функционирует в одном организме (*например*, бактерии), и другой эффективно функционирует в другом организме (*например*, эукариоте, таком как, например, дрожжи). В некоторых вариантах осуществления, ORI может эффективно функционировать у одного вида (*например*, *S. cerevisiae*), и другой ORI может эффективно функционировать у других видов (*например*, *S. pombe*). Реагент нуклеиновой кислоты также иногда включает один или несколько сайтов регуляции транскрипции.

Реагент нуклеиновой кислоты, *например*, кассета или вектор экспрессии, может включать последовательность нуклеиновой кислоты, кодирующую маркерный продукт. Маркерный продукт используют для определения того, был ли ген доставлен в клетку и экспрессируется ли он после доставки. Примеры маркерных генов включают ген *lacZ* *E. coli*, который кодирует β -галактозидазу и зеленый флуоресцентный белок. В некоторых вариантах осуществления, маркер может быть селективируемым маркером. Когда такие селективируемые маркеры успешно переносятся в клетку-хозяин, трансформированная клетка-хозяин может выжить при воздействии селективного давления. Существуют две широко используемые отдельные категории селективных схем. Первая категория основана на клеточном метаболизме и использовании мутантной клеточной линии, которая не способна расти независимо от дополненной среды. Вторая категория представляет собой доминантный отбор, который относится к схеме селекции, используемой для любого типа клеток, и не требует использования мутантной клеточной линии. В этих схемах обычно используется лекарственное средство для остановки роста клетки-хозяина. Те клетки, которые имеют новый ген, будут экспрессировать белок, передающий лекарственную резистентность, и выживут при селекции. Примеры такой доминантной селекции используют лекарственные препараты неомицин (Southern et al., J. Molec. Appl. Genet. 1: 327 (1982)), микофеноловую кислоту (Mulligan et al., Science 209: 1422 (1980)) или гигромицин (Sugden, et al., Mol. Cell. Biol. 5: 410-413 (1985)).

Реагент нуклеиновой кислоты может включать один или несколько элементов селекции (*например*, элементов для селекции присутствия реагента нуклеиновой кислоты, а не для активации промоторного элемента, который можно селективно регулировать). Элементы селекции часто используют с использованием известных процессов для определения того, включен ли реагент нуклеиновой кислоты в клетку. В некоторых вариантах осуществления, реагент нуклеиновой кислоты включает два или несколько элементов селекции, где один эффективно функционирует в одном организме, а другой эффективно функционирует в другом организме. Примеры элементов селекции включают, но не ограничиваются ими: (1) сегменты нуклеиновых кислот, которые кодируют продукты, обеспечивающие резистентность к токсичным соединениям (*например*, антибиотикам); (2) сегменты нуклеиновых кислот, которые кодируют продукты, которых

нет в клетке-реципиенте (*например*, жизненно важные продукты, гены тРНК, ауксотрофные маркеры); (3) сегменты нуклеиновых кислот, которые кодируют продукты, которые подавляют активность генного продукта; (4) сегменты нуклеиновых кислот, которые кодируют легко идентифицируемые продукты (*например*, фенотипические маркеры, такие как антибиотики (*например*, β -лактамаза), β -галактозидаза, зеленый флуоресцентный белок (GFP), желтый флуоресцентный белок (YFP), красный флуоресцентный белок (RFP), голубой флуоресцентный белок (CFP) и белки клеточной поверхности); (5) сегменты нуклеиновых кислот, которые связывают продукты, которые в противном случае вредны для выживания и/или функции клеток; (6) сегменты нуклеиновых кислот, которые иначе ингибируют активность любого из сегментов нуклеиновой кислоты, описанных выше в пп. 1-5 (*например*, антисмысловые олигонуклеотиды); (7) сегменты нуклеиновых кислот, которые связывают продукты, модифицирующие субстрат (*например*, рестрикционные эндонуклеазы); (8) сегменты нуклеиновых кислот, которые можно использовать для выделения или идентификации желаемой молекулы (*например*, специфические сайты связывания с белком); (9) сегменты нуклеиновых кислот, которые кодируют специфическую нуклеотидную последовательность, которая в противном случае может быть нефункциональной (*например*, для ПЦР-амплификации субпопуляций молекул); (10) сегменты нуклеиновых кислот, которые, при отсутствии, прямо или косвенно придают резистентность или чувствительность к определенным соединениям; (11) сегменты нуклеиновых кислот, которые кодируют продукты, которые либо являются токсичными, либо превращают относительно нетоксичное соединение в токсичное соединение (*например*, тимидинкиназа Herpes simplex, цитозиндезаминаза) в клетках-реципиентах; (12) сегменты нуклеиновых кислот, которые ингибируют репликацию, разделение или наследуемость молекул нуклеиновых кислот, которые их содержат; и/или (13) сегменты нуклеиновых кислот, которые кодируют условные функции репликации, *например*, репликацию в определенных хозяевах или штаммах клеток-хозяев или при определенных условиях окружающей среды (*например*, температуре, условиях питания и подобных).

Реагент нуклеиновой кислоты может быть в любой форме, пригодной для транскрипции и/или трансляции *in vivo*. Нуклеиновая кислота иногда представляет собой плазмиду, такую как сверхспиральная плазида, иногда представляет собой искусственную хромосому дрожжей (*например*, YAC), иногда представляет собой линейную нуклеиновую кислоту (*например*, линейную нуклеиновую кислоту, полученную ПЦР или рестриктазными фрагментами), иногда она является одноцепочечной, и иногда является двухцепочечной. Реагент нуклеиновой кислоты иногда получают с помощью процесса амплификации, такого как процесс полимеразной цепной реакции (ПЦР) или процесс амплификации, опосредованный транскрипцией (ТМА). В ТМА два фермента используются в изотермической реакции для продуцирования продуктов амплификации, определяемых испусканием света (*например*, Biochemistry 1996 Jun 25;35(25):8429-38). Стандартные ПЦР процессы известны (*например*, патенты США №№ 4,683,202; 4,683,195; 4,965,188; и 5,656,493) и, как правило, выполняются циклами. Каждый цикл включает

денатурацию нагреванием, при которой диссоциируют гибридные нуклеиновые кислоты, охлаждение, при котором олигонуклеотиды праймера гибридизируются; и удлинение олигонуклеотидов полимеразой (т.е. полимеразой Taq). Примером циклического процесса ПЦР является обработка образца при 9°C в течение 5 минут; повторение сорока пяти циклов при 95°C в течение 1 минуты, 59°C в течение 1 минуты, 10 секунд и 72°C в течение 1 минуты 30 секунд; и затем обработка образца при 72°C в течение 5 минут. Несколько циклов часто проводят с использованием коммерчески доступного термоциклера. Иногда продукты ПЦР амплификации хранят в течение время при более низкой температуре (*например*, при 4°C) и иногда перед анализом замораживают (*например*, при -20°C).

Стратегии клонирования, аналогичные описанным выше, могут быть использованы для получения ДНК, содержащей нуклеотиды не природного происхождения. Например, олигонуклеотиды, содержащие нуклеотиды не природного происхождения в нужных положениях, синтезируют с использованием стандартного твердофазного синтеза и очищаются с помощью ВЭЖХ. Затем олигонуклеотиды встраиваются в плазмиду, содержащую требуемый контекст последовательности (т.е. UTR и кодирующую последовательность) с использованием способа клонирования (такого как Golden Gate Assembly) с сайтами клонирования, такими как сайты BsaI (хотя можно использовать и другие, обсуждавшиеся выше).

Наборы/Готовые изделия

В некоторых вариантах осуществления, в настоящем документе описаны наборы и готовые изделия для использования с одним или несколькими способами, описанными в настоящем документе. Такие наборы включают носитель, упаковку или контейнер, разделенный на отсеки для приема одного или нескольких контейнеров, таких как флаконы, пробирки и подобные, где каждый из этих контейнеров содержит один из отдельных элементов, применяемых в способе, описанном в настоящем документе. Подходящие контейнеры включают, например, бутылки, флаконы, шприцы и пробирки. В одном варианте осуществления, контейнеры сформированы из различных материалов, таких как стекло или пластик.

В некоторых вариантах осуществления, набор включает подходящий упаковочный материал для размещения содержимого набора. В некоторых случаях, упаковочный материал предпочтительно изготавливается с использованием хорошо известных методов, чтобы обеспечить стерильную, свободную от загрязнений среду. Упаковочные материалы, используемые в настоящем документе, могут включать, например, те, которые обычно используются в коммерческих наборах, продаваемых для использования с системами секвенирования нуклеиновых кислот. Примеры упаковочных материалов включают, без ограничений, стекло, пластик, бумагу, фольгу и подобные, способные удерживать в фиксированных пределах набор компонентов, представленный в настоящем документе.

Упаковочный материал может включать этикетку, на которой указано конкретное использование компонентов. Использование набора, указанное на этикетке, может представлять собой один или несколько способов, изложенных в настоящем документе, в

зависимости от конкретной комбинации компонентов, присутствующих в наборе. Например, на этикетке может быть указано, что набор можно использовать для способа синтеза полинуклеотида или для способа определения последовательности нуклеиновой кислоты.

Инструкции по использованию упакованных реагентов или компонентов также могут быть включены в набор. Инструкции, как правило, включают материальное выражение, описывающее параметры реакции, такие как относительное количество компонентов набора и смешиваемого образца, периоды сохранения смесей реагента/образца, температуру, буферные условия и подобные.

Следует понимать, что не все компоненты, необходимые для конкретной реакции, должны присутствовать в конкретном наборе. Скорее, один или несколько дополнительных компонентов могут быть предоставлены из других источников. В инструкциях, прилагаемых к набору, могут быть указаны дополнительные компоненты, которые должны быть предоставлены, и где их можно получить.

В некоторых вариантах осуществления, представлен набор, пригодный для стабильного включения нуклеиновой кислоты не природного происхождения в клеточную нуклеиновую кислоту, *например*, с использованием способов, представленных настоящим изобретением, для получения генетически сконструированных клеток млекопитающих (*например*, клеток СНО или НЕК293Т). В одном варианте осуществления, набор, описанный в настоящем документе, включает генетически сконструированную клетку и одну или несколько нуклеиновых кислот не природного происхождения.

В дополнительных вариантах осуществления, набор, описанный в настоящем документе, представляет клетку и молекулу нуклеиновой кислоты, содержащую гетерологичный ген для введения в клетку, чтобы тем самым получить генетически сконструированную клетку, такую как векторы экспрессии, содержащие нуклеиновую кислоту из любого из вариантов осуществления, описанных выше в этом параграфе.

В некоторых вариантах осуществления, клетка, описанная в настоящем документе, доставляется в организм, который может быть многоклеточным организмом, таким как млекопитающее, например, человек. Как таковые, эукариотические клетки, содержащие полипептид, имеющий аминокислоту не природного происхождения, могут быть введены в организм.

Пронумерованные Варианты Осуществления

Настоящее описание включает следующие неограничивающие пронумерованные варианты осуществления:

Вариант осуществления 1. Способ получения полипептида, содержащего одну или несколько аминокислот не природного происхождения, в эукариотической клетке, включающий:

(а) предоставление эукариотической клетки, содержащей:

(i) транспортную РНК (тРНК) с антикодоном, содержащим первое основание не природного происхождения;

(ii) информационную РНК (иРНК) с кодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения образуют пару оснований не природного происхождения (UBP) в эукариотической клетке;

(b) трансляцию полипептида, содержащего одну или несколько аминокислот не природного происхождения, из иРНК с использованием тРНК с помощью рибосомы, которая является эндогенной для эукариотической клетки.

Вариант осуществления 2. Способ по варианту осуществления 1, где кодон иРНК содержит три непрерывных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

Вариант осуществления 3. Способ по варианту осуществления 1, где кодон иРНК содержит три непрерывных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

Вариант осуществления 4. Способ по варианту осуществления 1, где кодон иРНК содержит три непрерывных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

Вариант осуществления 5. Способ по любому из вариантов осуществления 1-4, где первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из:

(i) 2-тиоурацила, 2-тиотимина, 2'-дезоксинуридина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-галоурацила; 5-пропинилурацила, 6-азотимина, 6-азоурацила, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила, псевдоурацила, метилового эфира урацил-5-оксуксусной кислоты, урацил-5-оксуксусной кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, 5-метил-2-тиоурацила, 4-тиоурацила, 5-метилурацила, 5'-метоксикарбоксиметилурацила, 5-метоксиурацила, урацил-5-оксиуксусной кислоты, 5-(карбоксигидроксиметил)урацила, 5-карбоксиметиламинометил-2-тиуридина, 5-карбоксиметиламинометилурацила или дигидроурацила;

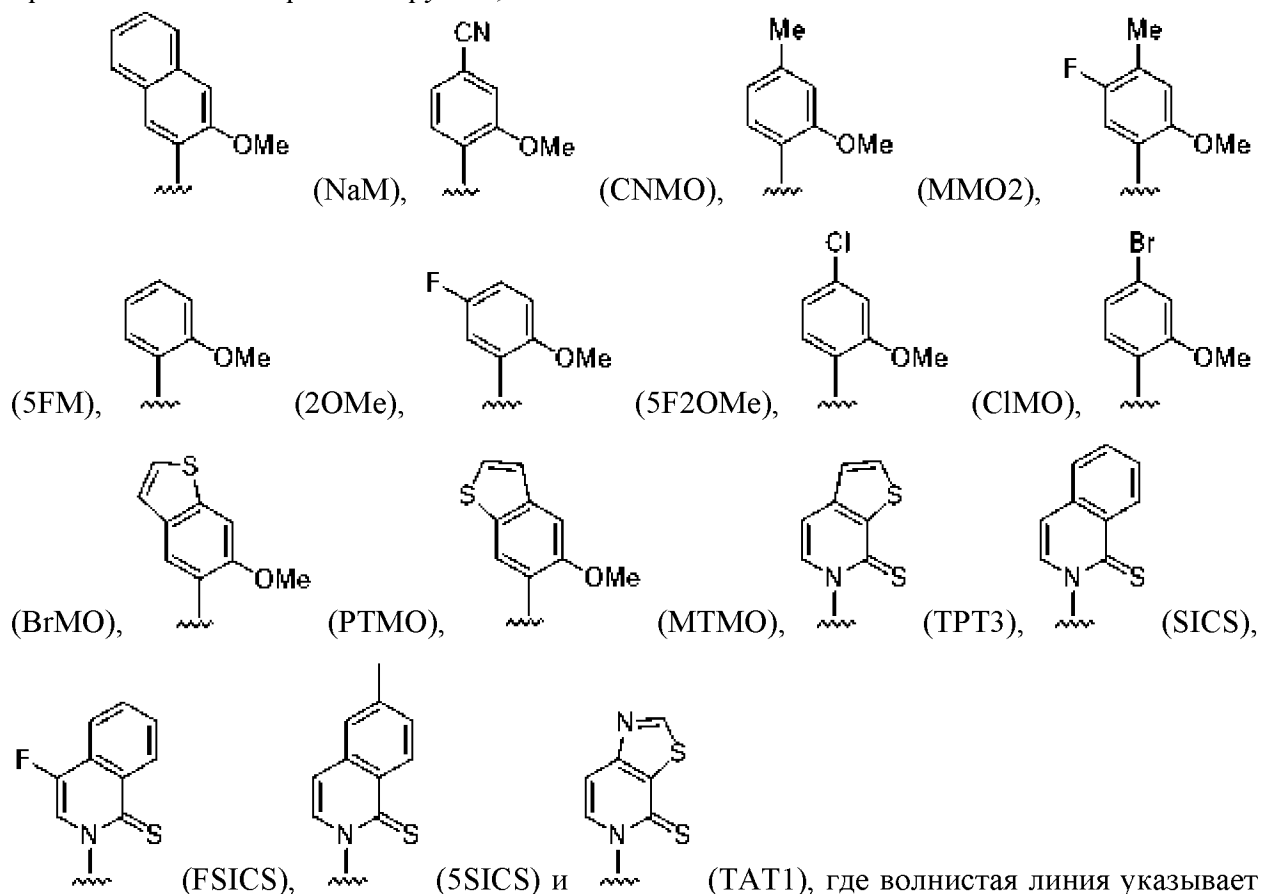
(ii) 5-гидроксиметилцитозина, 5-трифторметилцитозина, 5-галоцитозина, 5-пропинилцитозина, 5-гидроксицитозина, циклоцитозина, цитозинарабинозида, 5,6-дигидроцитозина, 5-нитроцитозина, 6-азоцитозина, азацитозина, N4-этилцитозина, 3-метилцитозина, 5-метилцитозина, 4-ацетилцитозина, 2-тиоцитозина, феноксазин цитидин ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазин цитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), феноксазин цитидин (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазол цитидин (2H-пиримидо[4,5-b]индол-2-она) или пиридоиндол цитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она);

(iii) 2-аминоаденина, 2-пропиладенина, 2-аминоаденина, 2-F-аденина, 2-аминопропиладенина, 2-амино-2'-дезоксияденозина, 3-дезааденина, 7-метиладенина, 7-дезааденина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных аденинов, N6-изопентениладенина, 2-метиладенина, 2,6-диаминопурин, 2-метилтио-N6-изопентениладенина или 6-азааденина;

(iv) 2-метилгуанина, 2-пропил- и алкилпроизводных гуанина, 3-деазагуанина, 6-тиогуанина, 7-метилгуанина, 7-деазагуанина, 7-деазагуанозина, 7-деаза-8-азагуанина, 8-азагуанина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных гуанинов, 1-метилгуанина, 2,2-диметилгуанина, 7-метилгуанина или 6-азагуанина; и

(v) гипоксантина, ксантина, 1-метилюнозина, кеозина, бета-D-галактозилкеозина, инозина, бета-D-маннозилкеозина, вибутоксозина, гидроксимочевины, (аср3)w, 2-аминопиридины или 2-пиридоны.

Вариант осуществления 6. Способ по любому из вариантов осуществления 1-4, где первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из

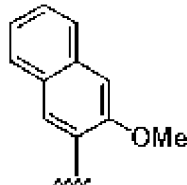


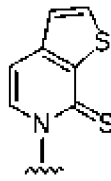
на связь с рибозильной группой.

Вариант осуществления 7. Способ по варианту осуществления 6, когда первое

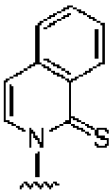
основание не природного происхождения представляет собой (TPT3), второе

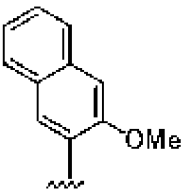
основание не природного происхождения представляет собой (NaM), и когда

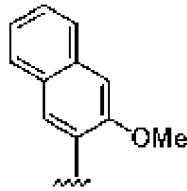
первое основание не природного происхождения представляет собой  (NaM),

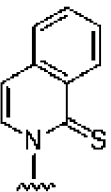
второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 8. Способ по варианту осуществления 6, когда первое

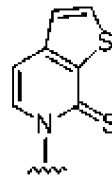
основание не природного происхождения представляет собой  (5SICS), второе

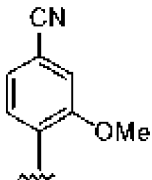
основание не природного происхождения представляет собой  (NaM), и когда

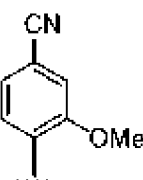
первое основание не природного происхождения представляет собой  (NaM),

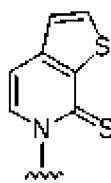
второе основание не природного происхождения представляет собой  (5SICS), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 9. Способ по варианту осуществления 6, когда первое

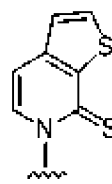
основание не природного происхождения представляет собой  (TPT3), второе

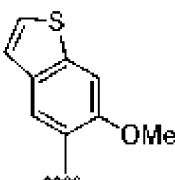
основание не природного происхождения представляет собой  (CNMO), и когда

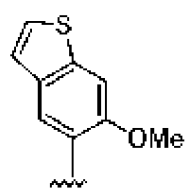
первое основание не природного происхождения представляет собой  (CNMO),

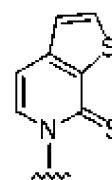
второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 10. Способ по варианту осуществления 6, когда первое

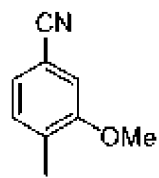
основание не природного происхождения представляет собой  (TPT3), второе

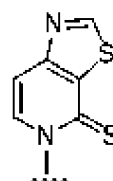
основание не природного происхождения представляет собой  (PTMO), и

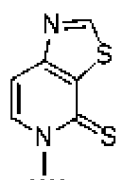
когда первое основание не природного происхождения представляет собой 

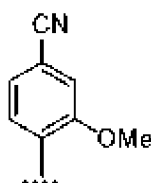
(PTMO), второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 11. Способ по варианту осуществления 6, когда первое

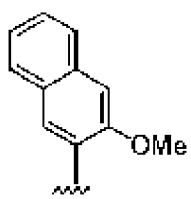
основание не природного происхождения представляет собой  (CNMO), второе

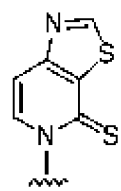
основание не природного происхождения представляет собой  (TAT1), и когда

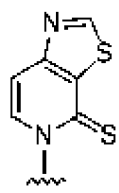
первое основание не природного происхождения представляет собой  (TAT1),

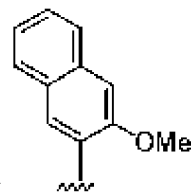
второе основание не природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 12. Способ по варианту осуществления 6, когда первое

основание не природного происхождения представляет собой  (NaM), второе

основание не природного происхождения представляет собой  (TAT1), и когда

первое основание не природного происхождения представляет собой  (TAT1),

второе основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 13. Способ по любому из вариантов осуществления 1-12, где первое основание не природного происхождения или второе основание не природного происхождения содержат модифицированную сахарную группу, выбранный из группы, состоящей из:

модификации в положении 2':

ОН, замещенный низший алкил, алкирил, аралкил, О-алкирил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F;

О-алкил, S-алкил, N-алкил;

О-алкенил, S-алкенил, N-алкенил;

О-алкинил, S-алкинил, N-алкинил;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, где алкил, алкенил и алкинил могут быть замещены или не замещены C₁-C₁₀ алкилом, C₂-C₁₀ алкенилом, C₂-C₁₀ алкинилом, -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-NH₂ и -O(CH₂)_nON[(CH₂)_nCH₃]₂, где n и m составляют от 1 до примерно 10;

и/или модификации в положении 5':

5'-винил, 5'-метил (R или S);

модификации в положении 4':

4'-S, гетероциклоалкил, гетероциклоалкарил, аминокламино, полиалкиламино, замещенный силлил, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для улучшения фармакодинамических свойств олигонуклеотида, и любая их комбинация.

Вариант осуществления 14. Способ по любому из вариантов осуществления 1-13, где способ представляет собой человеческую клетку.

Вариант осуществления 15. Способ по варианту осуществления 14, где клетка человека представляет собой клетку НЕК293Т.

Вариант осуществления 16. Способ по любому из вариантов осуществления 1-13, где клетка представляет собой клетку хомяка.

Вариант осуществления 17. Способ по варианту осуществления 16, где клетка хомяка представляет собой клетку яичника китайского хомяка (СНО).

Вариант осуществления 18. Способ по любому из вариантов осуществления 1-17, где аминокислота не природного происхождения:

представляет собой аналог лизина;

содержит ароматическую боковую цепь;

содержит азидогруппу;

содержит алкиновую группу; или

содержат альдегидную или кетонную группу.

Вариант осуществления 19. Способ по любому из вариантов осуществления, 1-17, отличающийся тем, что аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, ТСО-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононаноевой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтирозина, три-О-ацетил-GlcNAc-серины, L-фосфосерины, фосфоносерины, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-

оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 20. Способ по варианту осуществления 19, отличающийся тем, что аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 21. Способ получения полипептида в эукариотической клетке, где полипептид содержит одну или несколько аминокислот не природного происхождения, где способ включает:

(а) получение эукариотической клетки, где эукариотическая клетка содержит:

(i) иРНК, содержащую кодон, где кодон содержит одно или несколько оснований не природного происхождения;

(ii) тРНК, содержащую антикодон, где антикодон содержит одно или несколько оснований не природного происхождения, и где одно или несколько оснований не природного происхождения, содержащих кодон в иРНК, и одно или несколько оснований не природного происхождения содержащих антикодон в тРНК, образуют комплементарную пару оснований; и

(iii) тРНК синтетазу, где тРНК синтетаза предпочтительно аминоацилирует тРНК с помощью одной или нескольких аминокислот не природного происхождения по сравнению с природной аминокислотой; и

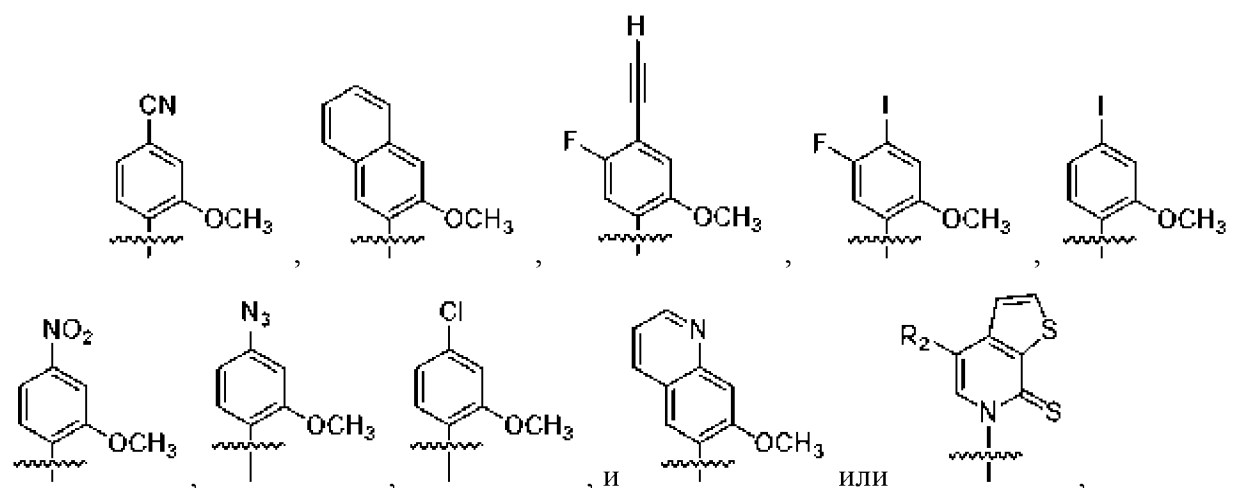
(b) предоставление еще одной аминокислоты не природного происхождения в эукариотическую клетку, где эукариотическая клетка продуцирует полипептид, содержащий одну или несколько аминокислот не природного происхождения.

Вариант осуществления 22. Способ по варианту осуществления 21, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

Вариант осуществления 23. Способ по варианту осуществления, 21, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

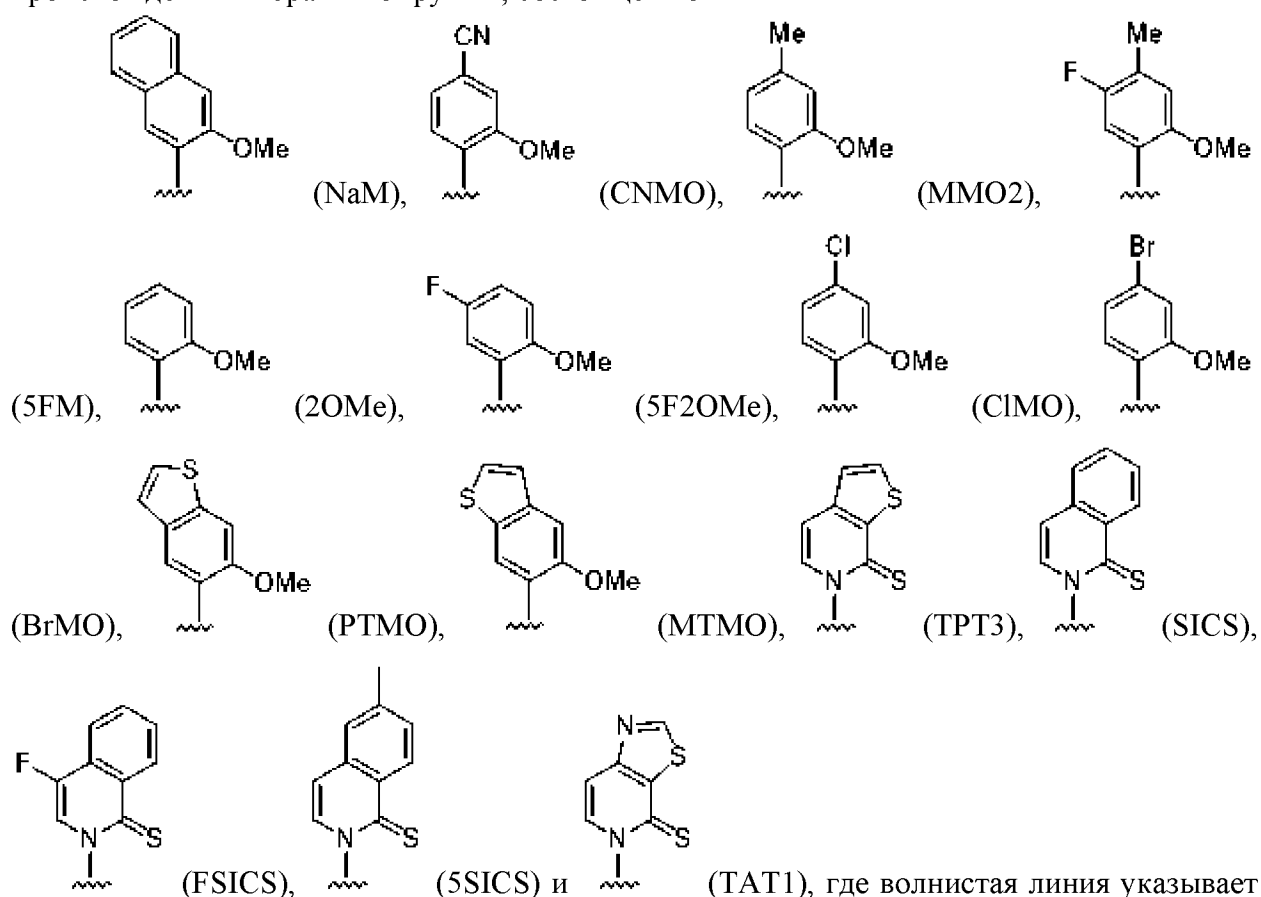
Вариант осуществления 24. Способ по варианту осуществления 21, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

Вариант осуществления 25. Способ по любому из вариантов осуществления, 21-24, отличающийся тем, что одно или несколько оснований не природного происхождения, содержащих кодон в иРНК, имеют формулу



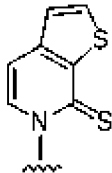
где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азидо, и волнистая линия указывает на связь с рибозильной группой.

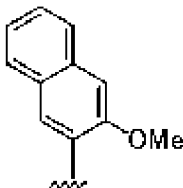
Вариант осуществления 26. Способ по любому из вариантов осуществления, 21-24, где первое основание не природного происхождения или второе основание не природного происхождения выбраны из группы, состоящей из

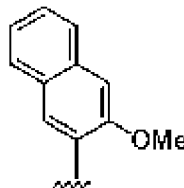


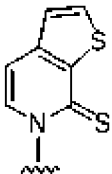
на связь с рибозильной группой.

Вариант осуществления 27. Способ по варианту осуществления 26, когда первое

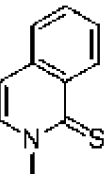
основание не природного происхождения представляет собой  (TPT3), второе

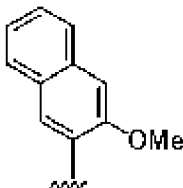
основание не природного происхождения представляет собой  (NaM), и когда

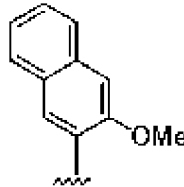
первое основание не природного происхождения представляет собой  (NaM),

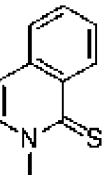
второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 28. Способ по варианту осуществления 26, когда первое

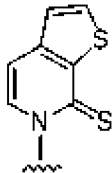
основание не природного происхождения представляет собой  (5SICS), второе

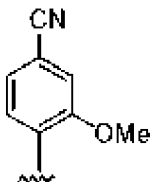
основание не природного происхождения представляет собой  (NaM), и когда

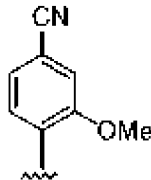
первое основание не природного происхождения представляет собой  (NaM),

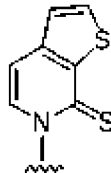
второе основание не природного происхождения представляет собой  (5SICS), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 29. Способ по варианту осуществления 26, когда первое

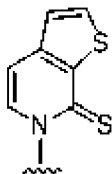
основание не природного происхождения представляет собой  (TPT3), второе

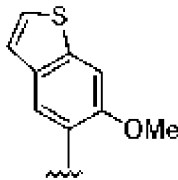
основание не природного происхождения представляет собой  (CNMO), и когда

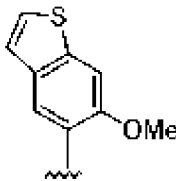
первое основание не природного происхождения представляет собой  (CNMO),

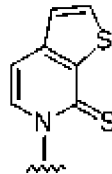
второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 30. Способ по варианту осуществления 26, когда первое

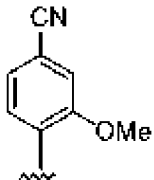
основание не природного происхождения представляет собой  (TPT3), второе

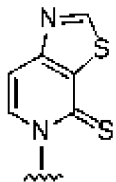
основание не природного происхождения представляет собой  (PTMO), и

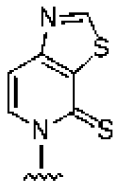
когда первое основание не природного происхождения представляет собой 

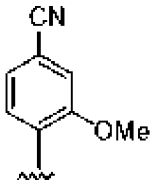
(PTMO), второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 31. Способ по варианту осуществления 26, когда первое

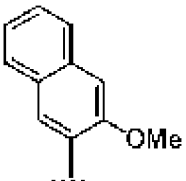
основание не природного происхождения представляет собой  (CNMO), второе

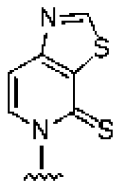
основание не природного происхождения представляет собой  (TAT1), и когда

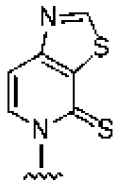
первое основание не природного происхождения представляет собой  (TAT1),

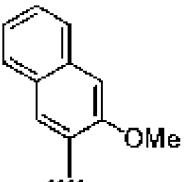
второе основание не природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 32. Способ по варианту осуществления 26, когда первое

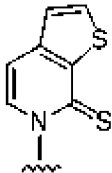
основание не природного происхождения представляет собой  (NaM), второе

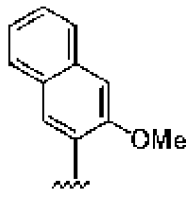
основание не природного происхождения представляет собой  (TAT1), и когда

первое основание не природного происхождения представляет собой  (TAT1),

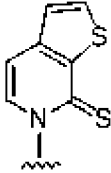
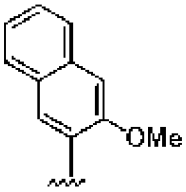
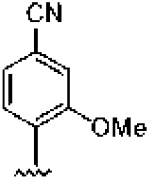
второе основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 33. Способ по варианту осуществления 26, когда первое

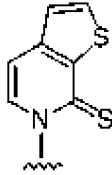
основание не природного происхождения представляет собой  (TPT3) и второе

основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

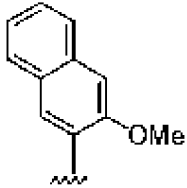
Вариант осуществления 34. Способ по любому из вариантов осуществления 21-24, отличающийся тем, что нуклеотид не природного происхождения, содержащий кодон в

иРНК, выбран из  (TPT3),  (NaM) и  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

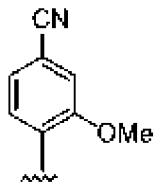
Вариант осуществления 35. Способ по варианту осуществления 34, где нуклеотид не

природного происхождения, содержащий кодон в иРНК, представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

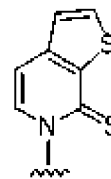
Вариант осуществления 36. Способ по варианту осуществления 34, отличающийся тем, что нуклеотид не природного происхождения, содержащий кодон в иРНК,

представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

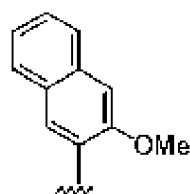
Вариант осуществления 37. Способ по варианту осуществления 34, отличающийся тем, что нуклеотид не природного происхождения, содержащий кодон в иРНК,

представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

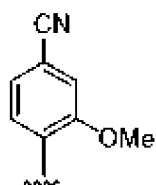
Вариант осуществления 38. Способ по варианту осуществления, 21, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание не природного происхождения (X) расположено в первом положении (X-N-N) в



(TPT3),



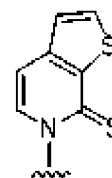
(NaM) и



(CNMO), где волнистая линия указывает на связь с

рибозильной группой.

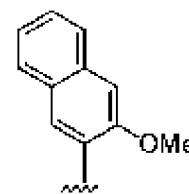
Вариант осуществления 39. Способ по варианту осуществления 38, в котором



(TPT3), где

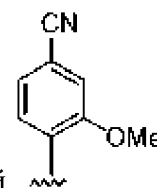
основание не природного происхождения представляет собой волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 40. Способ по варианту осуществления 38, отличающийся



тем, что основание не природного происхождения представляет собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

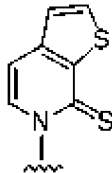
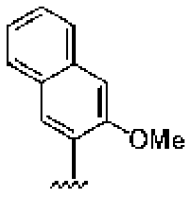
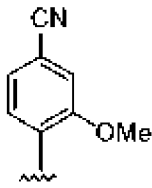
Вариант осуществления 41. Способ по варианту осуществления 38, отличающийся



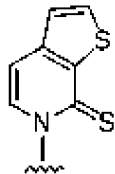
(CNMO),

где волнистая линия указывает на связь с рибозильной группой.

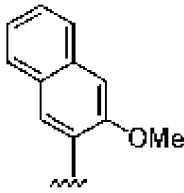
Вариант осуществления 42. Способ по варианту осуществления, 21, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание не природного происхождения (X) расположено в среднем положении (N-X-N)

в кодоне иРНК, где основание не природного происхождения выбрано из  (TPT3),  (NaM) и  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

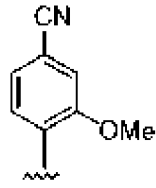
Вариант осуществления 43. Способ по варианту осуществления 42, отличающийся

тем, что основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

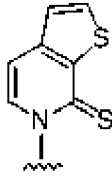
Вариант осуществления 44. Способ по варианту осуществления 42, отличающийся

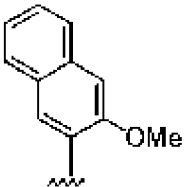
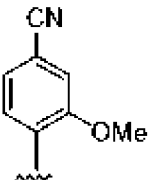
тем, что основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 45. Способ по варианту осуществления 42, отличающийся

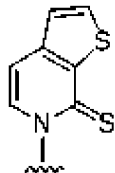
тем, что основание не природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 46. Способ по варианту осуществления 21, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где основание не природного происхождения (X) расположено в последнем положении (N-N-

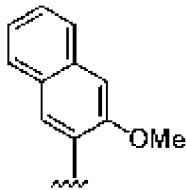
X) в кодоне иРНК, где основание не природного происхождения выбрано из 

(TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

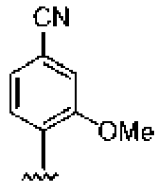
Вариант осуществления 47. Способ по варианту осуществления 46, отличающийся

тем, что основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 48. Способ по варианту осуществления 46, отличающийся

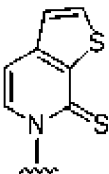
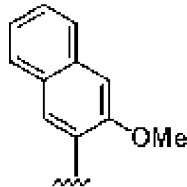
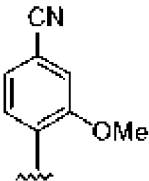
тем, что основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 49. Способ по варианту осуществления 46, отличающийся

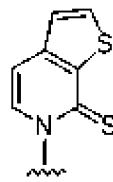
тем, что основание не природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 50. Способ по варианту осуществления 21, где антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в антикодоне тРНК.

Вариант осуществления 51. Способ по варианту осуществления 50, где основание не

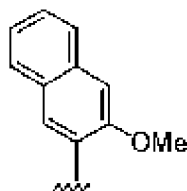
природного происхождения выбрано из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 52. Способ по варианту осуществления 51, где основание не



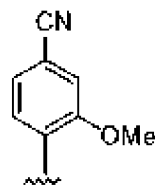
природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 53. Способ по варианту осуществления 51, где основание не



природного происхождения представляет собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

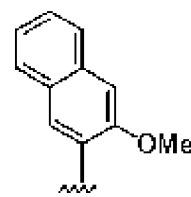
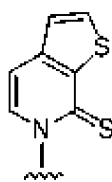
Вариант осуществления 54. Способ по варианту осуществления 51, где основание не



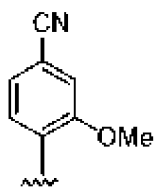
природного происхождения представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 55. Способ по варианту осуществления, 21, где антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в антикодоне тРНК.

Вариант осуществления 56. Способ по варианту осуществления 55, где основание не

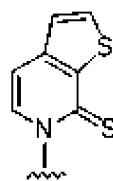


природного происхождения выбрано из (TPT3), (NaM) и



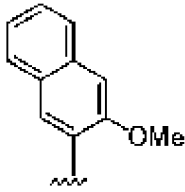
(CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 57. Способ по варианту осуществления 55, где основание не

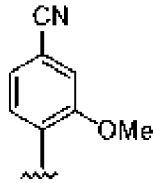


природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 58. Способ по варианту осуществления 55, где основание не

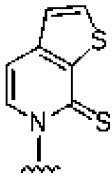
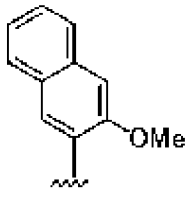
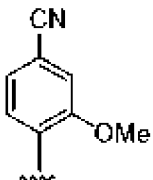
природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 59. Способ по варианту осуществления 55, где основание не

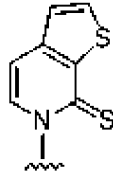
природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 60. Способ по варианту осуществления 21, где антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в антикодоне тРНК.

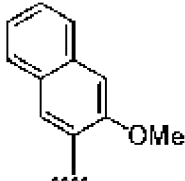
Вариант осуществления 61. Способ по варианту осуществления 60, где основание не

природного происхождения выбрано из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 62. Способ по варианту осуществления 61, где основание не

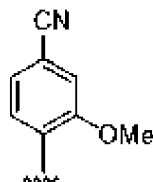
природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 63. Способ по варианту осуществления 61, где основание не

природного происхождения представляет собой  (NaM), где волнистая линия

указывает на связь с рибозильной группой.

Вариант осуществления 64. Способ по варианту осуществления 61, где основание не



природного происхождения представляет собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

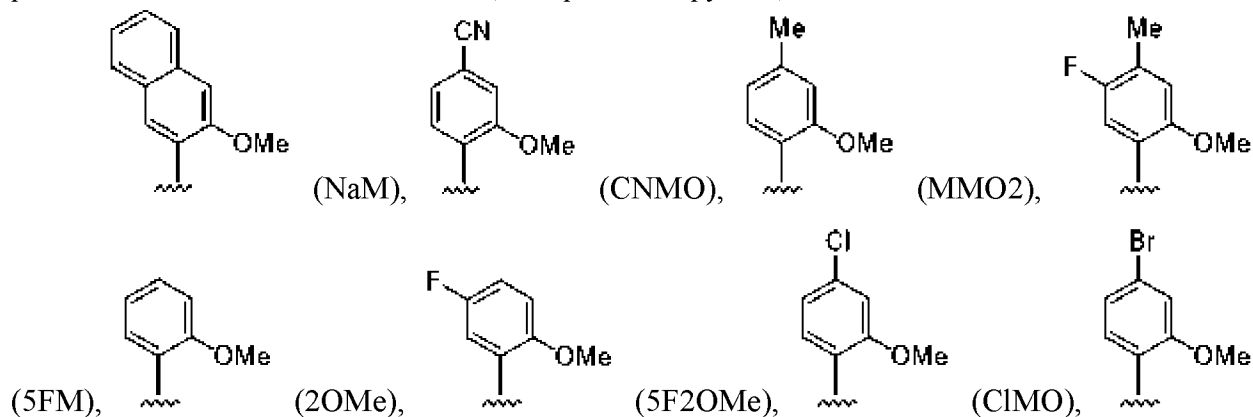
Вариант осуществления 65. Способ согласно варианту осуществления, 21, где кодон и антикодон каждый содержит три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в первом положении (X-N-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в последнем положении (N-N-Y) антикодона.

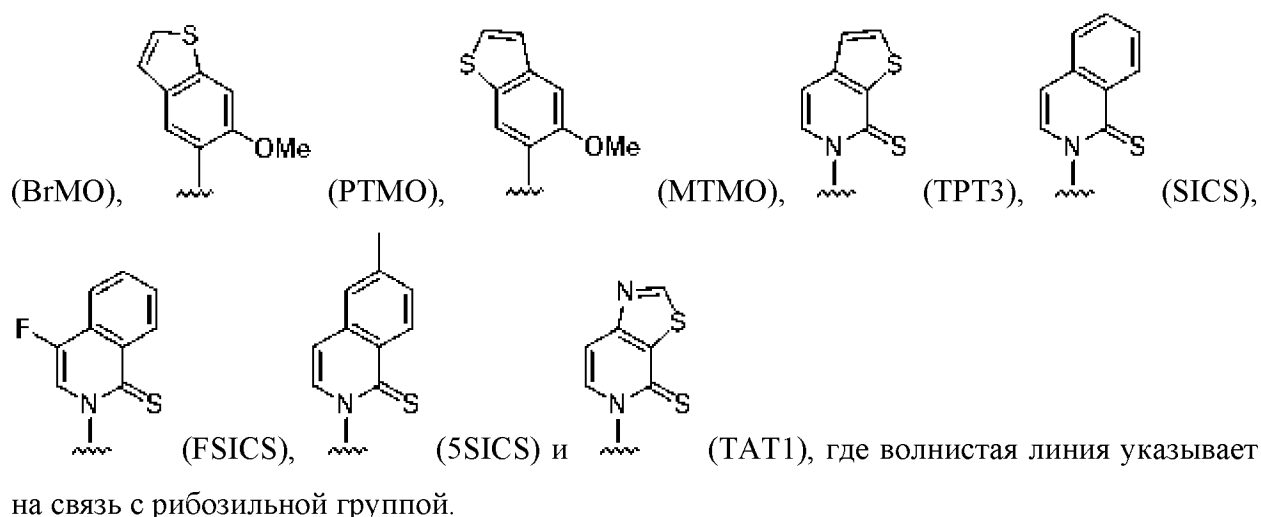
Вариант осуществления 66. Способ по варианту осуществления 65, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 67. Способ по варианту осуществления 66, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми.

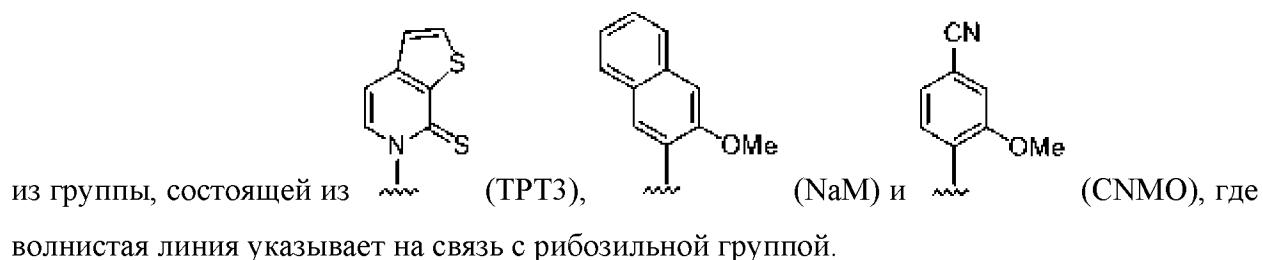
Вариант осуществления 68. Способ по варианту осуществления 66, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными.

Вариант осуществления 69. Способ по любому из вариантов осуществления 65-68, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбраны из группы, состоящей из

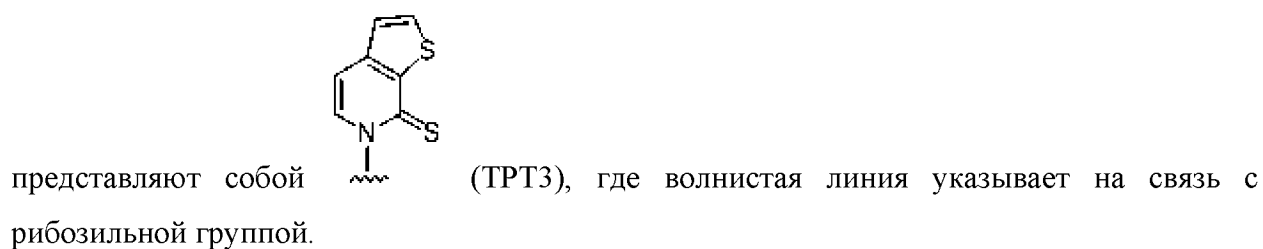




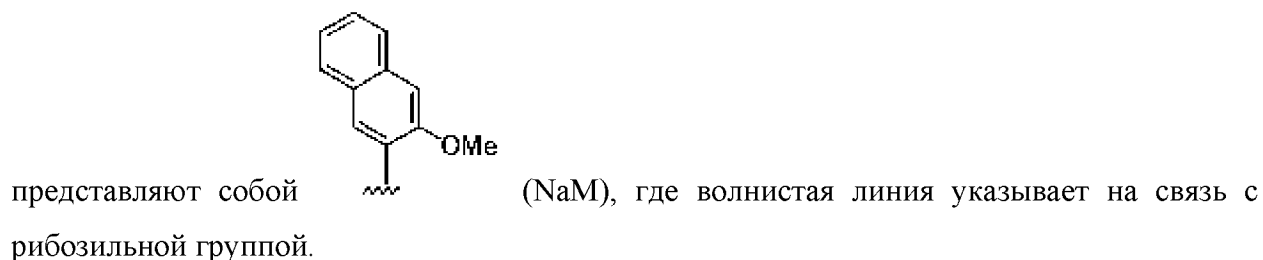
Вариант осуществления 70. Способ по варианту осуществления 69, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбраны



Вариант осуществления 71. Способ по варианту осуществления 70, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

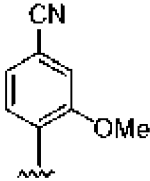


Вариант осуществления 72. Способ по варианту осуществления 70, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

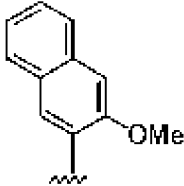
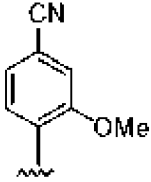


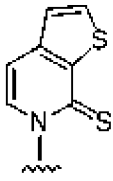
Вариант осуществления 73. Способ по варианту осуществления 70, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе

основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

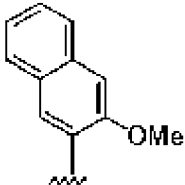
представляют собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 74. Способ по варианту осуществления 70, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, выбрано из

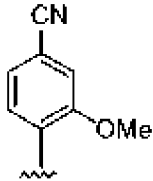
 (NaM) и  (CNMO), и второе основание не природного

происхождения (Y), расположенное в антикодоне тРНК, представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 75. Способ по варианту осуществления 74, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет

собой  (NaM).

Вариант осуществления 76. Способ по варианту осуществления 74, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет

собой  (CNMO).

Вариант осуществления 77. Способ по варианту осуществления, 21, где кодон и антикодон каждый содержит три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в среднем положении (N-X-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в среднем положении (N-Y-N) антикодона.

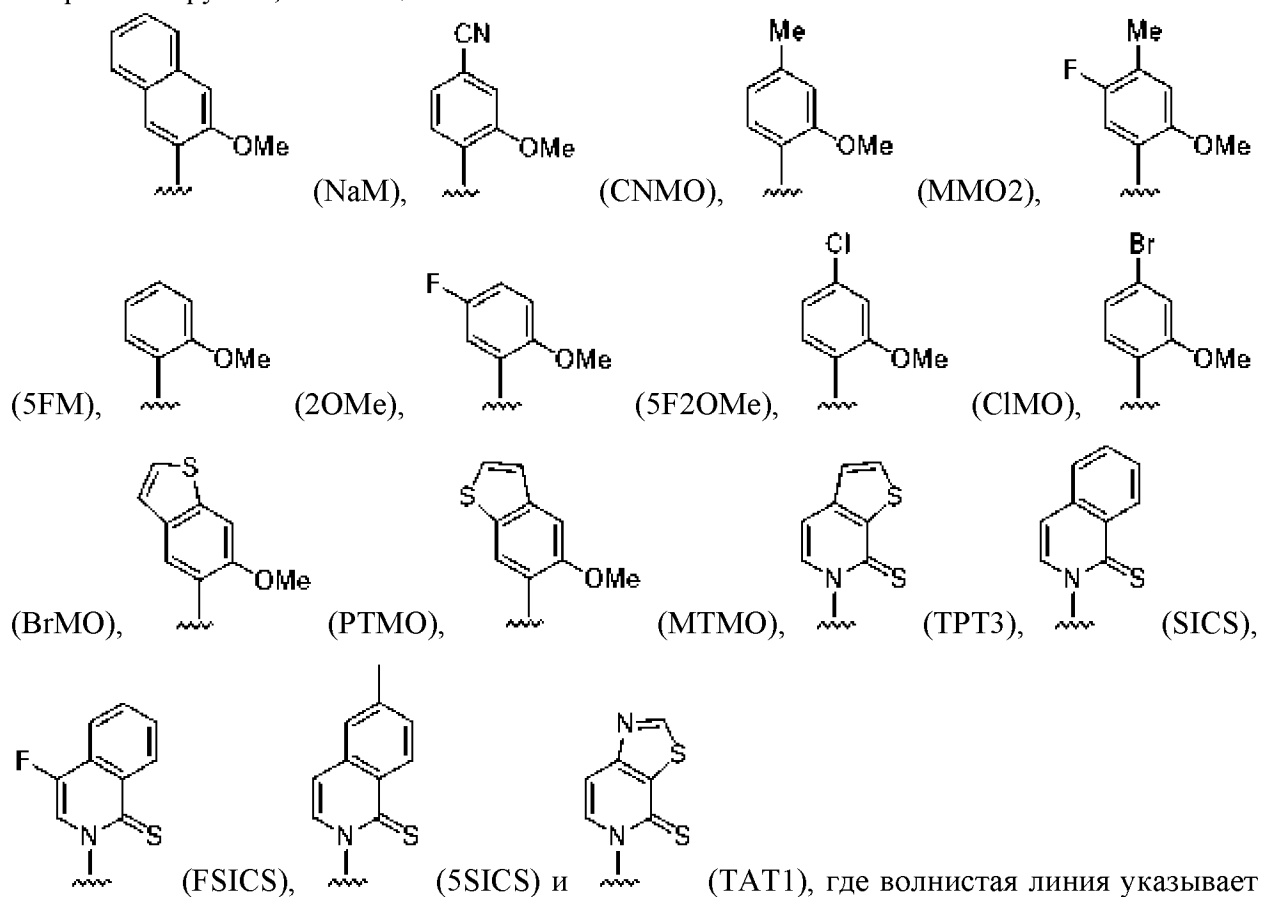
Вариант осуществления 78. Способ по варианту осуществления 77, где первое

основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными.

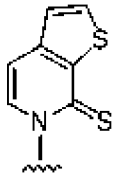
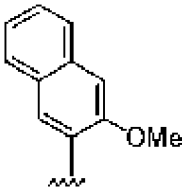
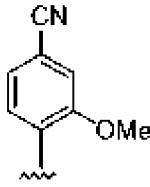
Вариант осуществления 79. Способ по варианту осуществления 78, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми.

Вариант осуществления 80. Способ по варианту осуществления 78, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными.

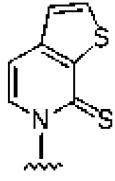
Вариант осуществления 81. Способ по любому из вариантов осуществления 77-79, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано из группы, состоящей из



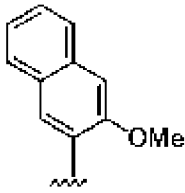
Вариант осуществления 82. Способ по варианту осуществления 81, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано

из группы, состоящей из  (TPT3),  (NaM) и  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

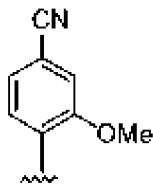
Вариант осуществления 83. Способ по варианту осуществления 82, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

представляют собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

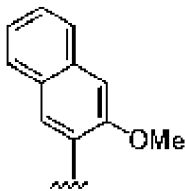
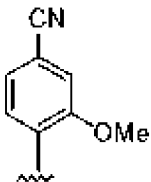
Вариант осуществления 84. Способ по варианту осуществления 82, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

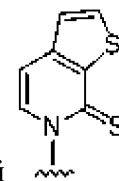
представляют собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 85. Способ по варианту осуществления 82, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

представляют собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

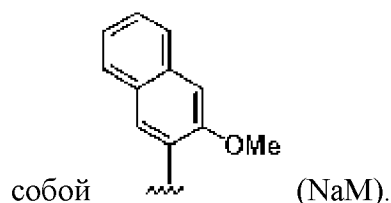
Вариант осуществления 86. Способ по варианту осуществления 82, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, выбрано из

 (NaM) и  (CNMO), и второе основание не природного

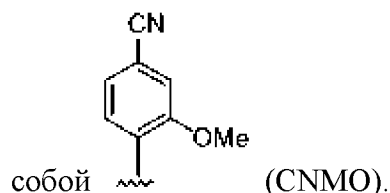


происхождения (Y), расположенное в антикодоне тРНК, представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 87. Способ по варианту осуществления 86, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет



Вариант осуществления 88. Способ по варианту осуществления 86, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет



Вариант осуществления 89. Способ согласно варианту осуществления, 21, где кодон и антикодон каждый содержит три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в последнем положении (N-N-X) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в первом положении (Y-N-N) антикодона.

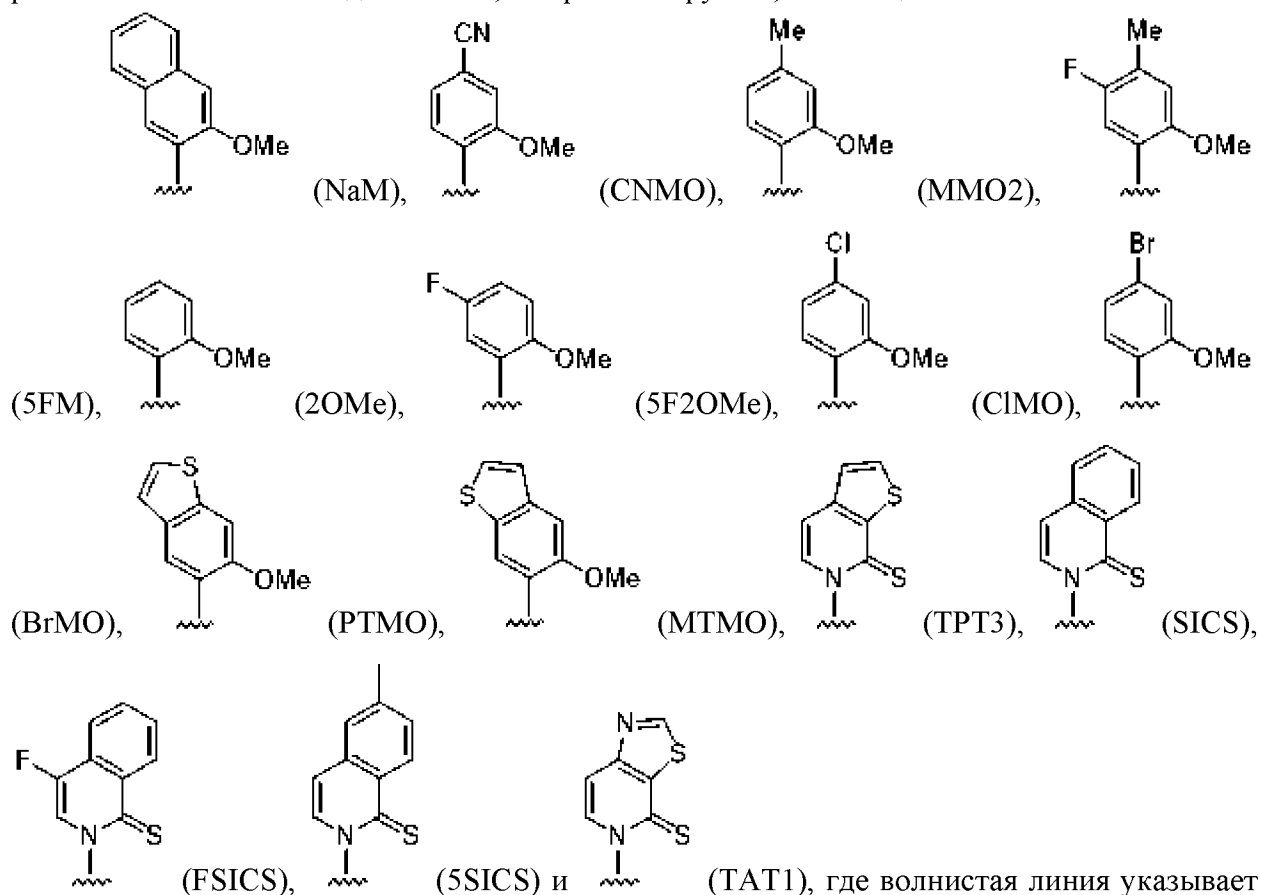
Вариант осуществления 90. Способ по варианту осуществления 89, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 91. Способ по варианту осуществления 89, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми.

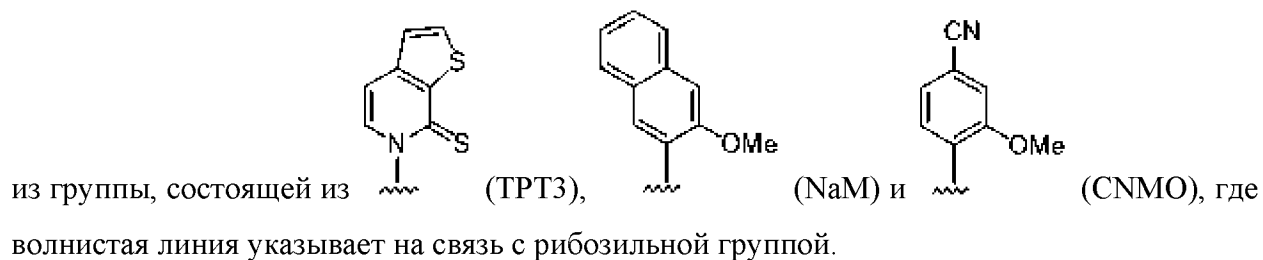
Вариант осуществления 92. Способ по варианту осуществления 89, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются разными.

Вариант осуществления 93. Способ по любому из вариантов осуществления 89-92, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y),

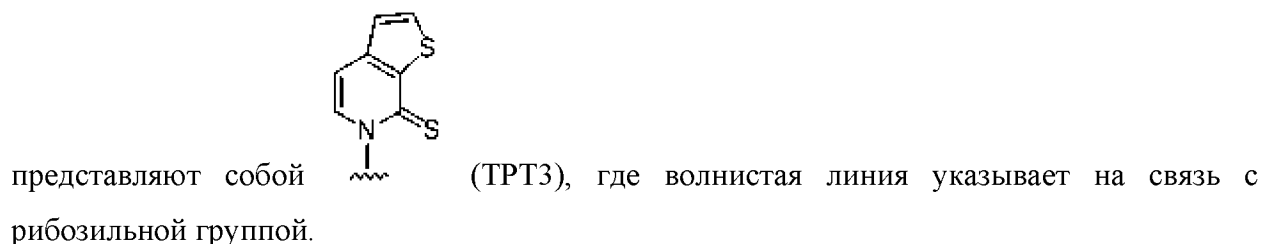
расположенное в антикодоне тРНК, выбрано из группы, состоящей из



Вариант осуществления 94. Способ по варианту осуществления 93, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано

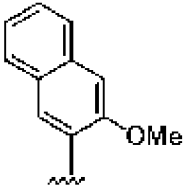


Вариант осуществления 95. Способ по варианту осуществления 94, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

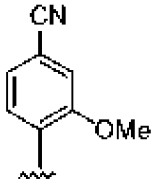


Вариант осуществления 96. Способ по варианту осуществления 94, где первое

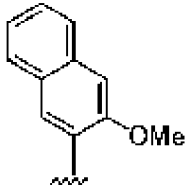
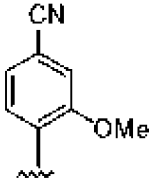
основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

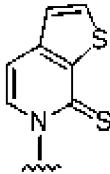
представляют собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 97. Способ по варианту осуществления 94, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба

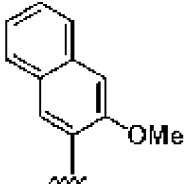
представляют собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 98. Способ по варианту осуществления 94, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, выбрано из

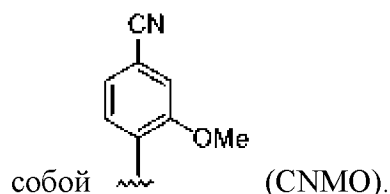
 (NaM) и  (CNMO), и второе основание не природного

происхождения (Y), расположенное в антикодоне тРНК, представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 99. Способ по варианту осуществления 98, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет

собой  (NaM).

Вариант осуществления 100. Способ по варианту осуществления 98, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, представляет



Вариант осуществления 101. Способ по любому из вариантов осуществления, 21, 23, 25-37, 42-45, 55-59 и 77-88, где кодон в иРНК выбран из АХС, GXC или GXU, где Х представляет собой основание не природного происхождения.

Вариант осуществления 102. Способ по варианту осуществления 101, где кодоном в иРНК является АХС, где Х представляет собой основание не природного происхождения.

Вариант осуществления 103. Способ по варианту осуществления 101, где кодон в иРНК представляет собой GXC, где Х представляет собой основание не природного происхождения.

Вариант осуществления 104. Способ по варианту осуществления 101, где кодон в иРНК представляет собой GXU, где Х представляет собой основание не природного происхождения.

Вариант осуществления 105. Способ по любому из вариантов осуществления, 21, 23, 25-37, 42-45, 55-59 и 77-88, где кодон в иРНК выбран из АХС, GXC или GXU, где антикодон в тРНК выбрано из GYU, GYC и AYC, где Х представляет собой первое основание не природного происхождения, и Y представляет собой второе основание не природного происхождения.

Вариант осуществления 106. Способ по варианту осуществления 105, где Х и Y являются одинаковыми или разными.

Вариант осуществления 107. Способ по варианту осуществления 106, где Х и Y являются одинаковыми.

Вариант осуществления 108. Способ по варианту осуществления, 106, где Х и Y являются разными.

Вариант осуществления 109. Способ по варианту осуществления 105, где кодон в иРНК представляет собой АХС, и антикодон в тРНК представляет собой GYU.

Вариант осуществления 110. Способ по варианту осуществления 109, где Х и Y являются одинаковыми или разными.

Вариант осуществления 111. Способ по варианту осуществления 109, где Х и Y являются одинаковыми.

Вариант осуществления 112. Способ по варианту осуществления 109, где Х и Y являются разными.

Вариант осуществления 113. Способ по варианту осуществления 106, где кодон в иРНК представляет собой GXC, и антикодон в тРНК представляет собой GYC.

Вариант осуществления, 114. Способ по варианту осуществления 113, где Х и Y являются одинаковыми или разными.

Вариант осуществления 115. Способ по варианту осуществления 113, где Х и Y являются одинаковыми.

Вариант осуществления 116. Способ по варианту осуществления 113, где X и Y являются разными.

Вариант осуществления 117. Способ по варианту осуществления 106, где кодон в иРНК представляет собой GXU, и антикодон представляет собой AYC.

Вариант осуществления, 118. Способ по варианту осуществления 117, где X и Y являются одинаковыми или разными.

Вариант осуществления 119. Способ по варианту осуществления 117, где X и Y являются одинаковыми.

Вариант осуществления 120. Способ по варианту осуществления 117, где X и Y являются разными.

Вариант осуществления 121. Способ по любому из вариантов осуществления, 21-120, где тРНК получают из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 122. Способ по любому из вариантов по варианту осуществления 21-120, где тРНК-синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 123. Способ по варианту осуществления 122, где тРНК и тРНК синтетазу получают из *Methanococcus jannaschii*.

Вариант осуществления 124. Способ по варианту осуществления 122, где тРНК и тРНК синтетазу получают из *Methanosarcina barkeri*.

Вариант осуществления 125. Способ по варианту осуществления 122, где тРНК и тРНК синтетазу получают из *Methanosarcina mazei*.

Вариант осуществления 126. Способ по варианту осуществления 122, где тРНК и тРНК синтетаза получают из *Methanosarcina acetivorans*.

Вариант осуществления 127. Способ по любому из вариантов осуществления 21-120, где тРНК получают из *Methanococcus jannaschii*, и тРНК-синтетаза получают из *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 128. Способ по любому из вариантов осуществления 21-120, где тРНК получают из *Methanosarcina barkeri*, и тРНК-синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 129. Способ по любому из вариантов осуществления 21-120, где тРНК получают из *Methanosarcina mazei*, и тРНК-синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina acetivorans*.

Вариант осуществления 130. Способ по любому из вариантов осуществления 21-120, где тРНК получают из *Methanosarcina acetivorans*, и тРНК-синтетазу получают из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina mazei*.

Вариант осуществления 131. Способ по любому из вариантов осуществления 21-120, где тРНК получают из *Methanosarcina mazei*, и тРНК-синтетаза получают из *Methanosarcina barkeri*.

Вариант осуществления 132. Способ по любому из вариантов осуществления 21-120,

где клетка представляет собой клетку человека.

Вариант осуществления 133. Способ по варианту осуществления 132, где клетка человека представляет собой клетку НЕК293Т.

Вариант осуществления 134. Способ по любому из вариантов осуществления 21-120, где клетка представляет собой клетку хомяка.

Вариант осуществления, 135. Способ согласно варианту осуществления 134, где клетка хомяка представляет собой клетку яичника китайского хомяка (СНО).

Вариант осуществления 136. Способ по любому из вариантов осуществления 21-135, где аминокислота не природного происхождения:

представляет собой аналог лизина;
содержит ароматическую боковую цепь;
содержит азидогруппу;
содержит алкиновую группу; или
содержат альдегидную или кетонную группу.

Вариант осуществления 137. Способ по любому из вариантов осуществления 21-135, где аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидозтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, ТСО-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилфенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацетил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-(((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 138. Способ по варианту осуществления 137, где аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 139. Система экспрессии полипептида не встречающегося в природе в эукариотической клетке, содержащая:

- (а) по меньшей мере, одну аминокислоту не природного происхождения;
- (б) иРНК, кодирующую полипептид не встречающийся в природе, где указанная иРНК содержит, по меньшей мере, один кодон, содержащий одно или несколько первых оснований не природного происхождения;

(с) тРНК, содержащую, по меньшей мере, один антикодон, содержащий одно или несколько вторых оснований не природного происхождения, где одно или несколько первых оснований не природного происхождения и одно или несколько вторых оснований не природного происхождения образуют одну или несколько комплементарных пар оснований;

(d) одну или несколько конструкций нуклеиновой кислоты, содержащих последовательность нуклеиновой кислоты, кодирующую тРНК-синтетазу, где тРНК-синтетаза предпочтительно аминоацилирует тРНК, по меньшей мере, с одной аминокислотой не природного происхождения; и

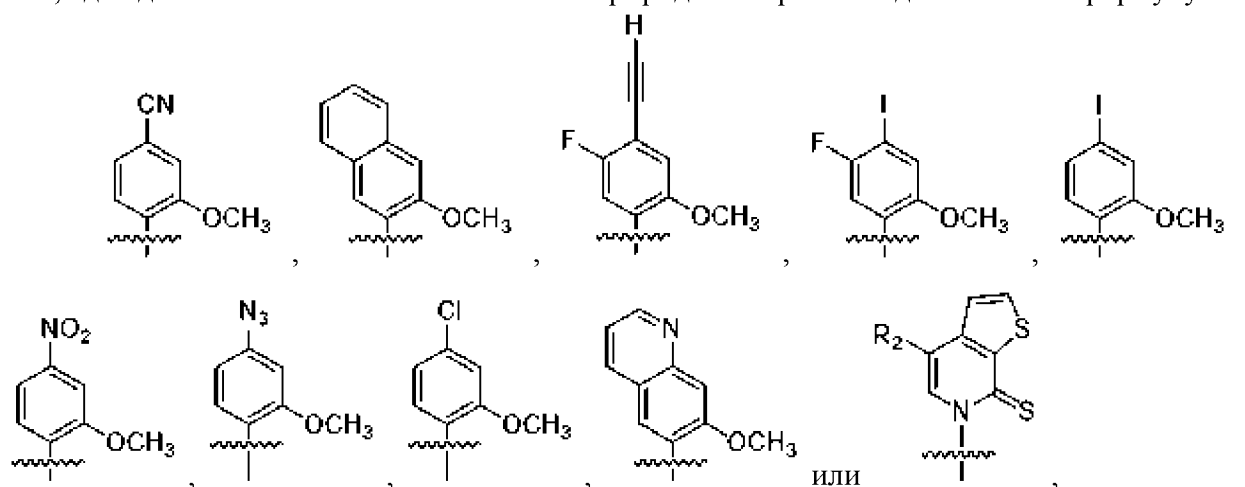
(е) эукариотическую клетку, способную транслировать иРНК в полипептид, содержащий аминокислоту не природного происхождения, с использованием тРНК и тРНК синтетазы.

Вариант осуществления 140. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в, по меньшей мере, одном кодоне иРНК.

Вариант осуществления 141. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК.

Вариант осуществления 142. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в, по меньшей мере, одном кодоне иРНК.

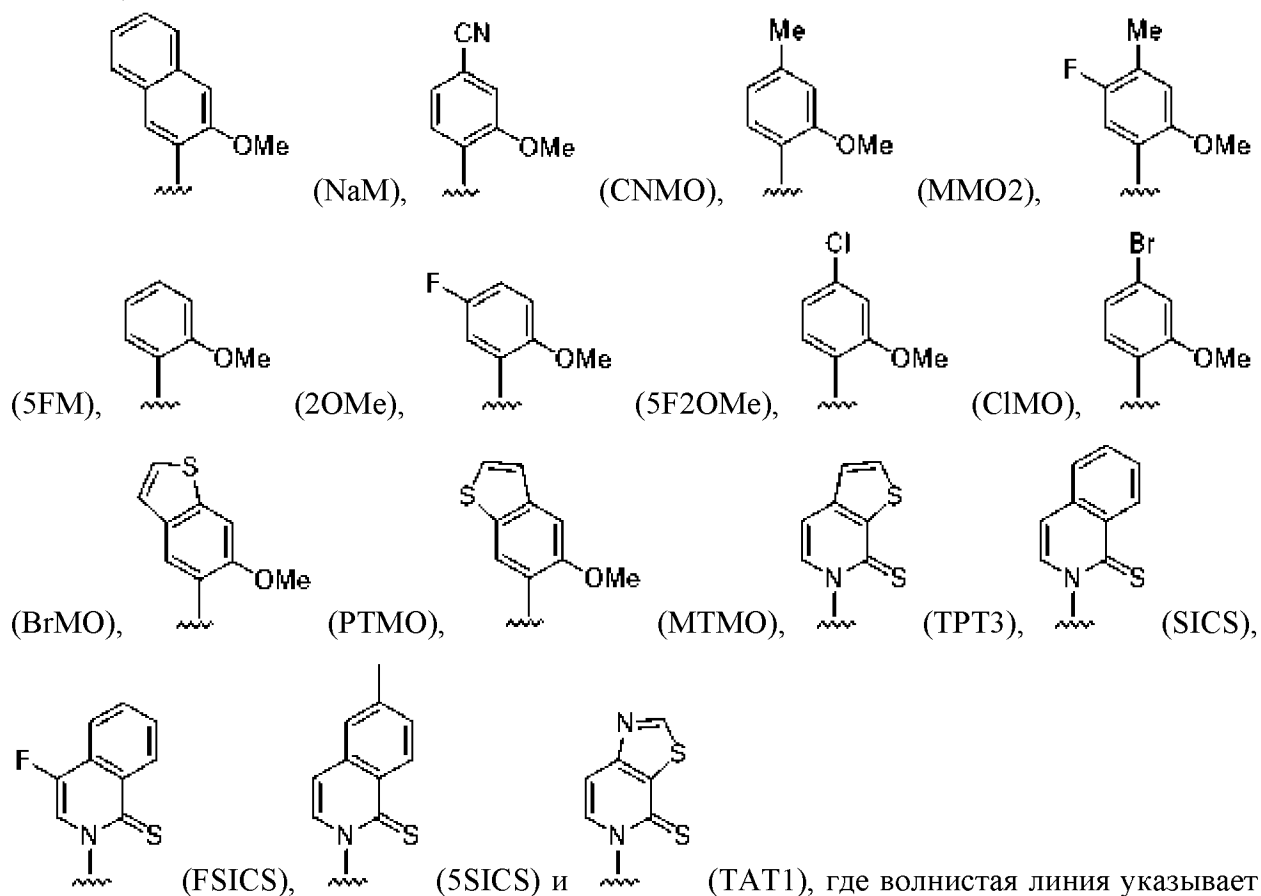
Вариант осуществления 143. Система по любому из вариантов осуществления 139-142, где одно или несколько оснований не природного происхождения имеют формулу



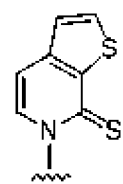
где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азидо, и волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 144. Система по любому из вариантов осуществления 139-

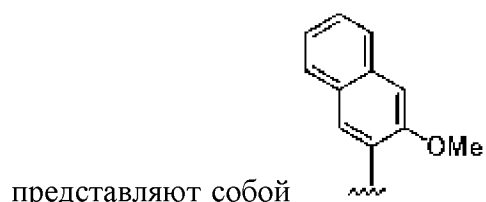
142, где одно или несколько первых оснований не природного происхождения или одно или несколько вторых оснований не природного происхождения выбраны из группы, состоящей из

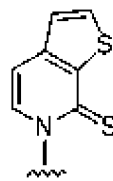


Вариант осуществления 145. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой



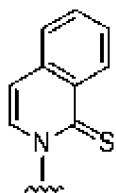
(TPT3), одно или несколько вторых оснований не природного происхождения



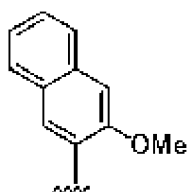


не природного происхождения представляет собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

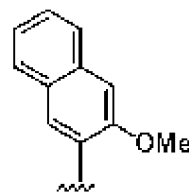
Вариант осуществления 146. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой



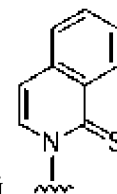
(5SICS), одно или несколько вторых оснований не природного происхождения



представляют собой (NaM), и когда одно или несколько первых оснований

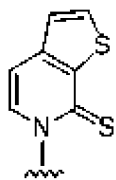


не природного происхождения представляют собой (NaM), одно или

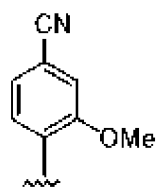


несколько вторых оснований не природного происхождения представляют собой (5SICS), где волнистая линия указывает на связь с рибозильной группой.

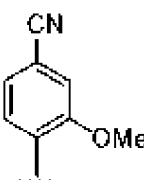
Вариант осуществления 147. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой

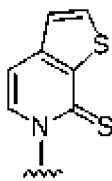


(TPT3), одно или несколько вторых оснований не природного происхождения

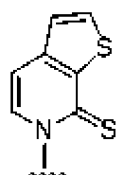


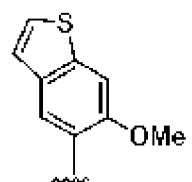
представляют собой (CNMO), и когда одно или несколько первых оснований

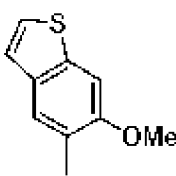
не природного происхождения представляют собой  (CNMO), одно или

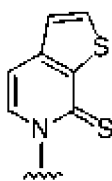
несколько вторых оснований не природного происхождения представляют собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 148. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой

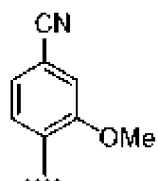
 (TPT3), одно или несколько вторых оснований не природного происхождения

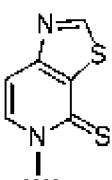
представляют собой  (PTMO), и когда одно или несколько первых оснований

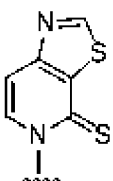
не природного происхождения представляют собой  (PTMO), одно или

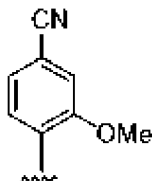
несколько вторых оснований не природного происхождения представляют собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 149. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой

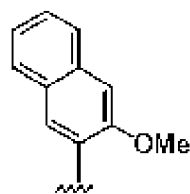
 (CNMO), одно или несколько вторых оснований не природного происхождения

представляют собой  (TAT1), и когда одно или несколько первых оснований не

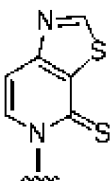
природного происхождения представляют собой  (TAT1), одно или несколько

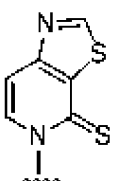
вторых оснований не природного происхождения представляют собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

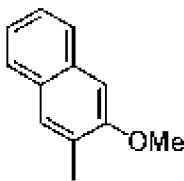
Вариант осуществления 150. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой



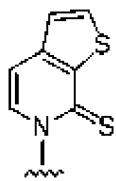
(NaM), одно или несколько вторых оснований не природного происхождения

представляют собой  (TAT1), и когда одно или несколько первых оснований не

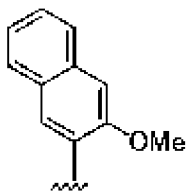
природного происхождения представляют собой  (TAT1), одно или несколько

вторых оснований не природного происхождения представляют собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 151. Система по варианту осуществления 144, когда одно или несколько первых оснований не природного происхождения представляют собой

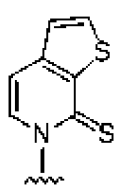


(TPT3) и одно или несколько вторых оснований не природного происхождения

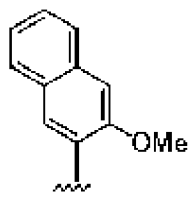


представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

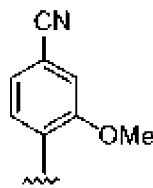
Вариант осуществления 152. Система по любому из вариантов осуществления 139-142, где одно или несколько первых оснований не природного происхождения выбраны из



(TPT3),



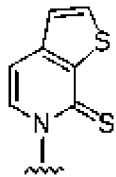
(NaM) и



(CNMO), где волнистая линия

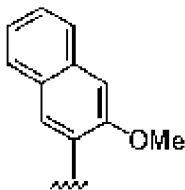
указывает на связь с рибозильной группой.

Вариант осуществления 153. Система по варианту осуществления 152, в которой одно или несколько первых оснований не природного происхождения представляют собой



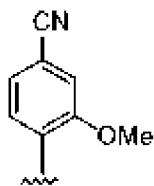
(TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 154. Система по варианту осуществления 152, где одно или несколько первых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

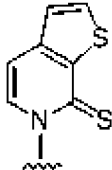
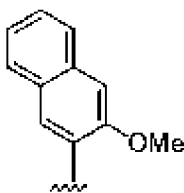
Вариант осуществления 155. Система по варианту осуществления 152, где одно или несколько первых оснований не природного происхождения представляют собой

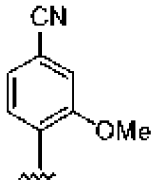


(CNMO), где волнистая линия указывает на связь с рибозильной группой.

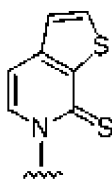
Вариант осуществления 156. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-

N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в кодоне иРНК, где одно или несколько первых

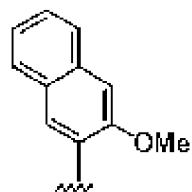
оснований не природного происхождения выбраны из  (TPT3), 

(NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

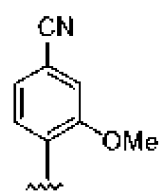
Вариант осуществления 157. Система по варианту осуществления 156, где одно или

несколько первых оснований не природного происхождения представляют собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 158. Система по варианту осуществления 156, где одно или несколько первых оснований не природного происхождения представляют собой

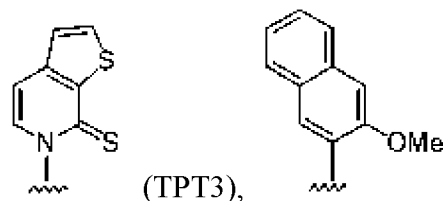
 (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 159. Система по варианту осуществления 156, где одно или несколько первых оснований не природного происхождения представляют собой

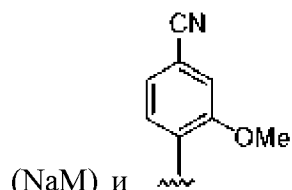
 (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 160. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК, где одно или несколько первых

оснований не природного происхождения выбраны из



(TPT3),

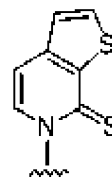


(NaM) и

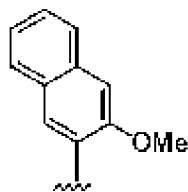
(CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 161. Система по варианту осуществления 160, где одно или

несколько первых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

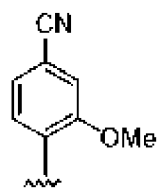


Вариант осуществления 162. Система по варианту осуществления 160, где одно или несколько первых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

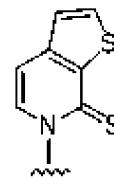
Вариант осуществления 163. Система по варианту осуществления 160, где одно или несколько первых оснований не природного происхождения представляют собой



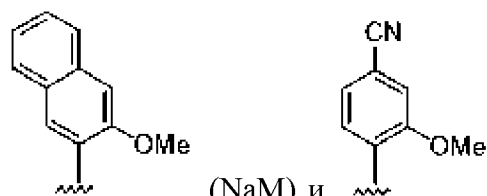
(CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 164. Система по варианту осуществления 139, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в кодоне иРНК, где одно или несколько

первых оснований не природного происхождения выбраны из



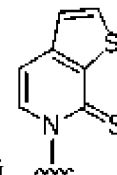
(TPT3),



(NaM) и (CNMO), и где волнистая линия указывает на связь с

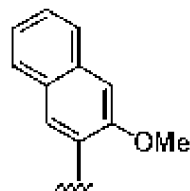
рибозильной группой.

Вариант осуществления 165. Система по варианту осуществления 164, где одно или



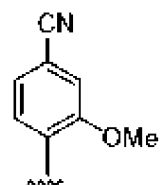
несколько первых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 166. Система по варианту осуществления 164, где одно или несколько первых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

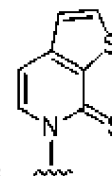
Вариант осуществления 167. Система по варианту осуществления 164, где одно или несколько первых оснований не природного происхождения представляют собой



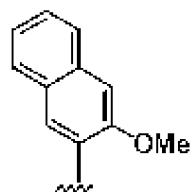
(CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 168. Система по варианту осуществления 139, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в антикодоне тРНК.

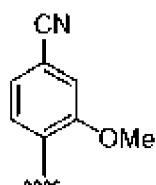
Вариант осуществления 169. Система по варианту осуществления 168, где одно или



несколько вторых оснований не природного происхождения выбраны из (TPT3),



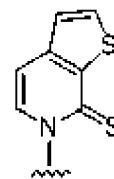
(NaM) и



(CNMO), и где волнистая линия указывает на связь с

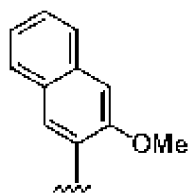
рибозильной группой.

Вариант осуществления 170. Система по варианту осуществления 168, где одно или



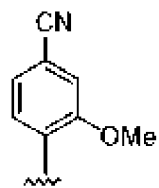
несколько вторых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 171. Система по варианту осуществления 168, где одно или несколько вторых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

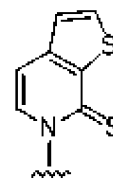
Вариант осуществления 172. Система по варианту осуществления 168, где одно или несколько вторых оснований не природного происхождения представляют собой



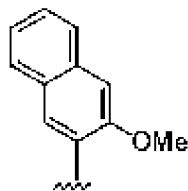
(CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 173. Система по варианту осуществления 139, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в антикодоне тРНК.

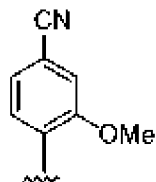
Вариант осуществления 174. Система по варианту осуществления 173, где одно или



несколько вторых оснований не природного происхождения выбраны из (TPT3),



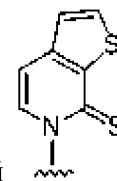
(NaM) и



(CNMO), и где волнистая линия указывает на связь с

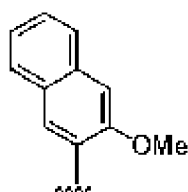
рибозильной группой.

Вариант осуществления 175. Система по варианту осуществления 173, где одно или



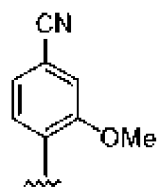
несколько вторых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 176. Система по варианту осуществления 173, где одно или несколько вторых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

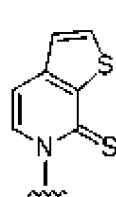
Вариант осуществления 177. Система по варианту осуществления 173, где одно или несколько вторых оснований не природного происхождения представляют собой



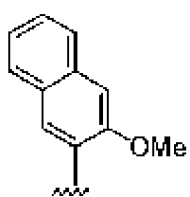
(CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 178. Система по варианту осуществления 139, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в антикодоне тРНК.

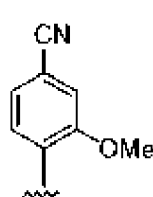
Вариант осуществления 179. Система согласно варианту осуществления 178, в которой одно или несколько вторых оснований не природного происхождения выбраны из



(TPT3),



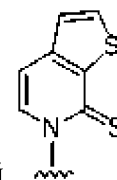
(NaM) и



(CNMO), и где волнистая линия

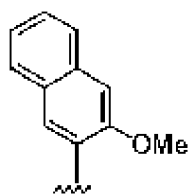
указывает на связь с рибозильной группой.

Вариант осуществления 180. Система по варианту осуществления 178, где одно или



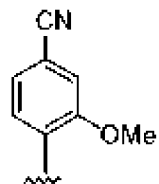
несколько вторых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 181. Система по варианту осуществления 178, где одно или несколько вторых оснований не природного происхождения представляют собой



(NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 182. Система по варианту осуществления 178, где одно или несколько вторых оснований не природного происхождения представляют собой



(CNMO), где волнистая линия указывает на связь с рибозильной группой.

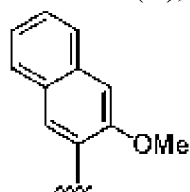
Вариант осуществления 183. Система по варианту осуществления 139, где, по меньшей мере, один кодон и по меньшей мере, один антикодон каждый независимо содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон содержит одно или несколько первых оснований не природного происхождения (X), расположенных в первом положении (X-N-N) кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в последнем положении (N-N-Y) антикодона.

Вариант осуществления 184. Система по варианту осуществления 183, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

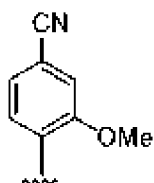
Вариант осуществления 185. Система по варианту осуществления 184, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми.

Вариант осуществления 186. Система по варианту осуществления 184, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными.

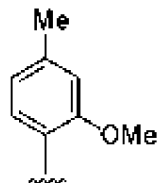
Вариант осуществления 187. Система по любому из вариантов осуществления 183-186, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбрано из группы, состоящей из



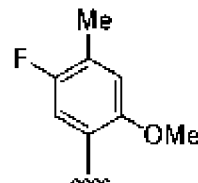
(NaM),

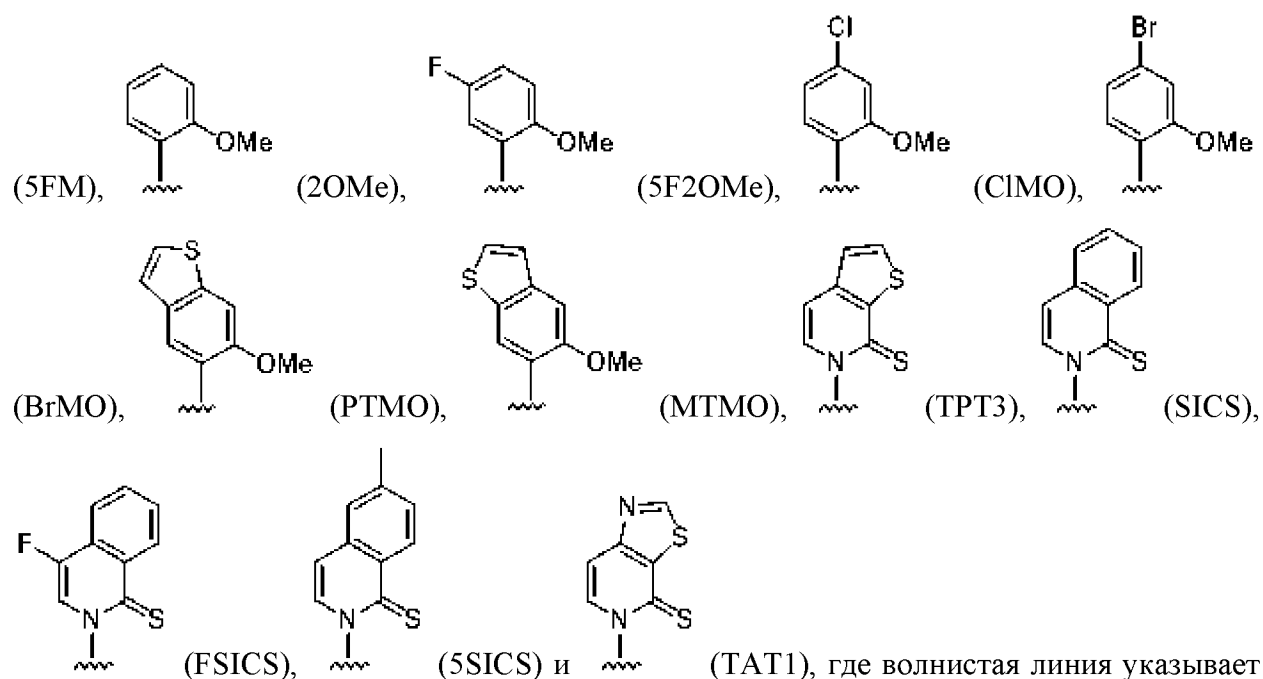


(CNMO),

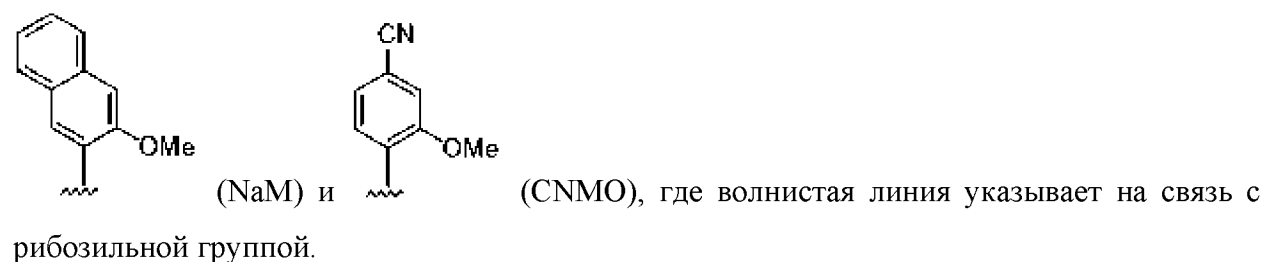
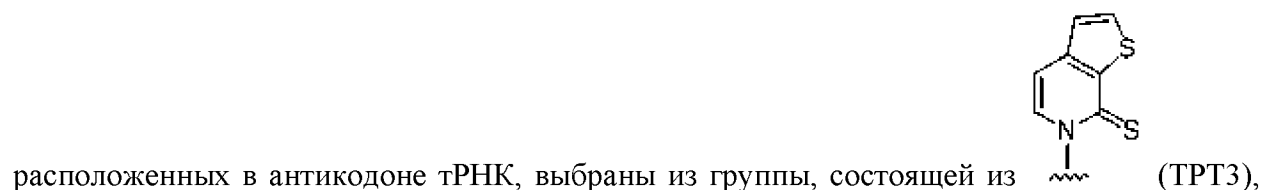


(MMO2),

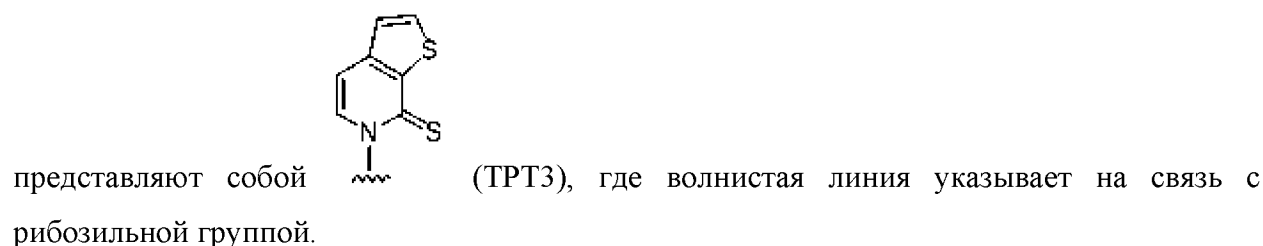




Вариант осуществления 188. Система по варианту осуществления 187, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

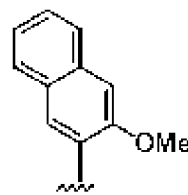


Вариант осуществления 189. Система по варианту осуществления 188, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, оба



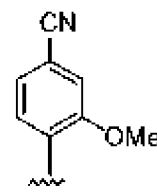
Вариант осуществления 190. Система по варианту осуществления 188, где одно или

несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



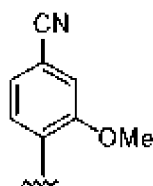
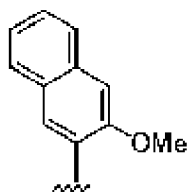
расположенных в антикодоне тРНК, оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 191. Система по варианту осуществления 188, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

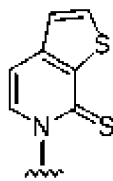


расположенных в антикодоне тРНК, оба представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 192. Система по варианту осуществления 188, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне

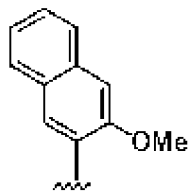


иРНК, выбраны из (NaM) и (CNMO), и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК



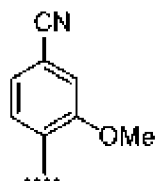
представляет собой , где волнистая линия в каждом случае указывает на связь с рибозильной группой.

Вариант осуществления 193. Система по варианту осуществления 192, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (NaM).

Вариант осуществления 194. Система по варианту осуществления 192, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (CNMO).

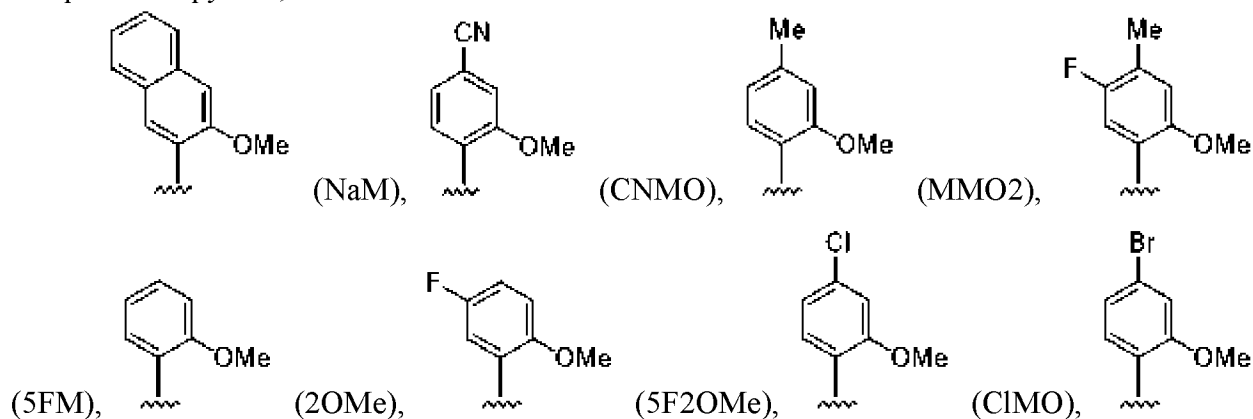
Вариант осуществления 195. Система согласно варианту осуществления 139, где, по меньшей мере, один кодон и, по меньшей мере, один антикодон, каждый независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в среднем положении (N-X-N), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в среднем положении (N-Y-N) антикодона.

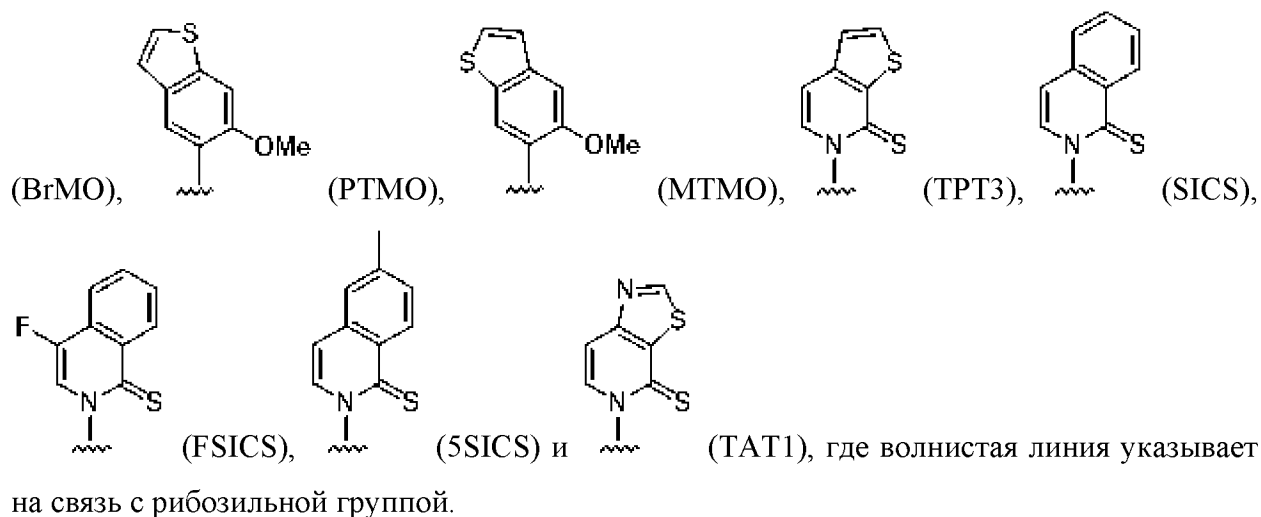
Вариант осуществления 196. Система по варианту осуществления 195, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 197. Система по варианту осуществления 195, в которой одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми.

Вариант осуществления 198. Система по варианту осуществления 195, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными.

Вариант осуществления 199. Система по любому из вариантов осуществления 195-198, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

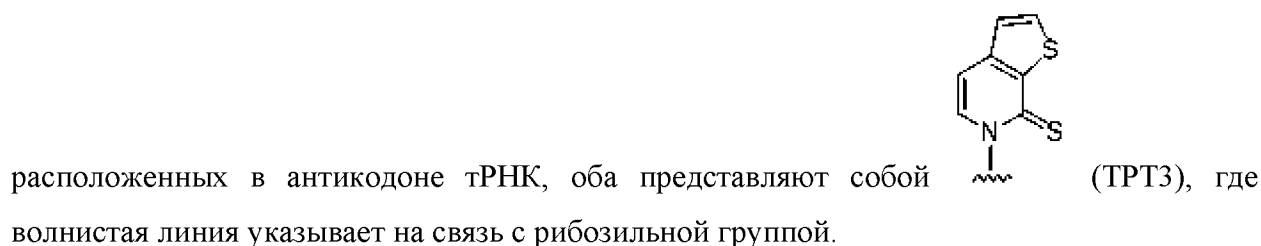




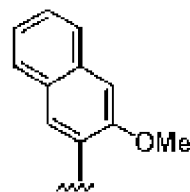
Вариант осуществления 200. Система по варианту осуществления 199, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



Вариант осуществления 201. Система по варианту осуществления 200, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

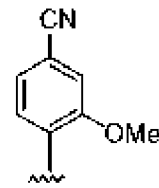


Вариант осуществления 202. Система по варианту осуществления 200, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



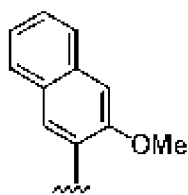
расположенных в антикодоне тРНК, оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 203. Система по варианту осуществления 200, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

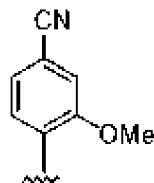


расположенных в антикодоне тРНК, оба представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 204. Система по варианту осуществления 200, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне

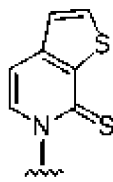


(NaM) и



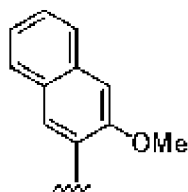
(CNMO), и одно или несколько

вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК



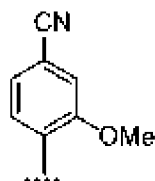
представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 205. Система по варианту осуществления 204, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (NaM).

Вариант осуществления 206. Система по варианту осуществления 204, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (CNMO).

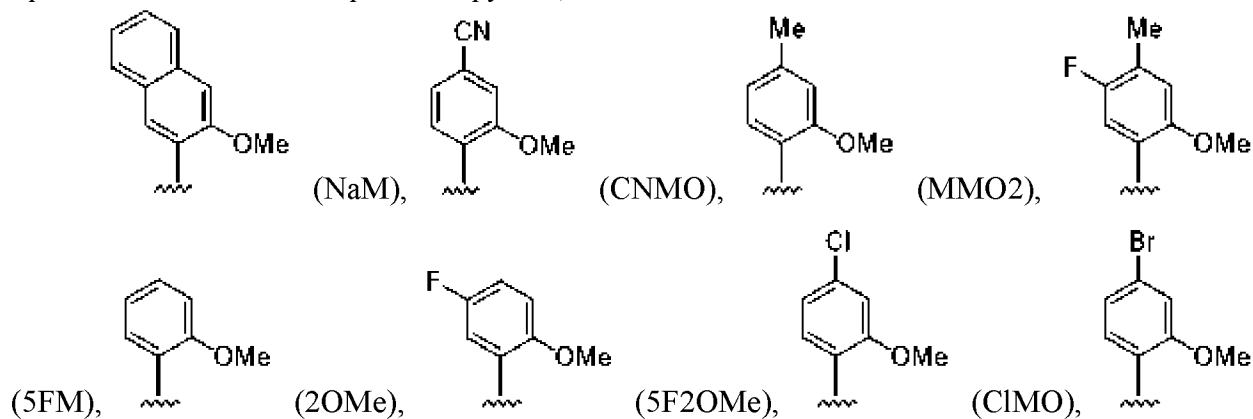
Вариант осуществления 207. Система по варианту осуществления 139, где, по меньшей мере, один кодон и, по меньшей мере, один антикодон, каждый независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в последнем положении (N-N-X), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в первом положении (Y-N-N) антикодона.

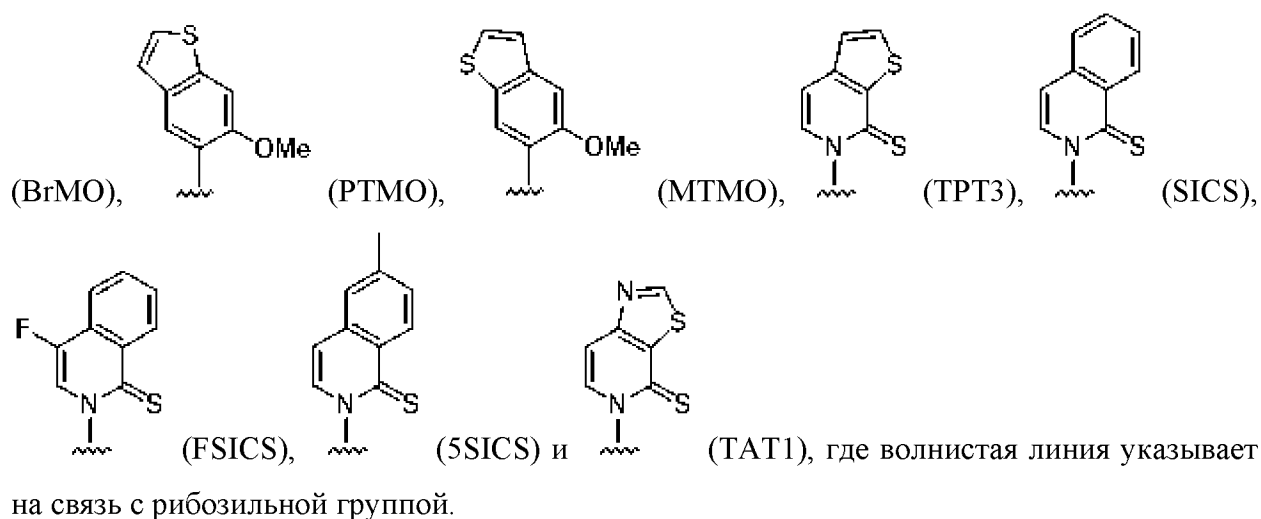
Вариант осуществления 208. Система по варианту осуществления 207, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 209. Система по варианту осуществления 208, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми.

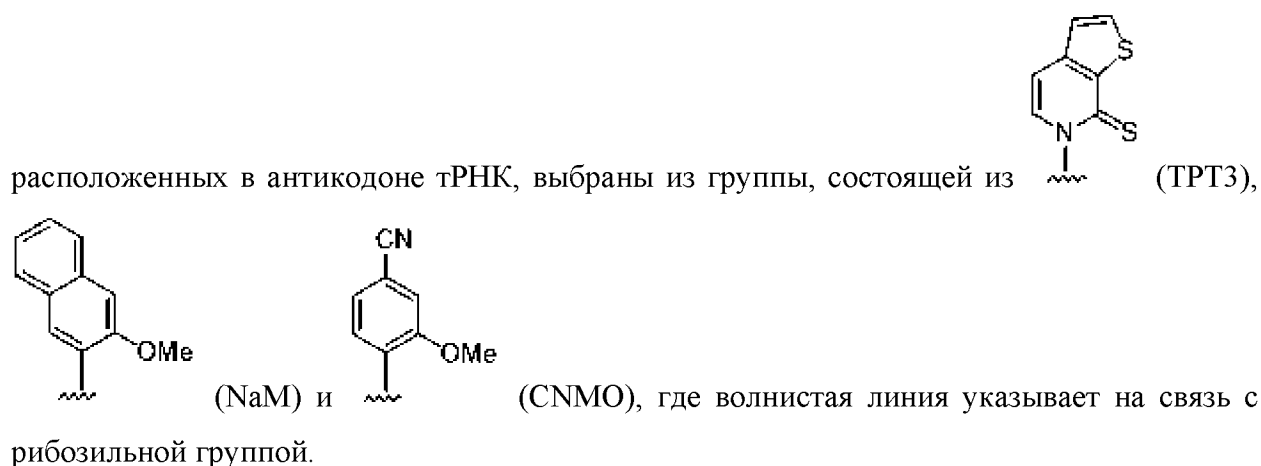
Вариант осуществления 210. Система по варианту осуществления 208, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются разными.

Вариант осуществления 211. Система по любому из вариантов осуществления 207-210, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, представляют собой выбрано из группы, состоящей из

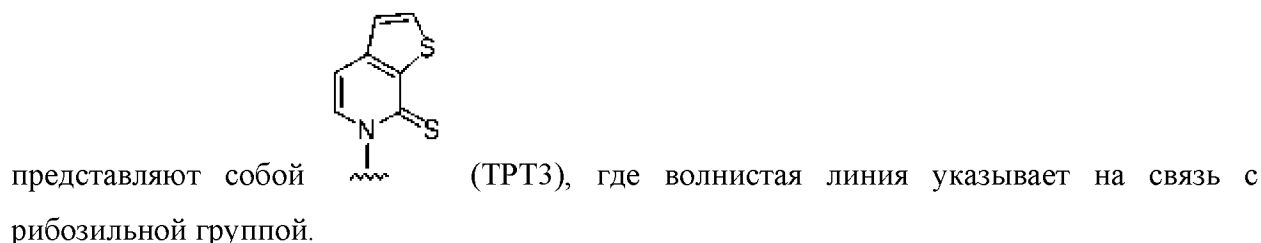




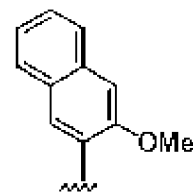
Вариант осуществления 212. Система по варианту осуществления 211, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



Вариант осуществления 213. Система по варианту осуществления 212, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, оба

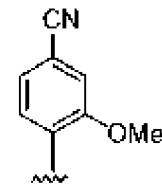


Вариант осуществления 214. Система по варианту осуществления 212, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



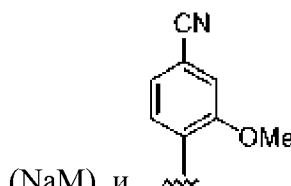
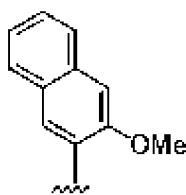
расположенных в антикодоне тРНК, оба представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 215. Система по варианту осуществления 212, в которой одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),

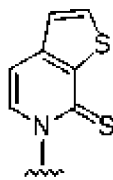


расположенных в антикодоне тРНК, оба представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 216. Система по варианту осуществления 212, в которой одно или несколько первых оснований не природного происхождения (X), расположенных

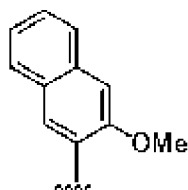


в кодоне иРНК, выбраны из (NaM) и (CNMO), и одно или несколько вторых оснований не природного происхождения (Y), расположенных в



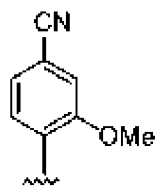
антикодоне тРНК представляет собой , где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 217. Система по варианту осуществления 216, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (NaM).

Вариант осуществления 218. Система по варианту осуществления 216, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне



иРНК, представляют собой (CNMO).

Вариант осуществления 219. Система по любому из вариантов осуществления 139-218, где, по меньшей мере, один кодон в иРНК выбран из АХС, GXC или GXU, где Х представляет собой основание не природного происхождения.

Вариант осуществления 220. Система по варианту осуществления 219, где, по меньшей мере, один кодон в иРНК представляет собой АХС, где Х представляет собой основание не природного происхождения.

Вариант осуществления 221. Система по варианту осуществления 219, где, по меньшей мере, один кодон в иРНК представляет собой GXC, где Х представляет собой основание не природного происхождения.

Вариант осуществления 222. Система по варианту осуществления 219, где, по меньшей мере, один кодон в иРНК представляет собой GXU, где Х представляет собой основание не природного происхождения.

Вариант осуществления 223. Система по любому из вариантов осуществления 139-218, где, по меньшей мере, один кодон в иРНК выбран из АХС, GXC или GXU, где, по меньшей мере, один антикодон в тРНК выбран из GYU, GYC и AYC, где Х представляет собой одно или несколько первых оснований не природного происхождения, и Y представляет собой одно или несколько вторых оснований не природного происхождения.

Вариант осуществления 224. Система по варианту 223, в которой Х и Y являются одинаковыми или разными.

Вариант осуществления 225. Система по варианту 224, где Х и Y являются одинаковыми.

Вариант осуществления 226. Система по варианту 224, где Х и Y являются разными.

Вариант осуществления 227. Система по варианту осуществления 223, где, по меньшей мере, один кодон в иРНК представляет собой АХС, и, по меньшей мере, один антикодон в тРНК представляет собой GYU.

Вариант осуществления 228. Система по варианту 227, где Х и Y являются одинаковыми или разными.

Вариант осуществления 229. Система по варианту 228, где Х и Y являются одинаковыми.

Вариант осуществления 230. Система по варианту 228, где Х и Y являются разными.

Вариант осуществления 231. Система по варианту осуществления 223, где, по меньшей мере, один кодон в иРНК представляет собой GXC, и, по меньшей мере, один антикодон в тРНК представляет собой GYC.

Вариант осуществления 232. Система по варианту 231, где Х и Y являются одинаковыми или разными.

Вариант осуществления 233. Система по варианту 232, где Х и Y являются

одинаковыми.

Вариант осуществления 234. Система по варианту 232, где X и Y являются разными.

Вариант осуществления 235. Система по варианту осуществления 223, где, по меньшей мере, один кодон в иРНК представляет собой GXU, и, по меньшей мере, один антикодон представляет собой AYC.

Вариант осуществления 236. Система по варианту 235, где X и Y являются одинаковыми или разными.

Вариант осуществления 237. Система по варианту 236, где X и Y являются одинаковыми.

Вариант осуществления 238. Система по варианту 236, где X и Y являются разными.

Вариант осуществления 239. Система по любому из вариантов осуществления 139-238, где тРНК получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 240. Система по любому из вариантов осуществления 139-238, где тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 241. Система по варианту осуществления 240, где тРНК и тРНК синтетаза получены из *Methanococcus jannaschii*.

Вариант осуществления 242. Система по варианту осуществления 240, где тРНК и тРНК синтетаза получены из *Methanosarcina barkeri*.

Вариант осуществления 243. Система по варианту осуществления 240, где тРНК и тРНК синтетаза получены из *Methanosarcina mazei*.

Вариант осуществления 244. Система по варианту осуществления 240, где тРНК и тРНК синтетаза получены из *Methanosarcina acetivorans*.

Вариант осуществления 245. Система по любому из вариантов осуществления 139-239, где тРНК получена из *Methanococcus jannaschii*, и тРНК синтетаза получена из *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 246. Система по любому из вариантов осуществления 139-239, где тРНК получена из *Methanosarcina barkeri*, и тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 247. Система по любому из вариантов осуществления 139-239, где тРНК получена из *Methanosarcina mazei*, и тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina acetivorans*.

Вариант осуществления 248. Система по любому из вариантов осуществления 139-239, где тРНК получена из *Methanosarcina acetivorans*, и тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri* или *Methanosarcina mazei*.

Вариант осуществления 249. Система по любому из вариантов осуществления 139-239, где тРНК получена из *Methanosarcina mazei*, и тРНК синтетаза получена из *Methanosarcina barkeri*.

Вариант осуществления 250. Система по любому из вариантов осуществления 139-

249, где клетка представляет собой клетку человека.

Вариант осуществления 251. Система согласно варианту осуществления 250, где клетка человека представляет собой клетку НЕК293Т.

Вариант осуществления 252. Система по любому из вариантов осуществления 139-239, в которой клетка представляет собой клетку хомяка.

Вариант осуществления 253. Система по варианту осуществления 252, где клетка хомяка представляет собой клетку яичника китайского хомяка (СНО).

Вариант осуществления 254. Система по любому из пп. 139-253, где аминокислота не природного происхождения:

- представляет собой аналог лизина;
- содержит ароматическую боковую цепь;
- содержит азидогруппу;
- содержит алкиновую группу; или
- содержат альдегидную или кетонную группу.

Вариант осуществления 255. Система по любому из вариантов осуществления 139-253, где аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, ТСО-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилфенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацетил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-(((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 256. Система по варианту осуществления 255, где аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 257. Способ по любому из вариантов осуществления 21-138, где иРНК и тРНК стабилизированы к деградации в эукариотической клетке.

Вариант осуществления 258. Способ по любому из вариантов осуществления 21-138 и 257, где полипептид получают путем трансляции иРНК с использованием тРНК с помощью рибосомы, эндогенной для эукариотической клетки.

Вариант осуществления 259. Система по любому из вариантов осуществления 139-

256, в которой иРНК и тРНК стабилизированы к деградации в эукариотической клетке.

Вариант осуществления 260. Система по любому из вариантов осуществления 139-256 и 259, где полипептид получают путем трансляции иРНК с использованием тРНК с помощью рибосомы, эндогенной для эукариотической клетки.

Вариант осуществления 261. Эукариотическая клетка, содержащая:

(а) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; и

(b) транспортную РНК (тРНК) с антикодоном второго основания не природного происхождения, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке, и где иРНК способна транслироваться в клетки для продуцирования полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения.

Вариант осуществления 262. Эукариотическая клетка по варианту осуществления 261, где тРНК заряжена аминокислотой не природного происхождения.

Вариант осуществления 263. Эукариотическая клетка по любому из вариантов осуществления 261-262, дополнительно содержащая полипептид, транслированный с иРНК, где полипептид содержит аминокислоту не природного происхождения, необязательно, где полипептид имеет эукариотический профиль гликозилирования.

Вариант осуществления 264. Эукариотическая клетка по любому из вариантов осуществления 261-263, дополнительно содержащая тРНК-синтетазу, где тРНК-синтетаза предпочтительно аминоацилирует тРНК с аминокислотой не природного происхождения.

Вариант осуществления 265. Эукариотическая клетка по любому из вариантов осуществления 261-264, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

Вариант осуществления 266. Эукариотическая клетка по любому из вариантов осуществления 261-265, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

Вариант осуществления 267. Эукариотическая клетка по любому из вариантов осуществления 261-266, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

Вариант осуществления 268. Эукариотическая клетка по любому из вариантов осуществления 261-267, где первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из:

(i) 2-тиоурацила, 2-тиотимина, 2'-дезоксинуридина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-галоурацила; 5-пропинилурацила, 6-азотимина, 6-азоурацила, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила, псевдоурацила, метилового эфира урацил-5-оксуксусной кислоты, урацил-5-оксуксусной

кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, 5-метил-2-тиоурацила, 4-тиоурацила, 5-метилурацила, 5'-метоксикарбоксиметилурацила, 5-метоксиурацила, урацил-5-оксиуксусной кислоты, 5-(карбоксигидроксиметил)урацила, 5-карбоксиметиламинометил-2-тиуридина, 5-карбоксиметиламинометилурацила или дигидроурацила;

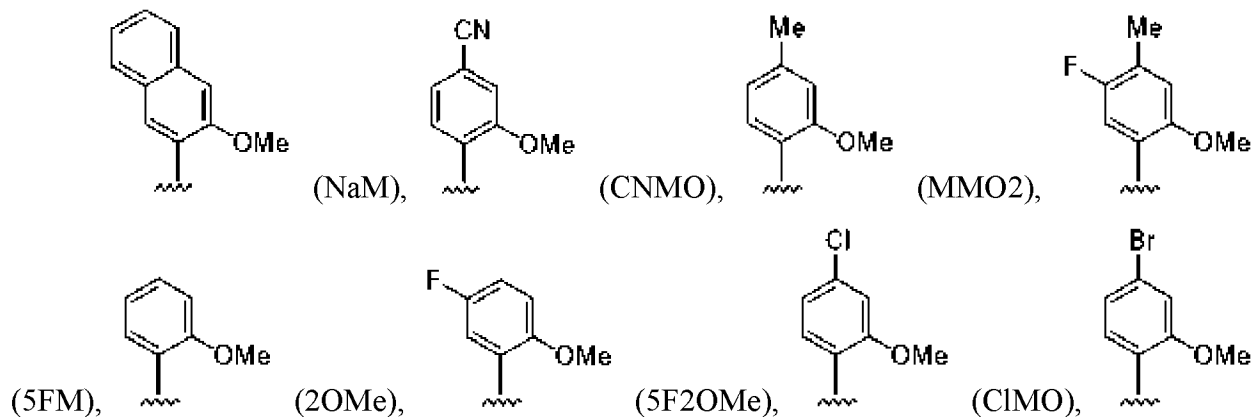
(ii) 5-гидроксиметилцитозина, 5-трифторметилцитозина, 5-галоцитозина, 5-пропинилцитозина, 5-гидроксицитозина, циклоцитозина, цитозинарабинозида, 5,6-дигидроцитозина, 5-нитроцитозина, 6-азоцитозина, азацитозина, N4-этилцитозина, 3-метилцитозина, 5-метилцитозина, 4-ацетилцитозина, 2-тиоцитозина, феноксазин цитидин ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазин цитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), феноксазин цитидин (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазол цитидин (2H-пиримидо[4,5-b]индол-2-она) или пиридоиндол цитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она);

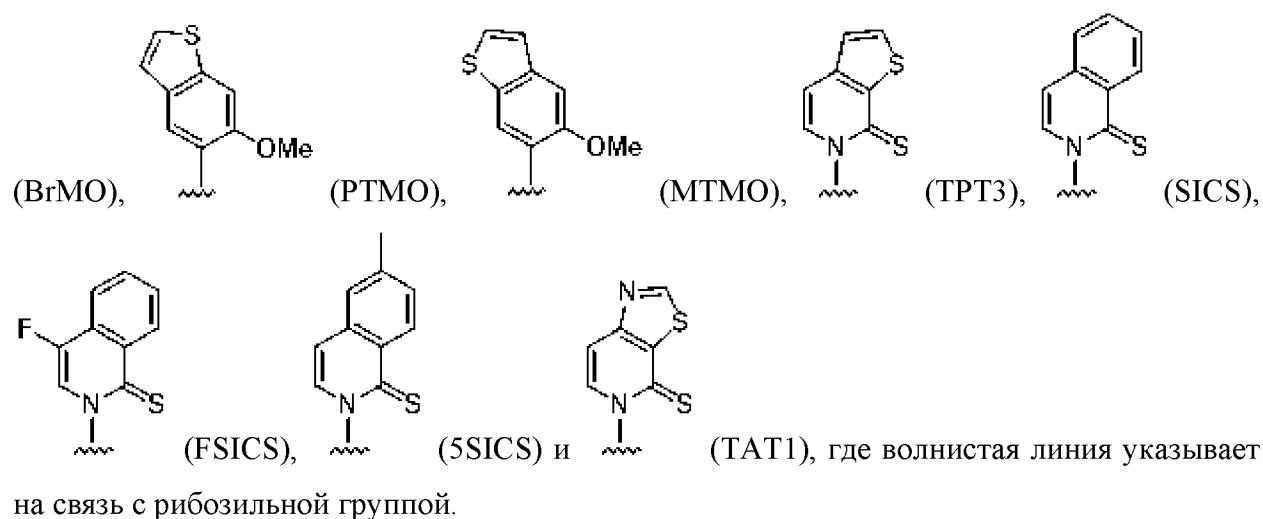
(iii) 2-аминоаденина, 2-пропиладенина, 2-аминоаденина, 2-F-аденина, 2-аминопропиладенина, 2-амино-2'-дезоксаденозина, 3-дезааденина, 7-метиладенина, 7-дезааденина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных аденинов, N6-изопентениладенина, 2-метиладенина, 2,6-диаминопурина, 2-метилтио-N6-изопентениладенина или 6-азааденина;

(iv) 2-метилгуанина, 2-пропил- и алкилпроизводных гуанина, 3-дезагуанина, 6-тиогуанина, 7-метилгуанина, 7-дезагуанина, 7-дезагуанозина, 7-деза-8-азагуанина, 8-азагуанина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных гуанинов, 1-метилгуанина, 2,2-диметилгуанина, 7-метилгуанина или 6-азагуанина; и

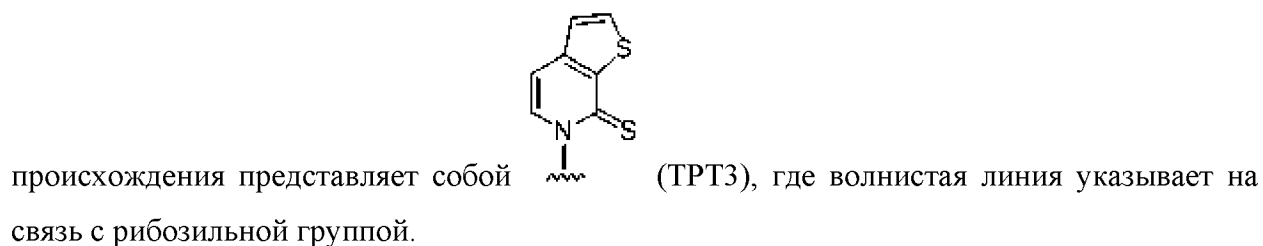
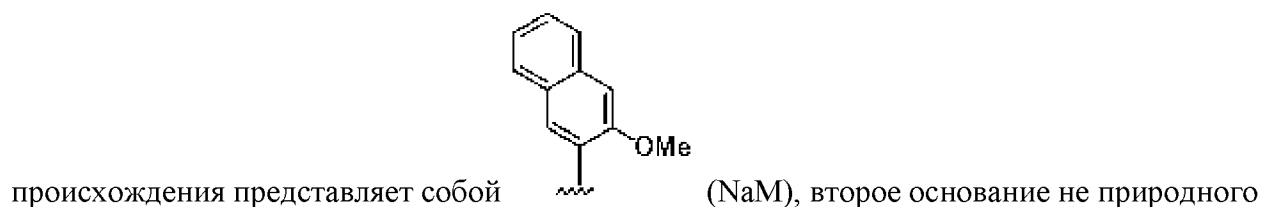
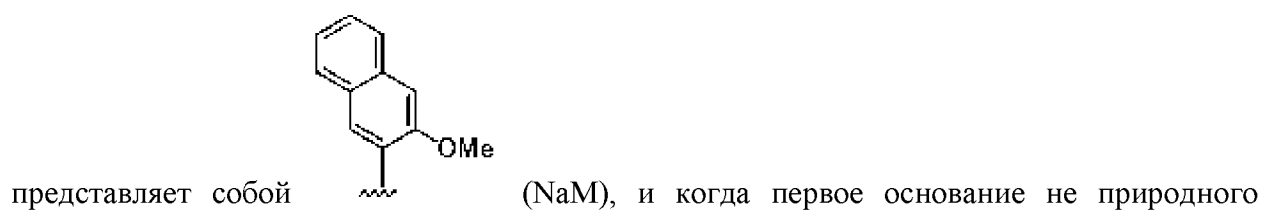
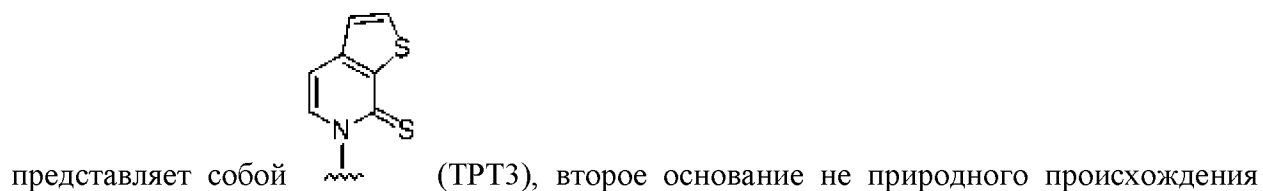
(v) гипоксантина, ксантина, 1-метилюозина, кеозина, бета-D-галактозилкеозина, инозина, бета-D-маннозилкеозина, вибутоксозина, гидроксимочевины, (аср3)w, 2-аминопиридины или 2-пиридоны.

Вариант осуществления 269. Эукариотическая клетка по любому из вариантов осуществления 261-267, где первое основание не природного происхождения и второе основание не природного происхождения, каждое независимо, выбраны из группы, состоящей из

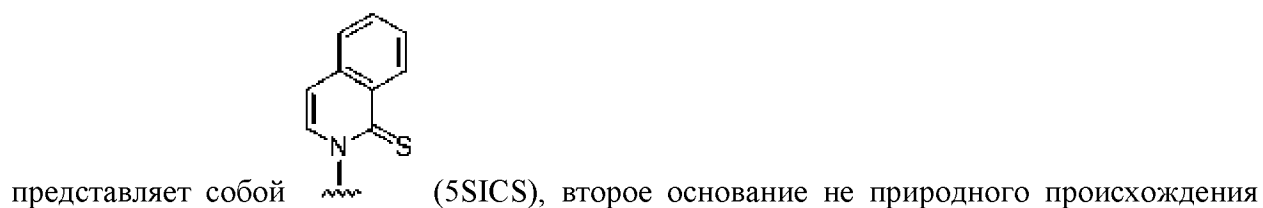


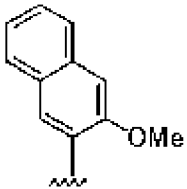
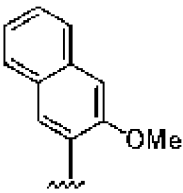
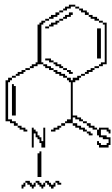


Вариант осуществления 270. Эукариотическая клетка по любому из вариантов осуществления 261-267, когда первое основание не природного происхождения

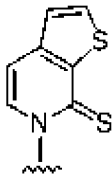
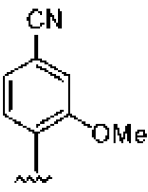
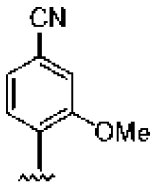
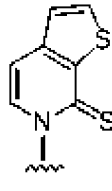


Вариант осуществления 271. Эукариотическая клетка по любому из вариантов осуществления 261-267, когда первое основание не природного происхождения

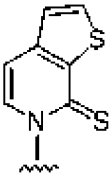


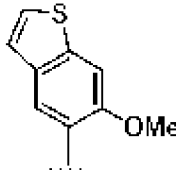
представляет собой  (NaM), и когда первое основание не природного происхождения представляет собой  (NaM), второе основание не природного происхождения представляет собой  (5SICS), где волнистая линия указывает на связь с рибозильной группой.

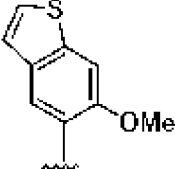
Вариант осуществления 272. Эукариотическая клетка по любому из вариантов осуществления 261-267, когда первое основание не природного происхождения

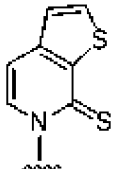
представляет собой  (TPT3), второе основание не природного происхождения представляет собой  (CNMO), и когда первое основание не природного происхождения представляет собой  (CNMO), второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 273. Эукариотическая клетка по любому из вариантов осуществления 261-267, когда первое основание не природного происхождения

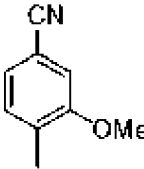
представляет собой  (TPT3), второе основание не природного происхождения

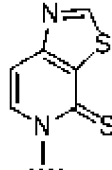
представляет собой  (PTMO), и когда первое основание не природного

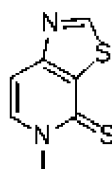
происхождения представляет собой  (PTMO), второе основание не

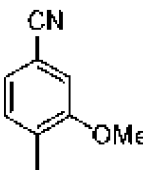
природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 274. Эукариотическая клетка по любому из вариантов осуществления 261-267, когда первое основание не природного происхождения

представляет собой  (CNMO), второе основание не природного происхождения

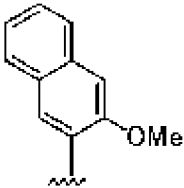
представляет собой  (TAT1), и когда первое основание не природного

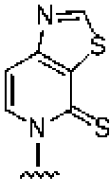
происхождения представляет собой  (TAT1), второе основание не природного

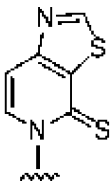
происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

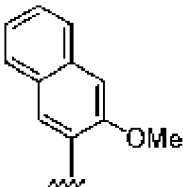
Вариант осуществления 275. Эукариотическая клетка по любому из вариантов

осуществления 261-267, когда первое основание не природного происхождения

представляет собой  (NaM), второе основание не природного происхождения

представляет собой  (TAT1), и когда первое основание не природного

происхождения представляет собой  (TAT1), второе основание не природного

происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 276. Эукариотическая клетка по любому из вариантов осуществления 261-275, где первое основание не природного происхождения или второе основание не природного происхождения содержат модифицированную сахарную группу, выбранную из группы, состоящей из:

модификации в положении 2':

ОН, замещенный низший алкил, алкарил, аралкил, О-алкарил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F;

О-алкил, S-алкил, N-алкил;

О-алкенил, S-алкенил, N-алкенил;

О-алкинил, S-алкинил, N-алкинил;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, где алкил, алкенил и алкинил могут быть замещены или не замещены C₁-C₁₀ алкилом, C₂-C₁₀ алкенилом, C₂-C₁₀ алкинилом, -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-NH₂ и -O(CH₂)_nON[(CH₂)_nCH₃]₂, где n и m составляют от 1 до примерно 10;

и/или модификации в положении 5':

5'-винил, 5'-метил (R или S);

модификации в положении 4':

4'-S, гетероциклоалкил, гетероциклоалкарил, аминокалкиламино, полиалкиламино, замещенный силлил, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для

улучшения фармакодинамических свойств олигонуклеотида, и любая их комбинация.

Вариант осуществления 277. Эукариотическая клетка по любому из вариантов осуществления 263-276, где, по меньшей мере, одна аминокислота не природного происхождения:

- представляет собой аналог лизина;
- содержит ароматическую боковую цепь;
- содержит азидогруппу;
- содержит алкиновую группу; или
- содержат альдегидную или кетонную группу.

Вариант осуществления 278. Эукариотическая клетка согласно варианту осуществления 277, где по меньшей мере, одна аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, O-аллилтирозина, O-метил-L-тирозина, O-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-O-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-(((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 279. Эукариотическая клетка по варианту осуществления 278, где, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 280. Эукариотическая клетка по любому из вариантов осуществления 261-279, где эукариотическая клетка представляет собой клетку человека.

Вариант осуществления 281. Эукариотическая клетка по непосредственно предшествующему варианту осуществления, где клетка человека представляет собой клетку HEK293T.

Вариант осуществления 282. Эукариотическая клетка по любому из вариантов осуществления 261-279, где клетка необязательно представляет собой клетку млекопитающего, где клетка млекопитающего представляет собой клетку хомяка.

Вариант осуществления 283. Эукариотическая клетка по непосредственно предшествующему варианту осуществления, где клетка млекопитающего представляет

собой клетку яичника китайского хомяка (СНО).

Вариант осуществления 284. Эукариотическая клетка по любому из вариантов осуществления 261-283, где клетка необязательно выделена, где клетка очищена.

Вариант осуществления 285. Эукариотическая клетка по любому из вариантов осуществления 261-284, дополнительно содержащая полипептид, транслированный с иРНК, где полипептид содержит аминокислоту не природного происхождения и профиль гликозилирования млекопитающих.

Вариант осуществления 285.1. Полусинтетический организм, содержащий эукариотическую клетку по любому из вариантов осуществления 261-285.

Вариант осуществления 286. Культура эукариотических клеток, содержащая множество эукариотических клеток по любому из вариантов осуществления 261-285.

Вариант осуществления 286.1. Способ доставки клетки в организм, включающий контакт организма с клеткой по любому из вариантов осуществления 261-285.

Вариант осуществления 286.2. Способ по варианту осуществления 286.1, где организм представляет собой млекопитающее, необязательно где млекопитающее представляет собой человека.

Вариант осуществления 287. Способ получения полипептида, содержащего по меньшей мере, одну аминокислоту не природного происхождения, в эукариотической клетке, включающий:

а) введение в клетку:

(i) информационной РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; и

(ii) транспортной РНК (тРНК) с антикодоном и вторым неприродным основанием в эукариотической клетке, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке; и

(b) трансляцию полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения из иРНК с использованием тРНК.

Вариант осуществления 288. Способ по предыдущему варианту осуществления, где тРНК заряжена аминокислотой не природного происхождения.

Вариант осуществления 289. Способ получения полипептида, содержащего по меньшей мере, одну аминокислоту не природного происхождения, в эукариотической клетке, включающий:

(a) получение эукариотической клетки, содержащей:

(i) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения;

(ii) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке;

(b) трансляцию полипептида, содержащего по меньшей мере, одну аминокислоту не природного происхождения, из иРНК с использованием тРНК посредством рибосомы, эндогенной для эукариотической клетки.

Вариант осуществления 290. Способ получения полипептида в эукариотической клетке, где полипептид содержит, по меньшей мере, одну аминокислоту не природного происхождения, где способ включает:

(a) получение эукариотической клетки, где эукариотическая клетка содержит:

(i) иРНК, содержащую кодон, где кодон содержит первое основание не природного происхождения;

(ii) тРНК, содержащую антикодон, где антикодон содержит второе основание не природного происхождения, и где первое и второе основания не природного происхождения способны образовывать комплементарную пару оснований; и

(b) тРНК синтетазу, где тРНК синтетаза предпочтительно аминоацилирует тРНК, по меньшей мере, с одной аминокислотой не природного происхождения, по сравнению с природной аминокислотой; и

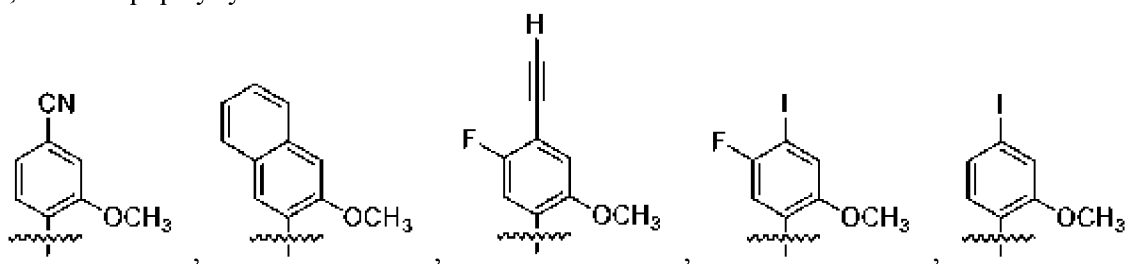
(c) предоставление еще одной аминокислоты не природного происхождения в эукариотическую клетку, где эукариотическая клетка продуцирует полипептид, содержащий, по меньшей мере, одну аминокислоту не природного происхождения.

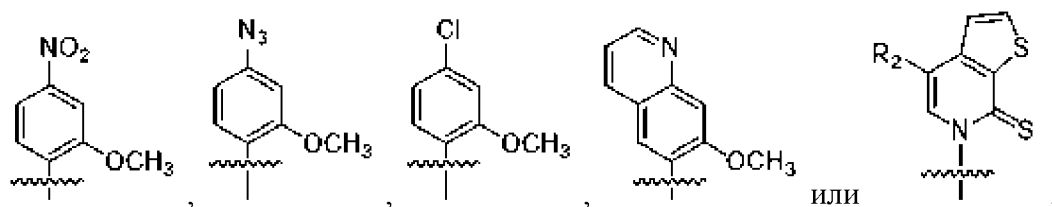
Вариант осуществления 291. Способ по любому из вариантов осуществления 287-290, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

Вариант осуществления 292. Способ по любому из вариантов осуществления 287-290, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

Вариант осуществления 293. Способ по любому из вариантов осуществления 287-290, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

Вариант осуществления 294. Способ по любому из вариантов осуществления 287-293, где одно или несколько оснований не природного происхождения, содержащих кодон в иРНК, имеют формулу





где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азида, и волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 295. Способ по любому из вариантов осуществления 287-293, где первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из:

(i) 2-тиоурацила, 2-тиотимина, 2'-дезоксуридина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-галоурацила; 5-пропинилурацила, 6-азотимина, 6-азоурацила, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила, псевдоурацила, метилового эфира урацил-5-оксуксусной кислоты, урацил-5-оксуксусной кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, 5-метил-2-тиоурацила, 4-тиоурацила, 5-метилурацила, 5'-метоксикарбоксиметилурацила, 5-метоксиурацила, урацил-5-оксиуксусной кислоты, 5-(карбоксигидроксиметил)урацила, 5-карбоксиметиламинометил-2-тиуридина, 5-карбоксиметиламинометилурацила или дигидроурацила;

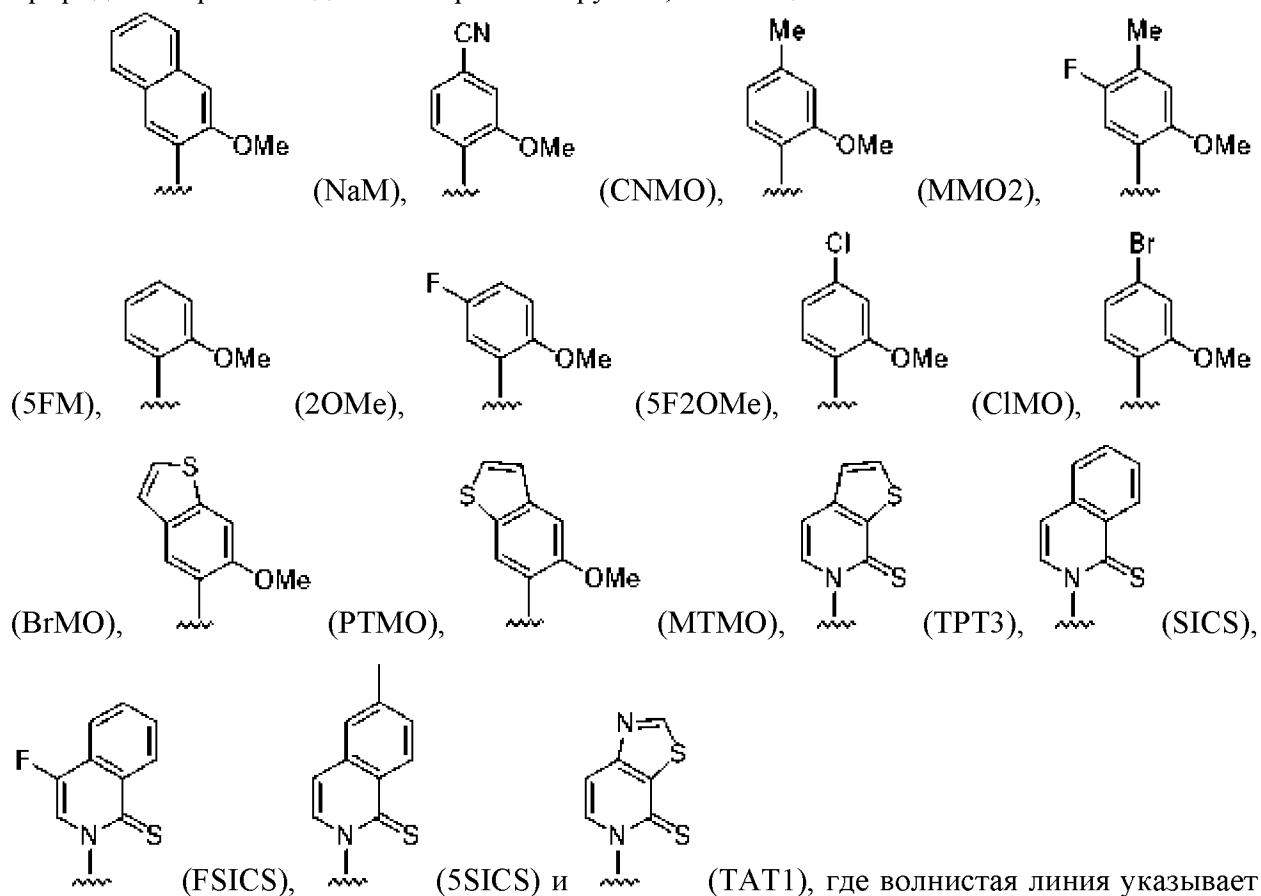
(ii) 5-гидроксиметилцитозина, 5-трифторметилцитозина, 5-галоцитозина, 5-пропинилцитозина, 5-гидроксицитозина, циклоцитозина, цитозинарабинозида, 5,6-дигидроцитозина, 5-нитроцитозина, 6-азоцитозина, азацитозина, N4-этилцитозина, 3-метилцитозина, 5-метилцитозина, 4-ацетилцитозина, 2-тиоцитозина, феноксазин цитидин ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазин цитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), феноксазин цитидин (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазол цитидин (2H-пиримидо[4,5-b]индол-2-она) или пиридоиндол цитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она);

(iii) 2-аминоаденина, 2-пропиладенина, 2-аминоаденина, 2-F-аденина, 2-аминопропиладенина, 2-амино-2'-дезоксаденозина, 3-дезааденина, 7-метиладенина, 7-дезааденина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных аденинов, N6-изопентениладенина, 2-метиладенина, 2,6-диаминопурина, 2-метилтио-N6-изопентениладенина или 6-азааденина;

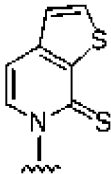
(iv) 2-метилгуанина, 2-пропил- и алкилпроизводных гуанина, 3-дезагуанина, 6-тиогуанина, 7-метилгуанина, 7-дезагуанина, 7-дезагуанозина, 7-деза-8-азагуанина, 8-азагуанина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксизамещенных гуанинов, 1-метилгуанина, 2,2-диметилгуанина, 7-метилгуанина или 6-азагуанина; и

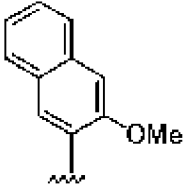
(v) гипоксантина, ксантина, 1-метилюридина, кеозина, бета-D-галактозилкеозина, инозина, бета-D-маннозилкеозина, вибутоксоина, гидроксимочевина, (аср3)w, 2-аминопиридины или 2-пиридоны.

Вариант осуществления 296. Способ по любому из вариантов осуществления 287-295, в котором первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из

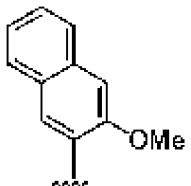


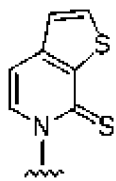
Вариант осуществления 297. Способ согласно варианту осуществления 296, где,

когда первое основание не природного происхождения представляет собой 

(TPT3), второе основание не природного происхождения представляет собой 

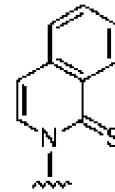
(NaM), и когда первое основание не природного происхождения представляет собой

 (NaM), второе основание не природного происхождения представляет собой

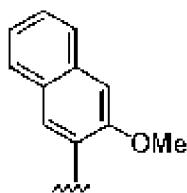


(TPT3), где волнистая линия указывает на связь с рибозильной группой.

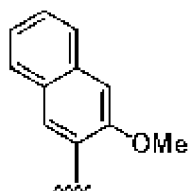
Вариант осуществления 298. Способ согласно варианту осуществления 296, где,



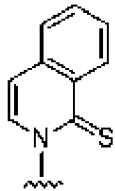
когда первое основание не природного происхождения представляет собой (5SICS), второе основание не природного происхождения представляет собой



(NaM), и когда первое основание не природного происхождения

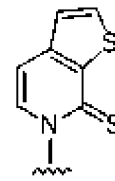


представляет собой (NaM), второе основание не природного происхождения

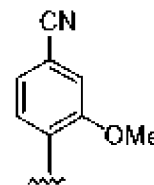


представляет собой (5SICS), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 299. Способ согласно варианту осуществления 296, где,

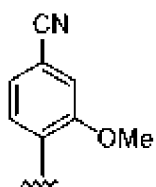


когда первое основание не природного происхождения представляет собой

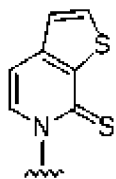


(TPT3), второе основание не природного происхождения представляет собой

(CNMO), и когда первое основание не природного происхождения представляет собой

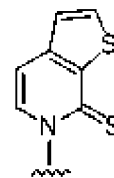


(CNMO), второе основание не природного происхождения представляет собой

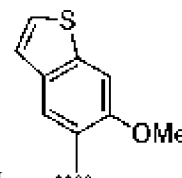


(TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 300. Способ согласно варианту осуществления 296, где,

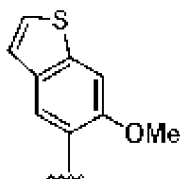


когда первое основание не природного происхождения представляет собой

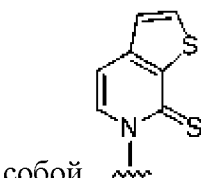


(TPT3), второе основание не природного происхождения представляет собой

(PTMO), и когда первое основание не природного происхождения представляет собой

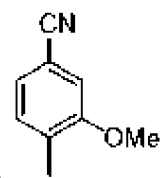


(PTMO), второе основание не природного происхождения представляет

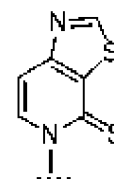


собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 301. Способ по варианту осуществления 296, где, когда

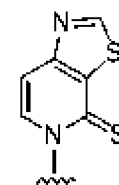


первое основание не природного происхождения представляет собой (CNMO),

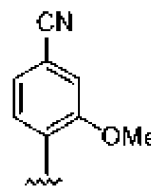


второе основание не природного происхождения представляет собой (TAT1), и

когда первое основание не природного происхождения представляет собой



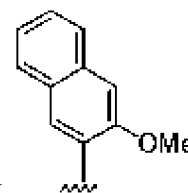
(TAT1), второе основание не природного происхождения представляет собой



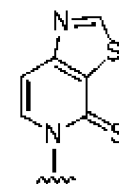
(CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 302. Способ согласно варианту осуществления 296, где,

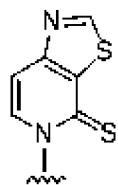
когда первое основание не природного происхождения представляет собой



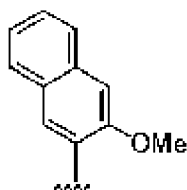
(NaM), второе основание не природного происхождения представляет собой



(TAT1), и когда первое основание не природного происхождения представляет собой



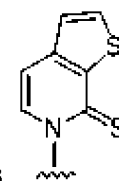
(TAT1), второе основание не природного происхождения представляет собой

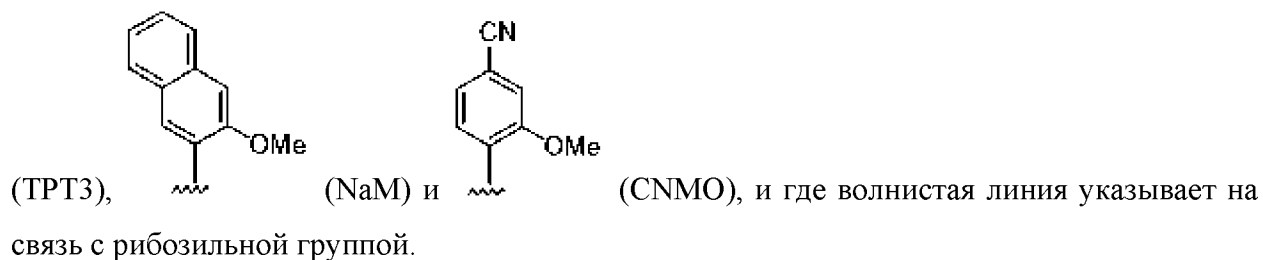


(NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 303. Способ по любому из вариантов осуществления 287-296, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в первом положении (X-

N-N) в кодоне иРНК, где основание не природного происхождения выбрано из

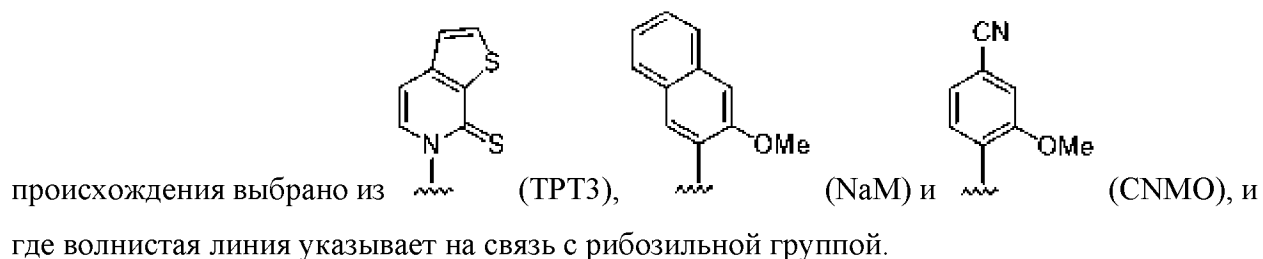




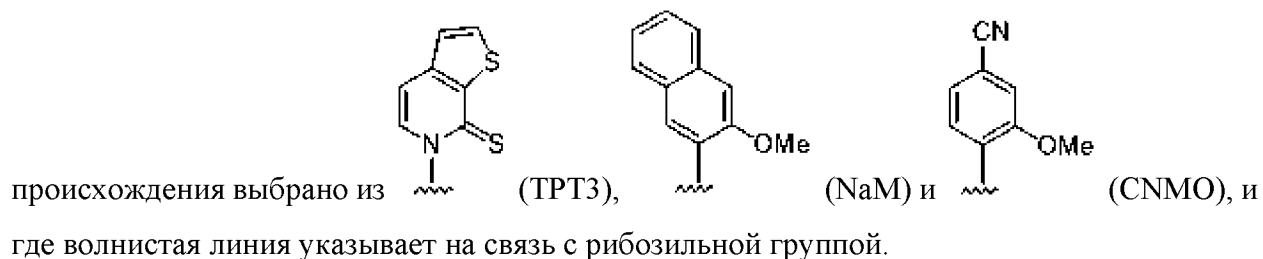
Вариант осуществления 304. Способ по любому из вариантов осуществления 287-296, где кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в среднем положении



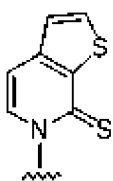
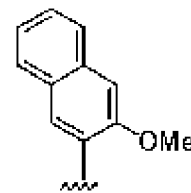
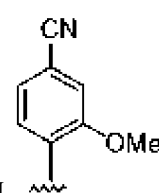
Вариант осуществления 305. Способ по любому из вариантов осуществления 287-296, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК, где основание не природного



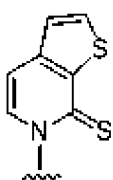
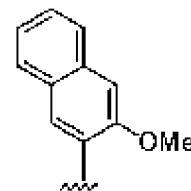
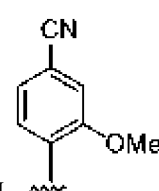
Вариант осуществления 306. Способ по любому из вариантов осуществления 287-296, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где второе основание не природного происхождения (X) расположено в первом положении (X-N-N) в антикодоне тРНК, где основание не природного



Вариант осуществления 307. Способ по любому из вариантов осуществления 287-296, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где второе основание не природного происхождения (X) расположено в среднем положении (N-X-N) в антикодоне тРНК, где основание не природного

происхождения выбрано из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 308. Способ по любому из вариантов осуществления 287-296, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где второе основание не природного происхождения (X) расположено в последнем положении (N-N-X) в антикодоне. тРНК, где основание не природного

происхождения выбрано из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 309. Способ по любому из вариантов по варианту осуществления 287-296, где кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) кодона в иРНК расположено в первом положении (X-N-N) кодона, и второе основание не природного происхождения (Y) антикодона тРНК расположено в последнем положении (N-N-Y) антикодона.

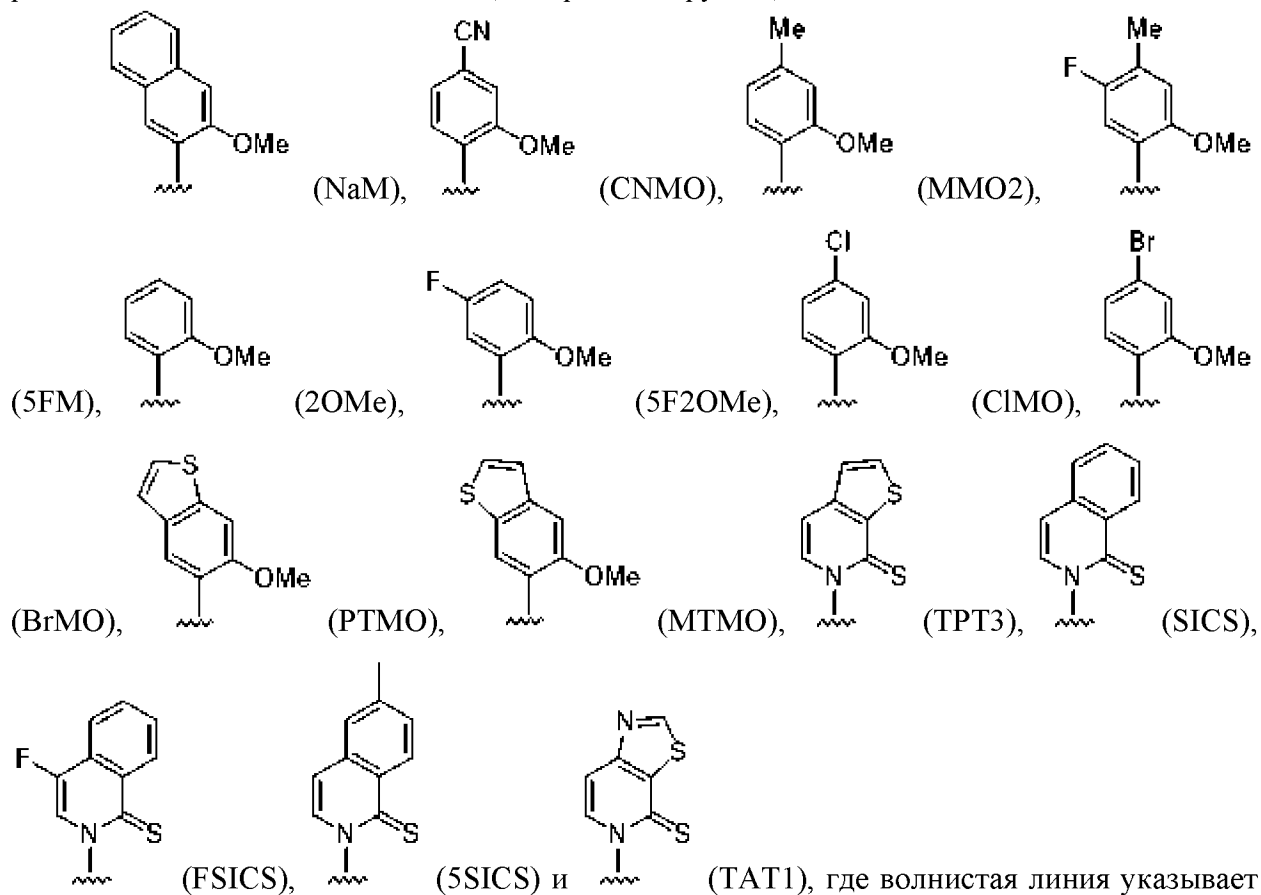
Вариант осуществления 310. Способ по любому из вариантов осуществления 287-296, отличающийся тем, что кодон и антикодон каждый содержит три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в среднем положении (N-X-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в среднем положении (N-Y-N) антикодона.

Вариант осуществления 311. Способ по любому из вариантов осуществления 287-296, где кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в последнем положении (N-N-X) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в первом положении (Y-N-N) антикодона.

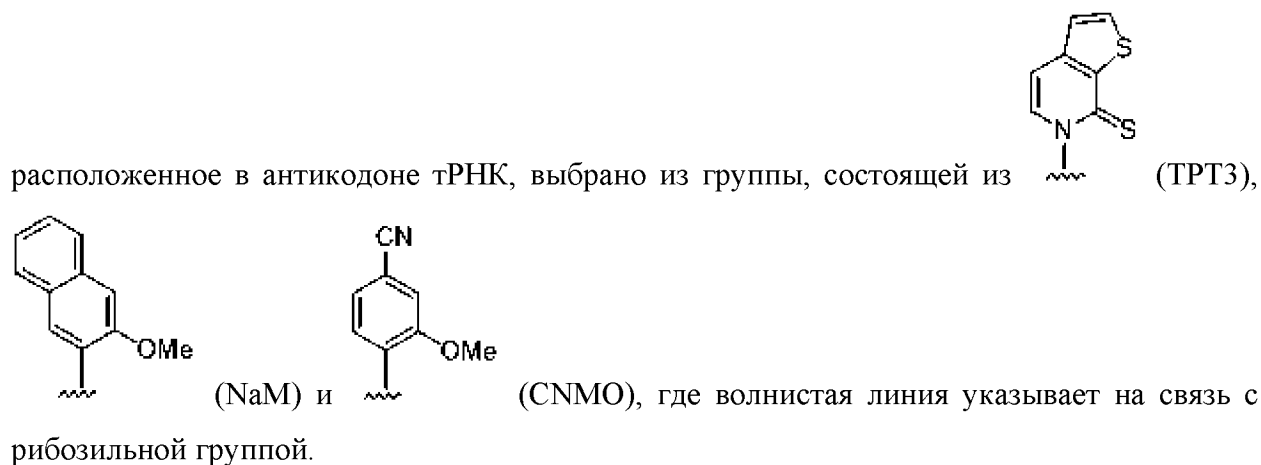
Вариант осуществления 312. Способ по любому из вариантов осуществления 309-

311, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными.

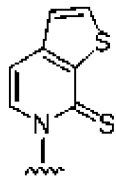
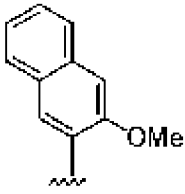
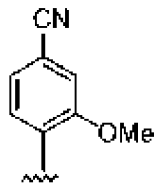
Вариант осуществления 313. Способ по любому из вариантов осуществления 309-312, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано из группы, состоящей из



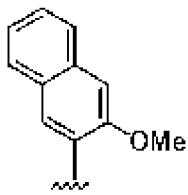
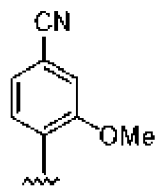
Вариант осуществления 314. Способ по варианту осуществления 313, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано из группы, состоящей из

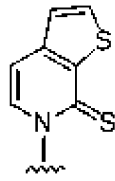


Вариант осуществления 315. Способ согласно варианту осуществления 314, где первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК,

оба представляют собой  (TPT3),  (NaM) или  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 316. Способ по варианту осуществления 314, отличающийся тем, что первое основание не природного происхождения (X),

расположенное в кодоне иРНК, выбрано из  (NaM) и  (CNMO), и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК,

представляет собой  (TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 317. Способ по любому из вариантов осуществления, 287-290, 292, 294-302, 304, 307 и 410, где кодон в иРНК выбран из AXC, GXC или GXU, где X представляет собой первое основание не природного происхождения.

Вариант осуществления 318. Способ в соответствии с непосредственно предшествующим вариантом осуществления, где антикодон в тРНК выбран из GYU, GYC и AYC, и Y представляет собой второе основание не природного происхождения.

Вариант осуществления 319. Способ по варианту осуществления 318, отличающийся тем, что кодон в иРНК представляет собой AXC, и антикодон в тРНК представляет собой GYU.

Вариант осуществления 320. Способ по варианту осуществления 318, отличающийся тем, что кодон в иРНК представляет собой GXC, и антикодон в тРНК представляет собой GYC.

Вариант осуществления 321. Способ по варианту осуществления 318, отличающийся тем, что кодон в иРНК представляет собой GXU, и антикодон представляет собой AYC.

Вариант осуществления 322. Способ по любому из вариантов осуществления 287-321, где первое основание не природного происхождения или второе основание не

природного происхождения содержат модифицированную сахарную группу, выбранную из группы, состоящей из:

модификации в положении 2':

ОН, замещенный низший алкил, алкарил, аралкил, О-алкарил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F;

О-алкил, S-алкил, N-алкил;

О-алкенил, S-алкенил, N-алкенил;

О-алкинил, S-алкинил, N-алкинил;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, где алкил, алкенил и алкинил могут быть замещены или не замещены C₁-C₁₀ алкилом, C₂-C₁₀ алкенилом, C₂-C₁₀ алкинилом, -O[(CH₂)_nO]_mCH₃, -O(CH₂)_nOCH₃, -O(CH₂)_nNH₂, -O(CH₂)_nCH₃, -O(CH₂)_n-NH₂ и -O(CH₂)_nON[(CH₂)_nCH₃]₂, где n и m составляют от 1 до примерно 10;

и/или модификации в положении 5':

5'-винил, 5'-метил (R или S);

модификации в положении 4':

4'-S, гетероциклоалкил, гетероциклоалкарил, аминоалкиламино, полиалкиламино, замещенный силлил, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для улучшения фармакодинамических свойств олигонуклеотида, и любая их комбинация.

Вариант осуществления 323. Способ по любому из вариантов осуществления 287-322, где, по меньшей мере, одна аминокислота не природного происхождения:

представляет собой аналог лизина;

содержит ароматическую боковую цепь;

содержит азидогруппу;

содержит алкиновую группу; или

содержат альдегидную или кетонную группу.

Вариант осуществления 324. Способ по любому из вариантов осуществления 287-322, отличающийся тем, что по меньшей мере, одна аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокс)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокс)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-

оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 325. Способ по варианту осуществления 324, где аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 326. Способ по любому из вариантов осуществления 287-325, где клетка представляет собой клетку человека.

Вариант осуществления, 327. Способ по варианту осуществления 326, где клетка человека представляет собой клетку НЕК293Т.

Вариант осуществления 328. Способ по любому из вариантов осуществления 287-325, где клетка представляет собой клетку хомяка.

Вариант осуществления, 329. Способ согласно варианту осуществления 328, где клетка хомяка представляет собой клетку яичника китайского хомяка (CHO).

Вариант осуществления 330. Способ по любому из вариантов осуществления 287-329, где тРНК получают из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 331. Способ по любому из вариантов осуществления 287-330, где клетка содержит тРНК синтетазу, полученную из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 332. Система для экспрессии полипептида не природного происхождения, содержащая:

- (а) по меньшей мере, одну аминокислоту не природного происхождения;
- (б) иРНК, кодирующую не встречающийся в природе полипептид, где указанная иРНК содержит, по меньшей мере, один кодон, содержащий одно или несколько первых оснований не природного происхождения;
- (с) тРНК, содержащую, по меньшей мере, один антикодон, содержащий одно или несколько вторых оснований не природного происхождения, где одно или несколько первых оснований не природного происхождения и одно или несколько вторых оснований не природного происхождения способны образовывать одну или несколько комплементарных пар оснований;
- (д) эукариотическую рибосому, способную транслировать иРНК в полипептид, содержащий аминокислоту не природного происхождения, с использованием тРНК и тРНК-синтетазы,

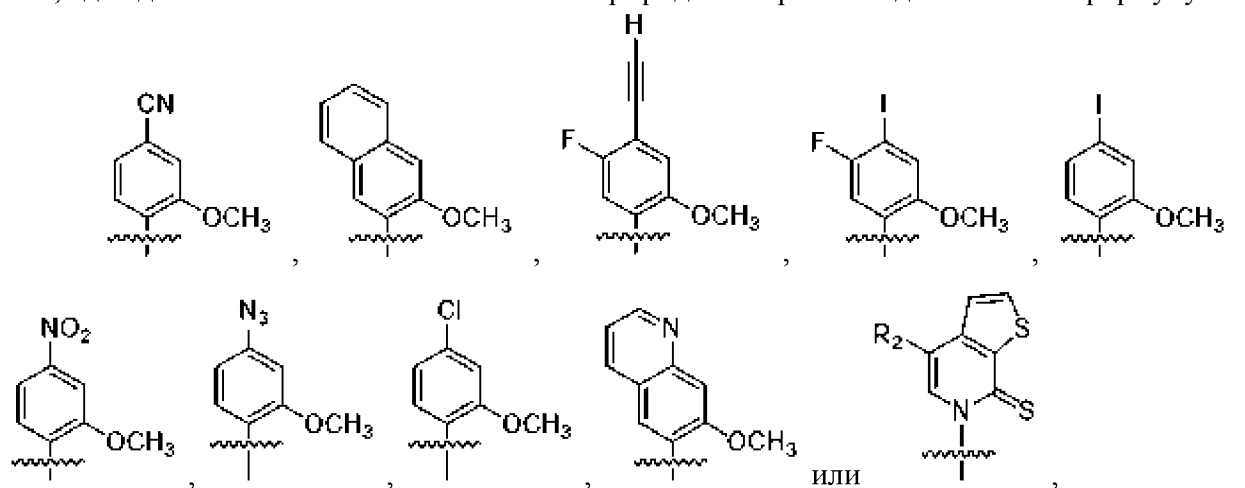
где тРНК заряжена аминокислотой не природного происхождения, или система дополнительно содержит тРНК синтетазу или одну или несколько конструкций нуклеиновых кислот, содержащих последовательность нуклеиновой кислоты, кодирующую тРНК синтетазу, где тРНК синтетаза предпочтительно аминоацилирует тРНК, по меньшей мере, с одной аминокислотой не природного происхождения.

Вариант осуществления 333. Система по варианту осуществления 332, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в, по меньшей мере, одном кодоне иРНК.

Вариант осуществления 334. Система по варианту осуществления 332, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК.

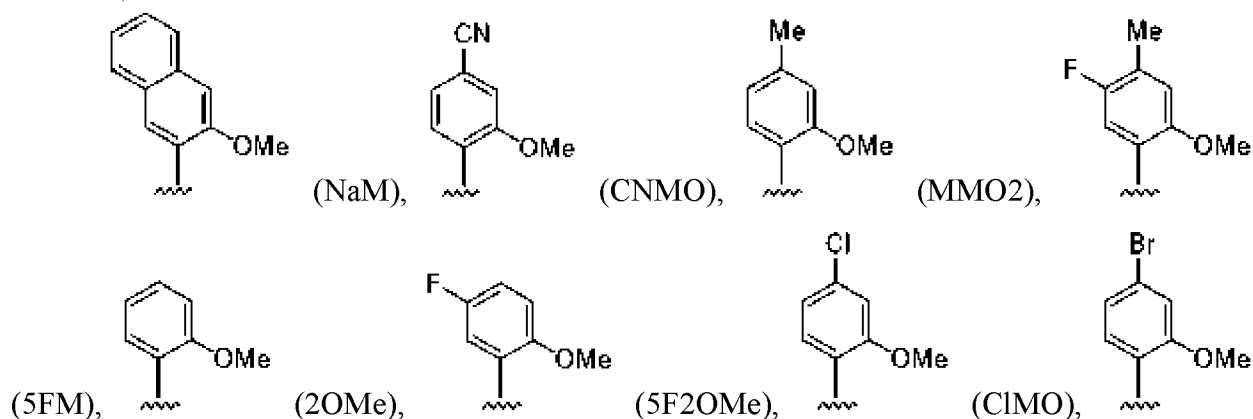
Вариант осуществления 335. Система по варианту осуществления 332, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в, по меньшей мере, одном кодоне иРНК.

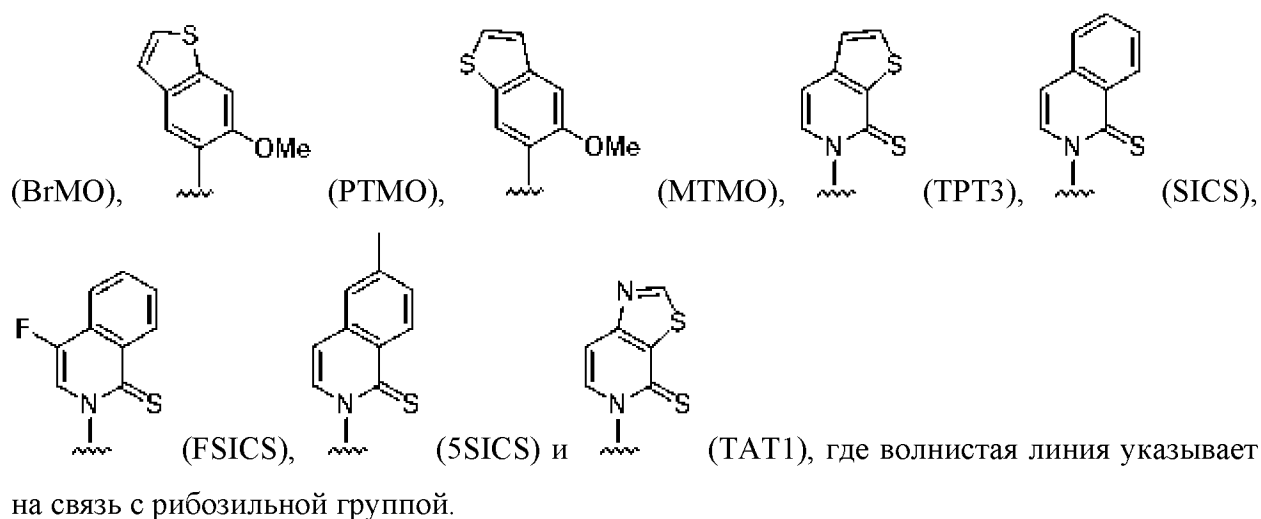
Вариант осуществления 336. Система по любому из вариантов осуществления 332-335, где одно или несколько оснований не природного происхождения имеют формулу



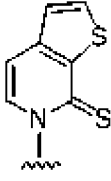
где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азида, и волнистая линия указывает на связь с рибозильной группой.

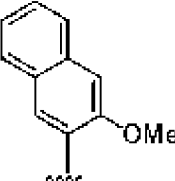
Вариант осуществления 337. Система по любому из вариантов осуществления 332-335, где одно или несколько первых оснований не природного происхождения или одно или несколько вторых оснований не природного происхождения выбраны из группы, состоящей из

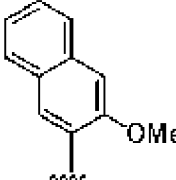


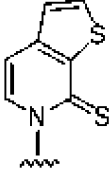


Вариант осуществления 338. Система по варианту осуществления 337, где одно или

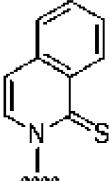
несколько первых оснований не природного происхождения представляют собой  (TPT3), одно или несколько вторых оснований не природного происхождения

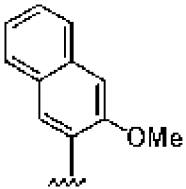
представляют собой  (NaM), и когда одно или несколько первых оснований

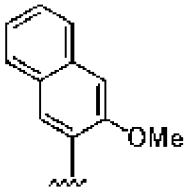
не природного происхождения представляют собой  (NaM), второе основание

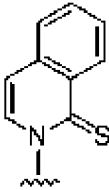
не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 339. Система по варианту осуществления 337, когда одно или несколько первых оснований не природного происхождения представляют собой

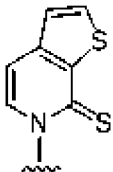
 (5SICS), одно или несколько вторых оснований не природного происхождения

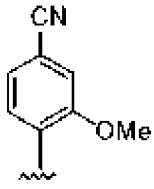
представляют собой  (NaM), и когда одно или несколько первых оснований

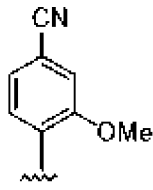
не природного происхождения представляют собой  (NaM), одно или

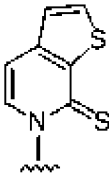
несколько вторых оснований не природного происхождения представляют собой  (5SICS), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 340. Система согласно варианту осуществления 337, когда одно или несколько первых оснований не природного происхождения представляют собой

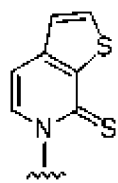
 (TPT3), одно или несколько вторых оснований не природного происхождения

представляют собой  (CNMO), и когда одно или несколько первых оснований

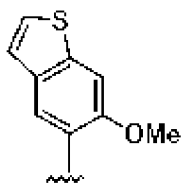
не природного происхождения представляют собой  (CNMO), одно или

несколько вторых оснований не природного происхождения представляют собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

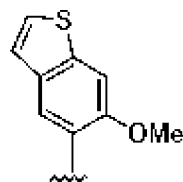
Вариант осуществления 341. Система по варианту осуществления 337, когда одно или несколько первых оснований не природного происхождения представляют собой



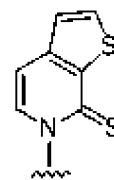
(TPT3), одно или несколько вторых оснований не природного происхождения



представляют собой (PTMO), и когда одно или несколько первых оснований

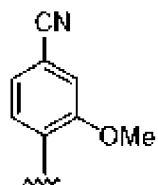


не природного происхождения представляют собой (PTMO), одно или

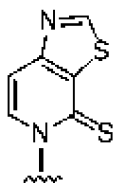


несколько вторых оснований не природного происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

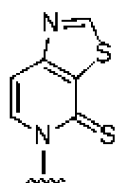
Вариант осуществления 342. Система по варианту осуществления 337, когда одно или несколько первых оснований не природного происхождения представляют собой



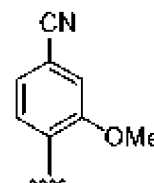
(CNMO), одно или несколько вторых оснований не природного происхождения



представляют собой (TAT1), и когда одно или несколько первых оснований не



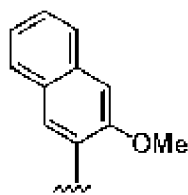
природного происхождения представляют собой (TAT1), одно или несколько



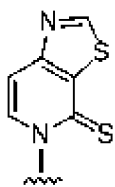
вторых оснований не природного происхождения представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 343. Система по варианту осуществления 337, когда одно

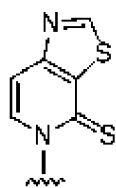
или несколько первых оснований не природного происхождения представляют собой



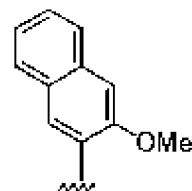
(NaM), одно или несколько вторых оснований не природного происхождения



представляют собой (TAT1), и когда одно или несколько первых оснований не



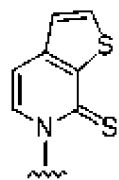
природного происхождения представляют собой (TAT1), одно или несколько



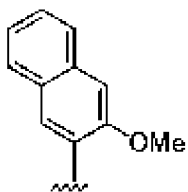
вторых оснований не природного происхождения представляют собой

(NaM), где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 344. Система по варианту осуществления 337, когда одно или несколько первых оснований не природного происхождения представляют собой

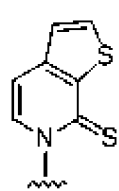


(TPT3) и одно или несколько вторых оснований не природного происхождения

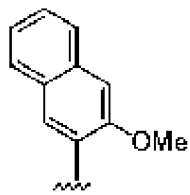


представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

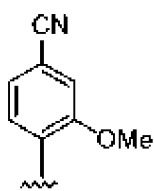
Вариант осуществления 345. Система по любому из вариантов осуществления 332-335, где одно или несколько первых оснований не природного происхождения выбраны из



(TPT3),



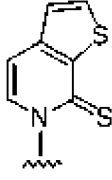
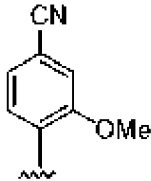
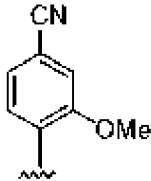
(NaM) и



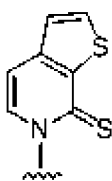
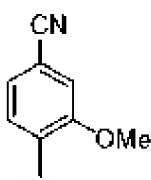
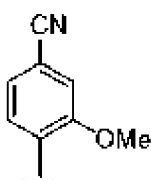
(CNMO), где волнистая линия

указывает на связь с рибозильной группой.

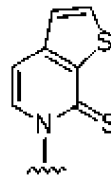
Вариант осуществления 346. Система по варианту осуществления 332, в которой по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в кодоне иРНК, где одно или несколько первых

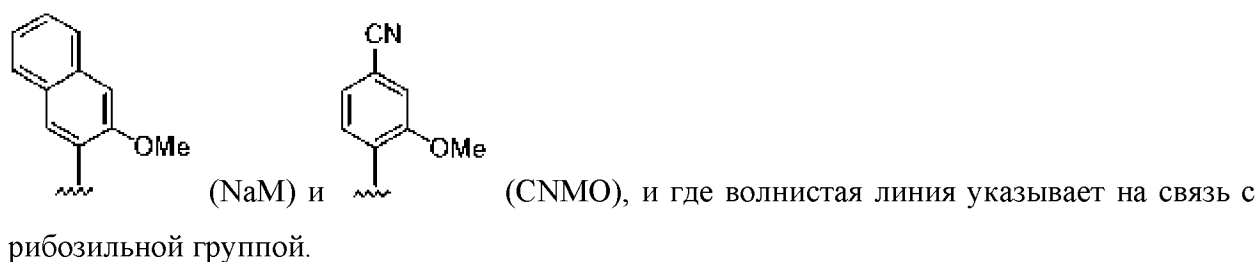
оснований не природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 347. Система по варианту осуществления 332, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК, где одно или несколько первых

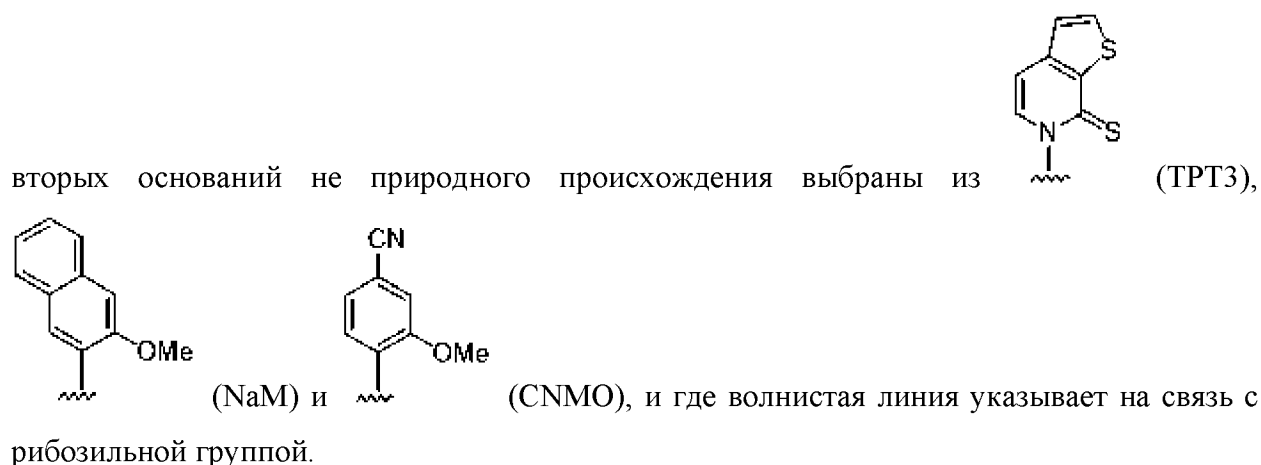
оснований не природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 348. Система по варианту осуществления 332, где, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в кодоне иРНК, где одно или несколько

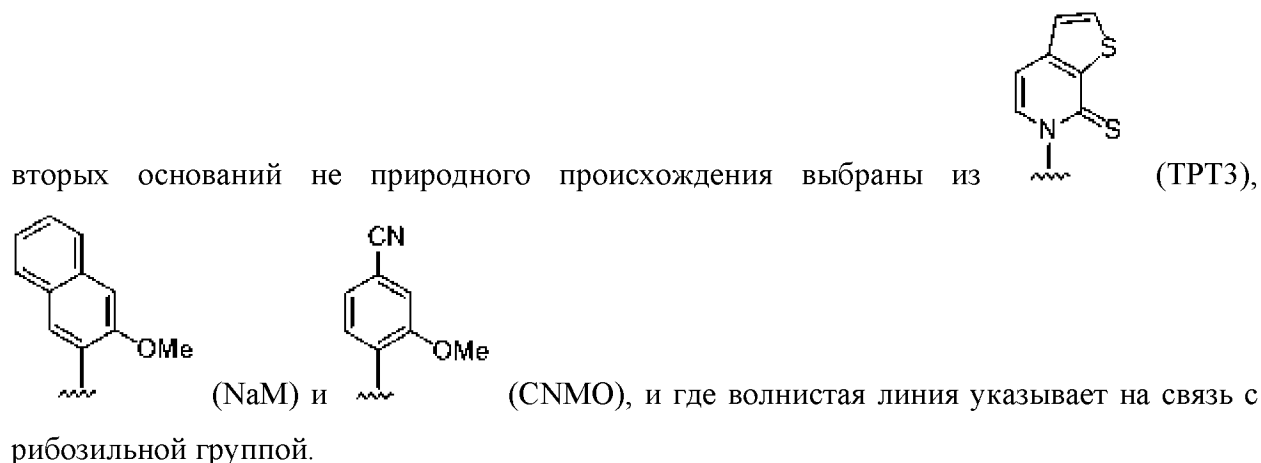
первых оснований не природного происхождения выбраны из  (TPT3),



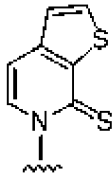
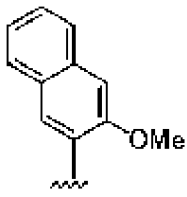
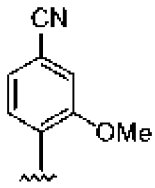
Вариант осуществления 349. Система по варианту осуществления 332, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в антикодоне тРНК, где одно или несколько



Вариант осуществления 350. Система по варианту осуществления 332, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в антикодоне тРНК, где одно или несколько



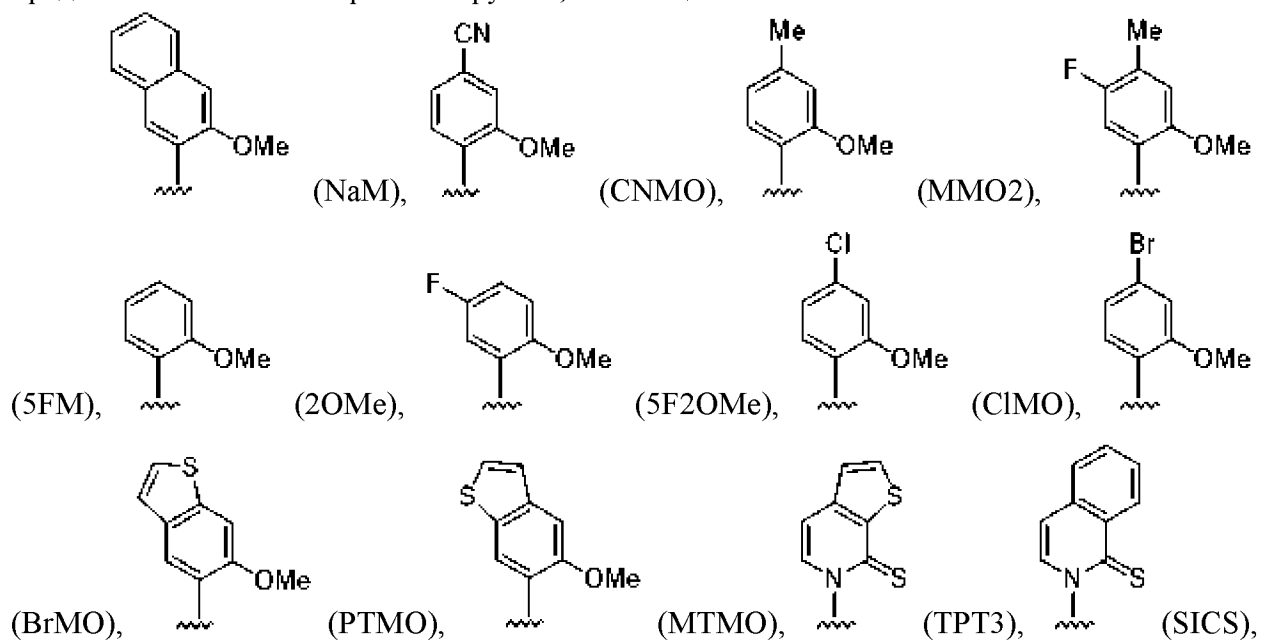
Вариант осуществления 351. Система по варианту осуществления 332, где, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N), и где одно или несколько вторых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в антикодоне тРНК, где одно или несколько

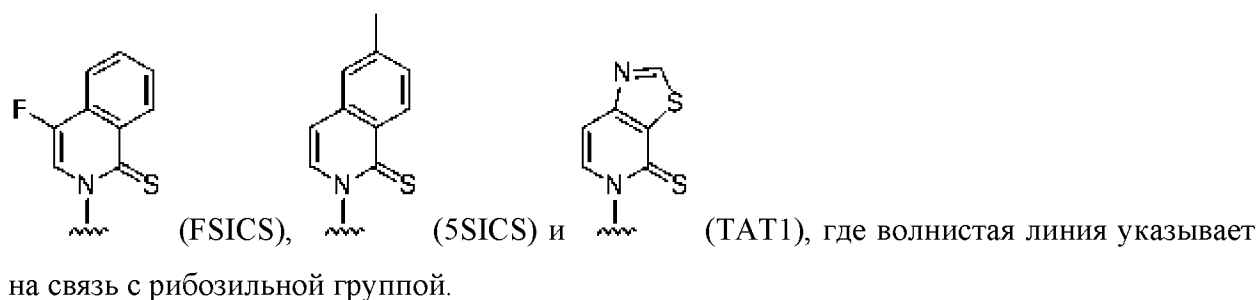
вторых оснований не природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 352. Система по варианту осуществления 332, где, по меньшей мере, один кодон и, по меньшей мере, один антикодон каждый независимо содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон содержит одно или несколько первых оснований не природного происхождения (X), расположенных в первом положении (X-N-N) кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в последнем положении (N-N-Y) антикодона.

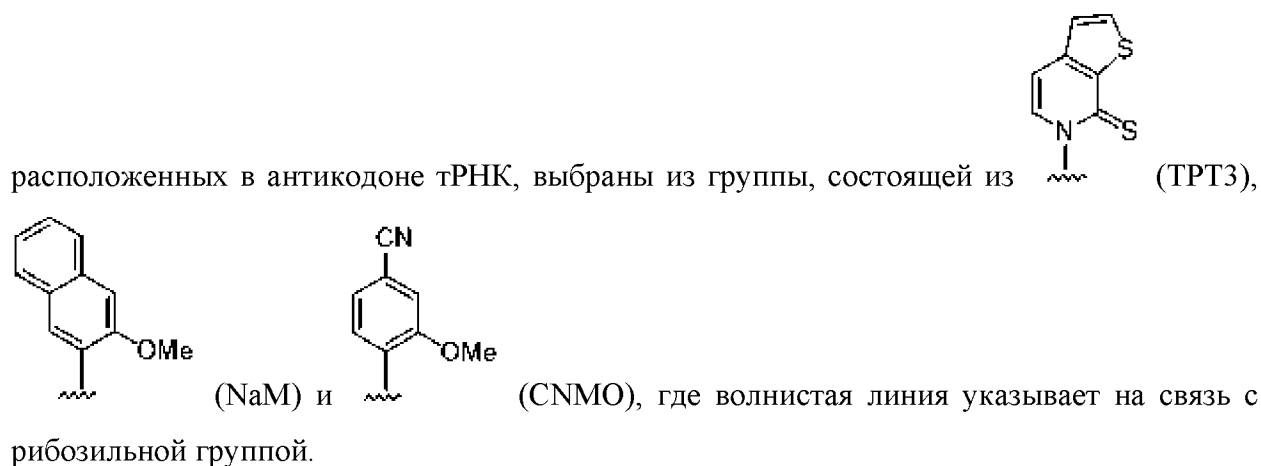
Вариант осуществления 353. Система по варианту осуществления 352, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 354. Система по любому из вариантов осуществления 352-353, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, представляют собой выбрано из группы, состоящей из

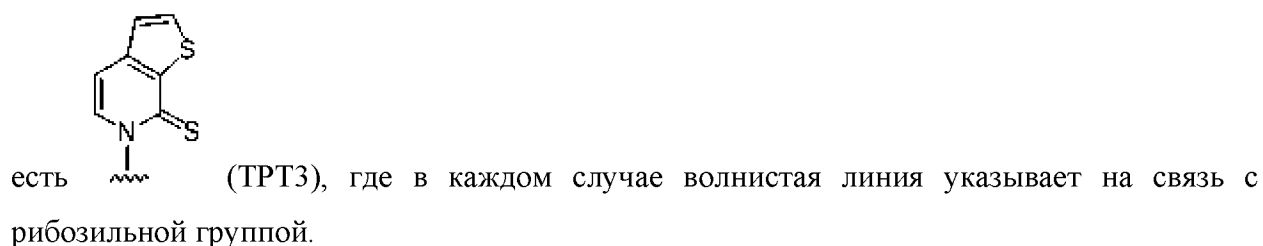
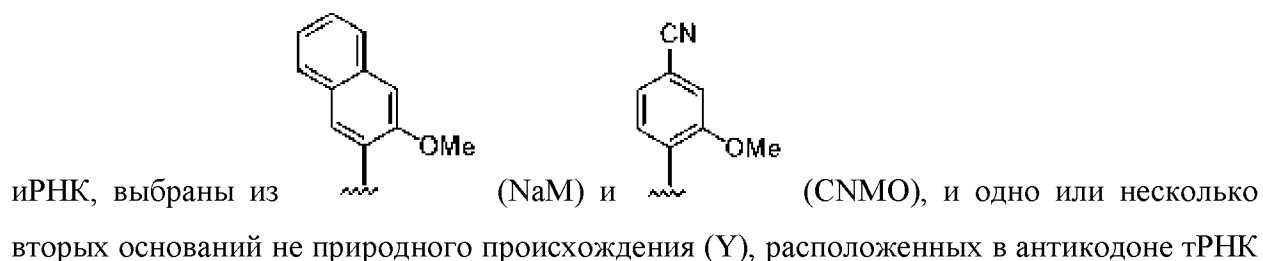




Вариант осуществления 355. Система по варианту осуществления 354, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



Вариант осуществления 356. Система по варианту осуществления 355, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне

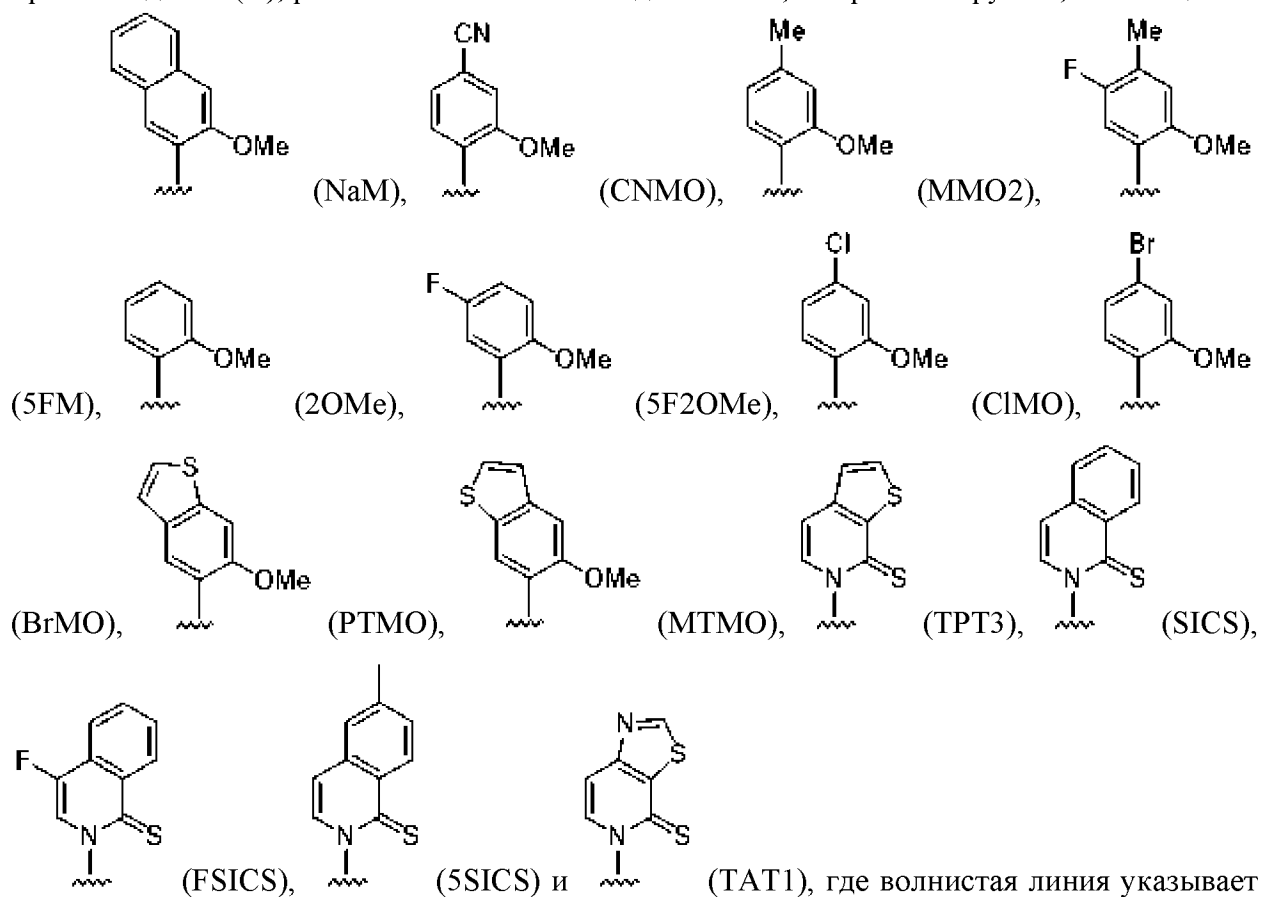


Вариант осуществления 357. Система по варианту осуществления 332, где, по меньшей мере, один кодон и, по меньшей мере, один антикодон, каждый независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в среднем положении (N-X-N), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в среднем

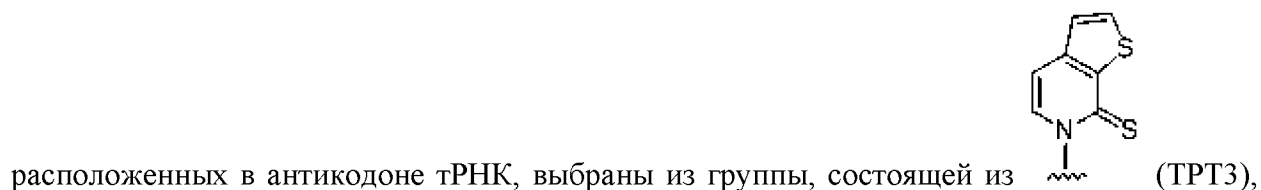
положении (N-Y-N) антикодона.

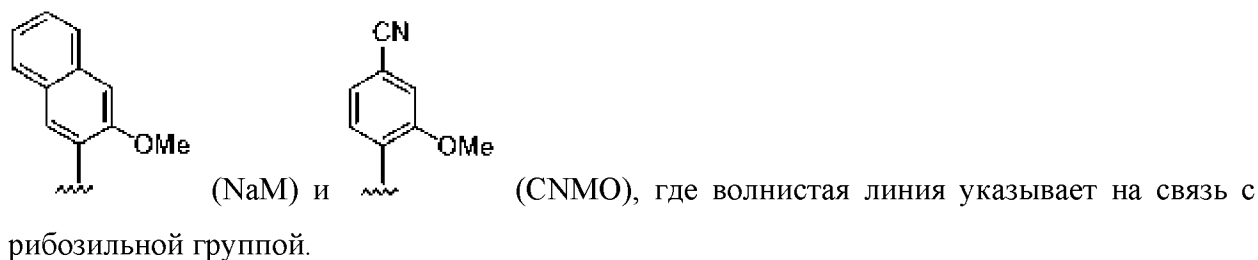
Вариант осуществления 358. Система по варианту осуществления 357, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 359. Система по любому из вариантов осуществления 357-358, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбрано из группы, состоящей из

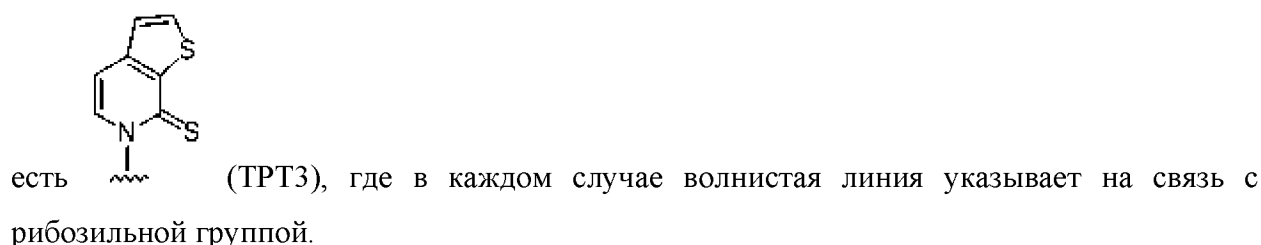
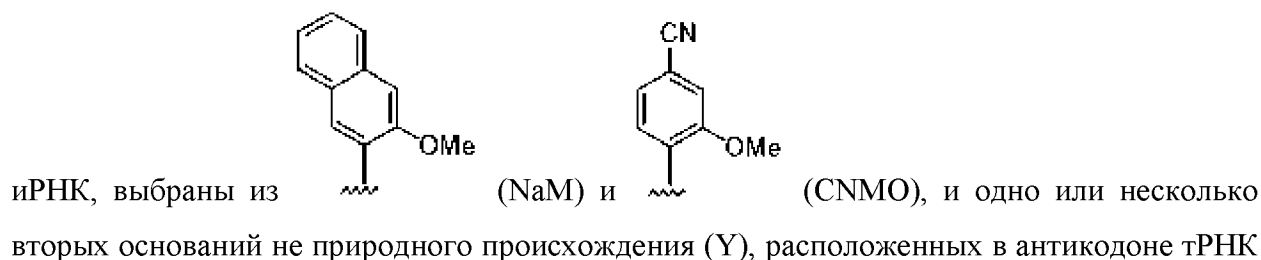


Вариант осуществления 360. Система по варианту осуществления 359, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),





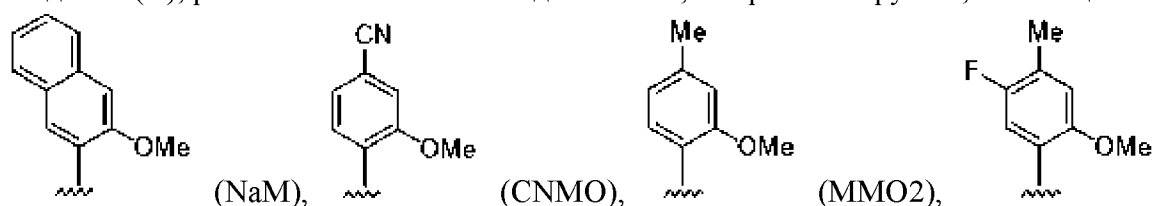
Вариант осуществления 361. Система по варианту осуществления 360, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне

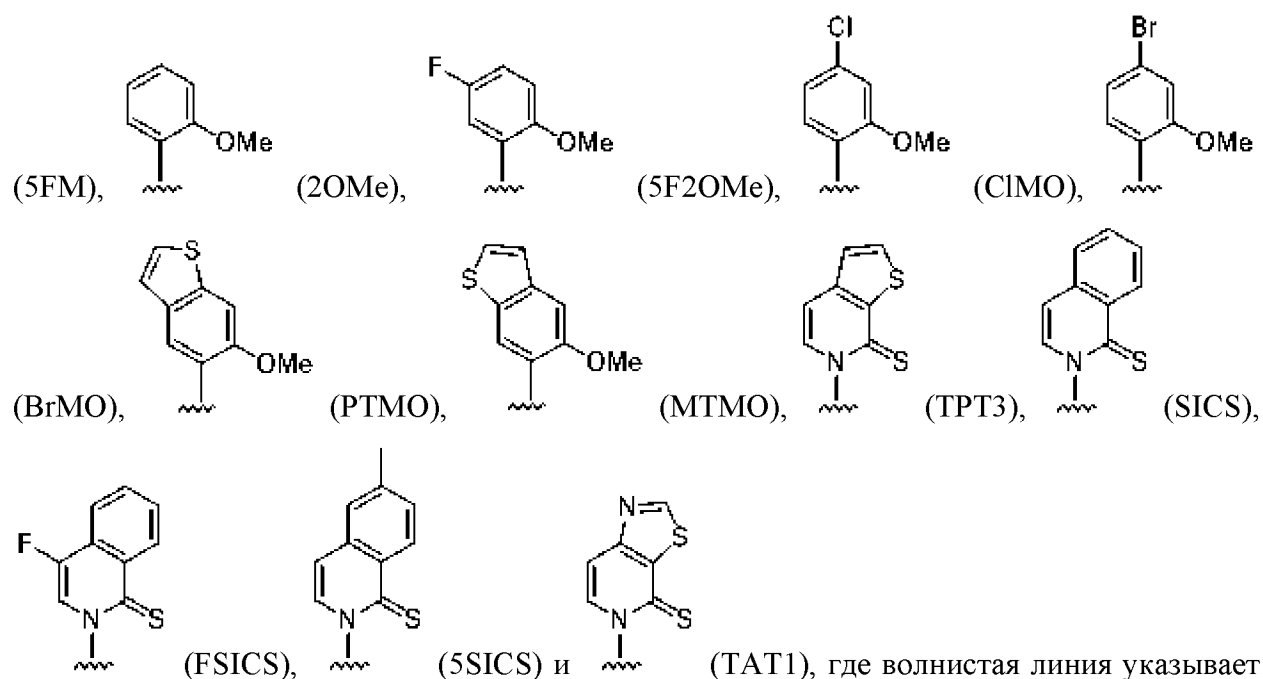


Вариант осуществления 362. Система по варианту осуществления 332, где, по меньшей мере, один кодон и, по меньшей мере, один антикодон, каждый независимо, содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенных в последнем положении (N-N-X), по меньшей мере, одного кодона, и по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в первом положении (Y-N-N) антикодона.

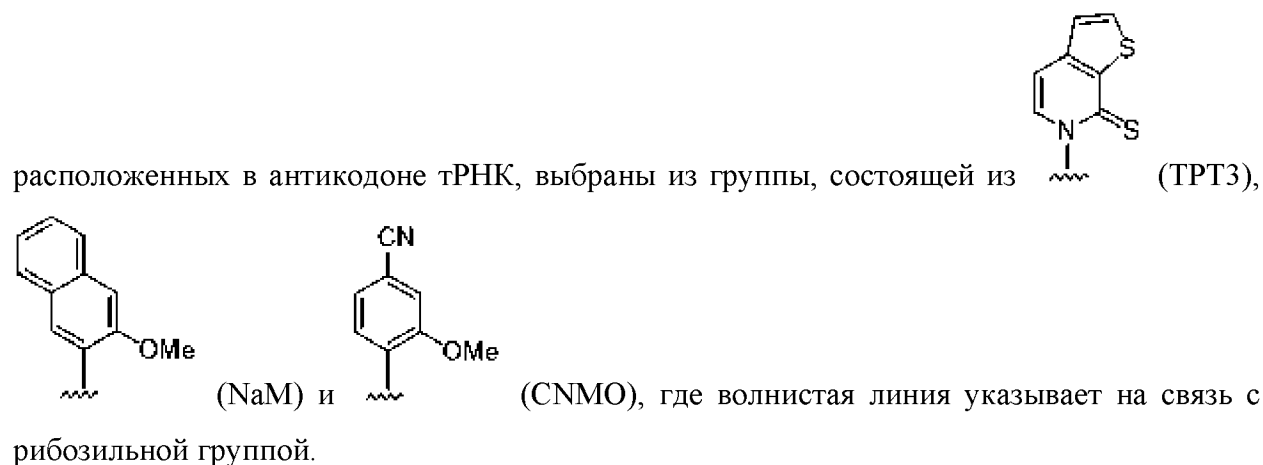
Вариант осуществления 363. Система по варианту осуществления 362, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

Вариант осуществления 364. Система по любому из вариантов осуществления 362-363, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

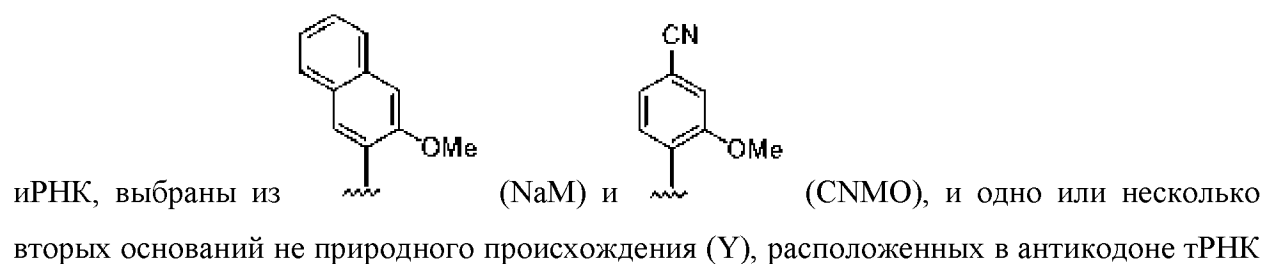


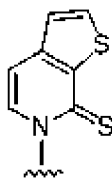


Вариант осуществления 365. Система по варианту осуществления 364, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y),



Вариант осуществления 366. Система по варианту осуществления 365, где одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне





представляет собой (TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

Вариант осуществления 367. Система по любому из вариантов осуществления 332-366, где, по меньшей мере, один кодон в иРНК выбран из AXC, GXC или GXU, где X представляет собой одно или несколько первых оснований не природного происхождения.

Вариант осуществления 368. Система по непосредственно предшествующему варианту осуществления, где, по меньшей мере, один антикодон в тРНК выбран из GYU, GYC и AYC, и Y представляет собой одно или несколько вторых оснований не природного происхождения.

Вариант осуществления 369. Система по варианту осуществления 368, где, по меньшей мере, один кодон в иРНК представляет собой AXC, и, по меньшей мере, один антикодон в тРНК представляет собой GYU.

Вариант осуществления 370. Система согласно варианту осуществления 368, где, по меньшей мере, один кодон в иРНК представляет собой GXC, и, по меньшей мере, один антикодон в тРНК представляет собой GYC.

Вариант осуществления 371. Система по варианту осуществления 368, где, по меньшей мере, один кодон в иРНК представляет собой GXU, и, по меньшей мере, один антикодон представляет собой AYC.

Вариант осуществления 372. Система по любому из вариантов осуществления 332-371, где тРНК получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 373. Система по любому из вариантов осуществления 332-372, где тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

Вариант осуществления 374. Система по любому из вариантов осуществления 332-373, которая находится в эукариотической клетке.

Вариант осуществления 374.1 Система по любому из вариантов осуществления 332-373, которая находится в клетке человека.

Вариант осуществления 375. Система по варианту осуществления 374.1, где клетка человека представляет собой клетку HEK293T.

Вариант осуществления 376. Система по любому из вариантов осуществления 332-373, которая находится в клетке млекопитающего.

Вариант осуществления 376.1 Система любого из вариантов осуществления 332-373, которая находится в клетке хомяка.

Вариант осуществления 377. Система по варианту осуществления 376.1, где клетка хомяка представляет собой клетку яичника китайского хомяка (CHO).

Вариант осуществления 377.1. Система по любому из вариантов осуществления 332-

377, где иРНК и тРНК стабилизированы к деградации в эукариотической клетке.

Вариант осуществления 377.2. Система по любому из вариантов осуществления 332-377.1, где полипептид продуцируется путем трансляции иРНК с использованием тРНК через рибосому, эндогенную для эукариотической клетки.

Вариант осуществления 377.3 Система по любому из вариантов осуществления 332-373, которая является *in vitro* или бесклеточной.

Вариант осуществления 378. Система по любому из вариантов осуществления 332-377.3, где аминокислота не природного происхождения:

представляет собой аналог лизина;
содержит ароматическую боковую цепь;
содержит азидогруппу;
содержит алкиновую группу; или
содержат альдегидную или кетонную группу.

Вариант осуществления 379. Система по любому из вариантов осуществления 332-378, где аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидозтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилфенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацетил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, O-аллилтирозина, O-метил-L-тирозина, O-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-O-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

Вариант осуществления 380. Система по варианту осуществления 379, где аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK).

Вариант осуществления 381. Система по любому из вариантов осуществления 332-380, где тРНК заряжена аминокислотой не природного происхождения.

Вариант осуществления 382. Способ по любому из вариантов осуществления 287-331, где иРНК и тРНК стабилизированы к деградации в эукариотической клетке.

Вариант осуществления 383. Способ по любому из вариантов осуществления 287-331 и 382, где полипептид получают путем трансляции иРНК с использованием тРНК с помощью рибосомы, эндогенной для эукариотической клетки.

ПРИМЕРЫ

Эти примеры представлены только в иллюстративных целях, и не для ограничения объема формулы изобретения, представленной в настоящем документе. Подробные способы представлены в настоящем документе в качестве последнего примера.

Пример 1: Трансляция кодонов не природного происхождения в клетках HEK293T

Плазмиды, кодирующие EGFP(AXC)¹⁵¹ и EGFP(GXC)¹⁵¹, конструируют с последовательностями CS2 3' и 5' UTR, фланкирующими кодирующую последовательность, для повышения стабильности иРНК. Кодоны AXC и GXC выбирают, поскольку было показано, что они хорошо декодируются в SSO *E. coli*. Желаемые иРНК и родственные тРНК получают с помощью реакций транскрипции *in vitro* с использованием T7 РНК полимеразы. ChPylRS вводят на плазмиде (pcDNA3.1_C211_IRES_mCherry), несущей бицистронную последовательность, кодирующую и ChPylRS, и mCherry маркер, связанный внутренним участком связывания рибосомы. Клетки HEK293T трансфицируют этой плазмидой, когда они достигают 50% конфлюэнтности. Клетки выращивают в течение 24 часов, чтобы обеспечить экспрессию ChPylRS, и затем в среду добавляют N6-((азидоэтокси)карбонил)-L-лизин (AzK) и клетки трансфицируют только иРНК в качестве контроля или иРНК и соответствующей родственной тРНК не природного происхождения. Клетки собирают через дополнительные 24 ч, и продуцирование EGFP в клетках, экспрессирующих mCherry маркер определяют количественно с помощью проточной цитометрии. В контроле без тРНК, трансфекция EGFP(AXC)¹⁵¹ и EGFP(GXC)¹⁵¹ иРНК приводит к низким, но обнаруживаемым уровням сигнала EGFP, предположительно в результате считывания кодонов не природного происхождения, когда их родственные тРНК отсутствуют. Напротив, клетки, трансфицированные как иРНК не природного происхождения, так и родственной тРНК не природного происхождения, демонстрируют повышенную флуоресценцию. В то время как увеличение было умеренным с EGFP(AXC)¹⁵¹, оно было более значительным с EGFP(GXC)¹⁵¹ (фиг. 5А).

Основываясь на относительно большем увеличении тРНК-зависимой флуоресценции, исследуют белок, продуцируемый конструкцией EGFP(GXC)¹⁵¹. Общий клеточный лизат подвергают клик-химии, стимулируемой штаммом, для присоединения красителя карбокситетраметилродамина (TAMRA) (DBCO-TAMRA), который, как было показано, изменяет электрофоретическую подвижность EGFP при анализе с помощью SDS-PAGE и, таким образом, позволяет оценить точность включения N6-((азидоэтокси)-карбонил)-L-лизина (AzK) вестерн-блоттингом. Отчетливый сигнал EGFP является очевидным (фиг. 5В) со сдвигом приблизительно на 70% для лизата, полученного из клеток, трансфицированных плазмидой синтазы, EGFP(GXC)¹⁵¹ иРНК и тРНК^{Pyl}(GYC), и выращивают в среде с добавлением N6-((азидоэтокси)-карбонил)-L-лизина (AzK). Напротив, сдвиг полосы от незначительного до отсутствующего наблюдают в лизате, полученном из клеток, трансфицированных без родственных тРНК не природного происхождения. Хотя низкий уровень экспрессии EGFP препятствует дальнейшей

характеризации, эти данные убедительно предполагают, что N6-((азидоэтокси)карбонил)-L-лизин (AzK) включается в EGFP через декодирование кодонов не природного происхождения с использованием тРНК с родственным антикодоном не природного происхождения.

Пример 2: Трансляция кодонов не природного происхождения в клетках CHO

Гетерогенную клеточную линию CHO-KS3, которая стабильно экспрессирует ChPylRS, конструируют с использованием системы рекомбинации FRT/Flp, что позволяет сократить трансфекцию до одной стадии котрансфекции РНК. Клетки CHO-KS3 трансфицируют EGFP(AXC)¹⁵¹, EGFP(GXC)¹⁵¹ или EGFP(GXC)¹⁵¹ иРНК и родственной тРНК; и N6-((азидоэтокси)-карбонил)-L-лизин (AzK) добавляют в среду для выращивания, когда клетки достигают 80% конфлюэнтности. Клетки собирают после однодневной инкубации, и затем непосредственно подвергают проточной цитометрии для обнаружения флуоресценции EGFP. Контрольные клетки, не имеющие родственную тРНК не природного происхождения, показывают аналогичные низкие, но обнаруживаемые уровни сигнала EGFP. Напротив, клетки, трансфицированные родственной тРНК не природного происхождения, демонстрируют значительно повышенную флуоресценцию, где EGFP(AXC)¹⁵¹ дает самый высокий сигнал флуоресценции на клетку, и EGFP(GXU)¹⁵¹ дает самый низкий, но во всех случаях флуоресценция была выше, чем у клеток HEK293T (фиг. 6А-6В).

Исследованные выше кодоны **NaM** выбирают, потому что они хорошо транслируются рибосомой *E. coli*. Напротив, рибосома *E. coli*, по-видимому, не способна транслировать кодоны, содержащие **TPT3**. Чтобы создать сравнительные отношения структура-активность между прокариотическими и эукариотическими рибосомами, EGFP(AYC)¹⁵¹, EGFP(GYC)¹⁵¹ и EGFP(GYU)¹⁵¹, а также родственные им тРНК tRNA^{Pyl}(GXU), tRNA^{Pyl}(GXC) и tRNA^{Pyl}(AXC) не природного происхождения, создают и используют для трансфекции клеток CHO-KS3. В отличие от SSO *E. coli*, все три кодона **TPT3** приводят к повышенной флуоресценции, когда клетки CHO-KS3 трансфицированы их родственными тРНК по сравнению с контролями, трансфицированными без тРНК, и фактически EGFP(GYU)¹⁵¹ достигает уровня флуоресценции, сходного с уровнем, наблюдаемым с аналогичным кодоном **NaM** (GXU) (фиг. 6А-6В).

Учитывая более высокие уровни экспрессии EGFP в клетках CHO-KS3, мы выбрали EGFP(AXC)¹⁵¹, EGFP(GXC)¹⁵¹, EGFP(GXU)¹⁵¹ и EGFP(GYC)¹⁵¹ для более количественной характеристики. EGFP аффинно очищают из клеточных лизатов с использованием тандемной С-концевой *Strep*-tag II и подвергают клик-химии с красителем DBCO-TAMRA, как описано выше. Затем очищенный EGFP анализируют с помощью вестерн-блоттинга. Из контрольных клеток, трансфицированных природной EGFP иРНК, наблюдают доминирующую полосу с более быстро мигрирующей, более слабой полосой (фиг. 6В). Более быстро мигрирующую полосу относят за счет разрушения *Strep* tag (данные не показаны). Как и ожидалось, ни одна полоса не показала сигнал TAMRA. При трансфекции каждой иРНК не природного происхождения ее родственной тРНК, наблюдают

аналогичный набор из двух полос, но обе смещены и показывают сигнал TAMRA. Эти результаты предполагают, что в клетках CHO, N6-((азидоэтокси)-карбонил)-L-лизин (AzK) включен в EGFP путем декодирования кодонов **NaM** или **TPT3** с родственными антикодонами не природного происхождения.

Для подтверждения правильного кодирования N6-((азидоэтокси)-карбонил)-L-лизина (AzK), жидкостную хроматографию - тандемную масс спектрометрию (ЖХ-МС/МС) используют для анализа белка, очищенного из клеток CHO-KS3, трансфицированных либо EGFP(GXC)¹⁵¹, либо EGFP(GYC)¹⁵¹ иРНК, и родственными им тРНК. EGFP очищают из трансфицированных клеток, как описано выше, и затем подвергают клик-химии, катализируемой медью, для присоединения 3-бутилилбензольной группы к AzK, чтобы способствовать МС анализу. Продукт реакции очищают с помощью SDS-PAGE и вырезают полосу между 25 кДа и 32 кДа, которая, на основе предыдущих анализов сдвига в геле, включает как сдвинутые, так и не сдвинутые полосы EGFP. Белки, восстановленные из срезов геля, переваривают трипсином и подвергают нано-ЖХ-МС/МС анализу. Пептидные фрагменты, содержащие аминокислотный сайт 151 EGFP, определяют с массами, соответствующими продукту клик-реакции, что подтверждает специфическое включение N6-((азидоэтокси)карбонил)-L-лизина (AzK) в сайт 151. Не модифицированный пептид не определяют, и, хотя оно не является количественным, это наблюдение подтверждает включение N6-((азидоэтокси)карбонил)-L-лизина (AzK) и предполагает, что это происходит, по меньшей мере, с разумной точностью. В то время как более тщательный анализ контекста последовательности остается для изучения, эти данные демонстрируют, что рибосомы млекопитающих, в отличие от их аналогов *E. coli*, способны декодировать кодоны не природного происхождения, содержащие либо **NaM**, либо **TPT3**.

Ранее было показано, что SSO *E. coli* также способен транслировать несколько кодонов с нуклеотидом не природного происхождения **NaM** в третьем положении, включая кодон AGX. Однако, в отличие от второго положения, декодирование происходит либо с «гетероспариванием» тРНК^{PyI}(YCT), либо «самоспариванием» тРНК^{PyI}(XCT) (фиг. 5). Самоспаривание **NaM-NaM** в третьем положении может быть усилено образом, подобным неоднозначному спариванию природных кодонов в третьем положении. Чтобы изучить декодирование со самоспаривающимися родственными тРНК в клетках млекопитающих, кодон AGX затем тестируют в том же контексте иРНК. Клетки CHO-KS3 трансфицируют только иРНК EGFP(AGX)¹⁵¹, или котрансфицируют вместе с тРНК^{PyI}(YCT) или тРНК^{PyI}(XCT). Как и в случае с кодонами не природного происхождения во втором положении, проточная цитометрия выявила небольшое количество сквозной экспрессии EGFP в клетках, трансфицированных без каких-либо тРНК. Котрансфекция с тРНК^{PyI}(YCT) приводит к значительному увеличению флуоресценции, в то время как котрансфекция с тРНК^{PyI}(XCT), самоспаривающейся тРНК, приводит к еще большему увеличению флуоресценции (фиг. 6А). Затем используют тот же анализ белкового сдвига, описанный выше, для дальнейшей оценки EGFP, продуцируемого из кодона не природного происхождения AGX. Сдвинутые полосы определяют в белках, очищенных из клеток,

котрансфицированных либо с тРНК^{Pyl}(YCT), либо с тРНК^{Pyl}(XCT) (фиг. 6В). В обоих случаях, снова наблюдают две сдвинутые полосы, при отсутствии или незначительной видимой не сдвинутой полосы. Эти результаты демонстрируют что, по меньшей мере, с кодоном AGX, декодирование через либо гетероспаривание, либо самоспаривание является, по меньшей мере, достаточно эффективным.

Результаты с кодонами TPTЗ демонстрируют явные различия между прокариотическими и эукариотическими рибосомами. Для дальнейшего сравнения этих рибосом, трансляция кодонов с нуклеотидом не природного происхождения в первом положении, который рибосома *E. coli*, по-видимому, не может декодировать, иРНК EGFP(XCC)¹⁵¹ и EGFP(YCC)¹⁵¹ продуцируют *in vitro* и трансфицируют в клетки CHO-KS3 без или с их родственной тРНК не природного происхождения, тРНК^{Pyl}(GGY) или тРНК^{Pyl}(GGX), соответственно. Анализ с использованием проточной цитометрии показывает небольшое количество считывания, когда тРНК не добавлены в обоих случаях, где EGFP(YCC)¹⁵¹ дает относительно более высокий сигнал EGFP, чем EGFP(XCC)¹⁵¹. Когда добавляют соответствующие тРНК, небольшое увеличение сигнала EGFP наблюдают с EGFP(XCC)¹⁵¹, но не наблюдают значительное увеличение сигнала EGFP с EGFP(YCC)¹⁵¹ (фиг. 6). В обоих случаях, выходы EGFP являются слишком низкими для вестерн-блоттинга. Эти данные предполагают, что, как и в случае с рибосомой *E. coli*, кодоны не природного происхождения в первом положении плохо декодируются. Вероятно, это связано с взаимодействием А-минор типа I, при котором рибосома выбирает структуру, подобную Уотсону-Крику, в первом положении кодона.

Пример 3: Соотношение экспрессии белка между иРНК с CYBA UTR и иРНК с CS2 UTR

Исследуют использование чередующихся 5' и 3' UTR. Сообщалось, что комбинированное использование 5' и 3' UTR CYBA увеличивает продуцирование белка, не влияя на его период полужизни в клетках человека. Конструируют последовательности EGFP со всеми 9 протестированными выше кодонами не природного происхождения, с UTR CS2, замененными на UTR CYBA (CYBA-EGFP(NX/YN)¹⁵¹). Клетки CHO-KS3 трансфицируют этими вновь сконструированными иРНК без или с родственной тРНК не природного происхождения. Затем клетки анализируют с помощью проточной цитометрии, и результаты сравнивают с их аналогами с CS2 UTR. Данные проточной цитометрии показали, что во всех случаях с CYBA UTR продуцируется меньше белка, чем с их аналогами CS2. Для клеток, трансфицированных CYBA-EGFP(GXC)¹⁵¹ и CYBA-EGFP(GYC)¹⁵¹, мы также оцениваем точность декодирования кодонов не природного происхождения, используя анализ сдвига в геле, как описано выше. Наблюдаемые сдвиги были аналогичны сдвигам, наблюдаемым с аналогами CS2 UTR (EGFP(GXC)¹⁵¹ и EGFP(GYC)¹⁵¹), соответственно (фиг. 7А-В), демонстрируя, что на точность декодирования существенно не влияет изменение фланкирующих UTR.

В то время как сниженный уровень экспрессии, наблюдаемый с CYBA UTR, может быть связан с использованием клеток хомяка вместо клеток человека, мы также отметили,

что величина эффекта, совершенно неожиданно, значительно отличалась для разных кодонов не природного происхождения. При трансфекции их родственными неприродными тРНК (самоспаривающуюся тРНК используют с кодоном AGX), кодоны ХСС, YCC, GXU и GYU с UTR CYBA демонстрируют уровни экспрессии, которые составляют ~60% от их аналогов CS2, в то время как уровни экспрессии для кодонов AXC, AYC, GXC, GYC и AGX с UTR CYBA составляют только ~30% от их аналогов CS2 (**фиг. 7A-D**). Янтарную конструкцию CYBA-EGFP(TAG)¹⁵¹ и природную конструкцию CYBA-EGFP(TAC)¹⁵¹ используют в качестве контролей. CYBA-EGFP(TAG)¹⁵¹ и CYBA-EGFP(TAC)¹⁵¹ демонстрируют уровни экспрессии, которые составляют ~60% и ~80% их аналогов CS2 UTR.

Чтобы проверить, может ли этот эффект UTR, зависимый от кодона не природного происхождения, возникать из-за различий в стабильности иРНК, уровень иРНК через 8 ч после трансфекции сравнивают с уровнем через 4 ч после трансфекции для EGFP(UAC)¹⁵¹, EGFP(GXC)¹⁵¹, EGFP(GXU)¹⁵¹, CYBA-EGFP(UAC)¹⁵¹, CYBA-EGFP(GXC)¹⁵¹ и CYBA-EGFP(GXU)¹⁵¹ с использованием обратной транскрипции в сочетании с количественной ПЦР. Различия, наблюдаемые в деградации между этими разными конструкциями, не учитывают резкие различия в соотношениях, описанные выше (**фиг. 6**), и, следовательно, должны быть ответственны другие факторы. Считается, что один из способов, которым UTR влияют на трансляцию, заключается в регулировании эффективности рекрутирования рибосом. Однако трудно объяснить, как это может влиять на трансляцию кодона, который удален от 5' или 3' UTR (в настоящем случае на, по меньшей мере, 350 нуклеотидов). Интересно, известно, что в одной клетке существует несколько субпопуляций рибосом, и они могут, например, дифференцироваться по вариабельным способностям элонгации трансляции. В отличие от трансляции природных кодонов, это в принципе может иметь более значительное влияние на то, как рибосома обрабатывает разные кодоны не природного происхождения, возможно, аналогично нашему наблюдению, что рибосомы из прокариотов и эукариотов по-разному декодируют разные кодоны не природного происхождения. Требуются дальнейшие эксперименты для того, чтобы опираться на эту увлекательную возможность.

Результаты, описанные в настоящем документе, демонстрируют, что кодоны не природного происхождения могут быть декодированы, по меньшей мере, с разумной эффективностью и точностью как в клетках HEK293T, так и в клетках CHO. Интересно, что распознавание эукариотическими рибосомами показывает как сходство, так и различия с распознаванием, опосредованным рибосомой *E. coli*. Кодоны первого положения ХСС и YCC не могут быть декодированы с хорошей эффективностью ни в клетках *E. coli*, ни в клетках CHO; **NaM** кодоны AXC, GXC и GXU во втором положении могут быть декодированы с хорошей эффективностью как в клетках *E. coli*, так и в клетках CHO; **TRP3** кодоны AYC, GYC и GYU во втором положении кодона не могут быть декодированы в клетках *E. coli*, но, что интересно, могут быть декодированы в клетках CHO; и кодон AGX в третьем положении может быть декодирован как в клетках *E. coli*, так и в клетках CHO и

его родственной гетероспаренной тРНК, и его не родственной самоспаривающейся тРНК.

Пример 4: Способы

Материалы и способы, использованные в примерах 1-3, следующие:

Материалы. Плазмиды и праймеры, использованные в примерах 1-4, можно найти в таблицах 1 и 2. Праймеры и природные олигонуклеотиды приобретены у IDT (Coralville, Iowa). Секвенирование проводит Genewiz (San Diego, CA). Плазмиды очищают с использованием коммерческого набора miniprep (Продукт № D4013, Zymo Research; Irvine, CA). Продукты ПЦР очищают с использованием коммерческого набора для очистки ДНК (D4054, Zymo Research) и количественно оценивают с помощью планшетного ридера Infinite M200 Pro (TECAN). Все эксперименты, включающие виды РНК, проводят с реагентами, не содержащими РНКазу, наконечниками пипеток, пробирками и перчатками, чтобы избежать загрязнения. Нуклеозиды dNaM, dTPT3, NAM, TPT3, d5SICS и dMMO2bio синтезируют (WuXi AppTec; Shanghai, China) и трифосфорилируют (TriLink BioTechnologies LLC; San Diego, CA и MyChem LLC; San Diego, CA) коммерчески. Вне олигонуклеотиды на природного происхождения синтезируют в Biosearch Technologies (Petaluma, CA) с очисткой ВЭЖХ.

Конструирование плазмид синтетазы. Последовательность химеры синтетазы ChPylPS_C211 клонируют из pGEX_ChPylRS, которая описана в Fischer et al., Nat. Chem. Biol. 16:570-576 (2020). pcDNA3.1_C211_IRES_mCh получают клонированием последовательностей ChPylRS, IRES и mCherry, одну за другой, в вектор pcDNA3.1 с использованием ряда рестрикционных ферментов.

Конструирование матриц EGFP и тРНК. Плазмиды EGFP матрицы, pUCCS2_EGFP(NNN) и pUCCYBA_EGFP(NNN), получают с помощью сборки Golden Gate, как описано ранее, но с контекстом последовательности EGFP вместо контекста sfGFP (см. Zhang et al., Nature 551:644-647 (2017)). Вставки, использованные во всех сборках Golden Gate, представляют собой продукты ПЦР, созданные с синтезированными dNaM-содержащими олигонуклеотидами и праймерами YZ73 и YZ74 (см. таблицу 1). Плазмиды pUCCS2_EGFP(NNN) и pUCCYBA_EGFP(NNN) получают после сборки Golden Gate и количественно оценивают с использованием Qubit (ThermoFisher). Плазмиды EGFP матрицы (2 нг) используют в реакции ПЦР, создающей матрицу, с праймерами ED101 и AZ38 для pUCCS2_EGFP(NNN) и праймерами ED101 и AZ87 для pUCCYBA_EGFP(NNN). Продукты ПЦР подвергают перевару DpnI и затем очищают с получением EGFP матриц для транскрипции *in vitro* (см. ниже). Матрицы тРНК получают с помощью прямой ПЦР из синтезированных dNaM-содержащих олигонуклеотидов с праймерами AZ01 и AZ67. Продукты ПЦР очищают с получением матриц тРНК для транскрипции *in vitro*.

Анализ биотинового сдвига. Сохранение пары оснований не природного происхождения в матрицах видов РНК анализируют, как описано в предыдущей работе, с использованием d5SICSTP и dMMO2bio-TP с праймерами YZ73 и YZ7 (см. Zhang et al., Nature 551:644-647 (2017)). Изображения количественно оценивают с использованием Image Lab (BioRad). Сохранение пар оснований не природного происхождения

нормализуют путем деления доли исходного сдвига каждого образца на долю исходного сдвига синтезированного dNaM-содержащей олигонуклеотидной матрицы, используемой в сборке Golden Gate при конструировании EGFP плазмиды.

In vitro транскрипция иРНК EGFP. Матрицы (500-1000 нг) используют в каждой реакции in vitro транскрипции (HiScribe T7 ARCA с Tailing, E2060S, New England Biolabs, (NEB)) с или без 1,25 мМ рибонуклеотрифосфата не природного происхождения, соответственно, с последующей очисткой (D7010, Zymo Research). Продукты иРНК количественно оценивают с помощью Qubit, и затем хранят в 5 мкг аликвотах при -80°C.

In vitro транскрипция тРНК. Матрицы (500-1000 нг) используют в каждой реакции in vitro транскрипции (T7 РНК полимеразы, E0251L, NEB) с или без 2 мМ рибонуклеотрифосфата не природного происхождения, соответственно, с последующей очисткой (D7010, Zymo). тРНК продукты количественно определяют с помощью Qubit, и затем подвергают рефолдингу (95°C в течение 1 мин, 37°C в течение 1 мин, 10°C в течение 2 мин). Все тРНК хранят в аликвотах по 1800 нг при -80°C.

Конструирование стабильной клеточной линии. Плазмиду, содержащую синтетазу pcDNA3.1_FRT_HygroResist_C211_IRES_mCherry, получают путем замены кассеты резистентности к канамицину KanR в pcDNA3.1_C211_IRES_mCherry кассетой резистентности к гигромицину HygroResist путем клонирования с лигированием тупых концов. Гетерогенную клеточную линию CHO-KS3 модифицируют для стабильной экспрессии ChPylRS C211 с использованием системы Flp-In™ T-REx™ (ThermoFisher) в соответствии с инструкциями производителя. Исходные клетки Flp-in™ CHO-K1 выделяют в культуре 10% FBS, 1% PS DMEM/F12. Клетки котрансфицируют pOG44 и pcDNA3.1_C211_IRES_mCherry (контроль) или pcDNA3.1_FRT_HygroResist_C211_IRES_mCherry. Успешные рекомбинантные клетки селективируют с 100 мкг/мл гигромицина В (Sigma Aldrich) в течение двух недель (обновление среды для культивирования клеток каждые четыре дня), пока все клетки в контрольной группе не погибают. Клетки, трансфицированные pcDNA3.1_FRT_HygroResist_C211_IRES_mCherry, затем открепляют трипсиновым (25200056, Life Technology Invitrogen) переваром (5 мин при 37°C) и пассируют в течение еще двух циклов со средой для культивирования клеток, содержащей 100 мкг/мл гигромицина В.

Трансфекция клеток. Свежую клеточную культуру, содержащую 1 мМ AzK, добавляют в планшеты для культивирования клеток после истощения предыдущей среды. Для трансфекции РНК, клетки трансфицируют с видами РНК с использованием Lipofectamine MessengerMax (ThermoFisher) в соответствии с руководством по реагентам. Для каждого эксперимента по трансфекции, 300 нг иРНК и 900 нг тРНК каждый смешивают с 0,75 мкл липофектаминаминовыми реагентами и добавляют в культуру клеток (1 лунка 24-луночного плоскодонного полистиролового микролуночного планшета) по отдельности. Для трансфекции ДНК, клетки трансфицируют с видами ДНК с использованием Липофектамина. 3000 (LMRNA008, ThermoFisher) в соответствии с руководством по

реагентам. Для каждого эксперимента по трансфекции, 500 нг плазмиды ДНК смешивают с 1,5 мкл липофектаминных реагентов и добавляют к культуре клеток (1 лунка 24-луночного планшета). В некоторых случаях, клетки трансфицируют в 12-луночном планшете, объемы реагентов для трансфекции и РНК удваивают.

Проточная цитометрия. Клетки отделяют трипсиновым переваром (5 мин при 37°C), и затем промывают 1' физиологическим раствором с фосфатным буфером по Дульбекко (DPBS). Затем клетки собирают и разводят в буфере для сортировки (1' DPBS с 1% FBS) и затем анализируют с помощью проточной цитометрии для определения сигнала EGFP с использованием аналитического проточного цитометра LSR II (BD; сигнал EGFP обнаруживают с помощью лазера 488 нм и фильтра 530/30).

Приготовление цельноклеточного лизата. Клетки из экспериментов по трансфекции отделяют трипсиновым переваром (5 мин при 37°C) с последующей промывкой DPBS. Затем клетки собирают и лизируют с использованием M-PER (78503, Thermo Fisher) с добавлением ингибитора протеазы HALT (78430, Thermo Fisher) в соответствии с инструкциями по реагентам. Лизаты подвергают суперфильтрации с использованием центрифужных фильтров (Amicon Ultra - 0,5 мл центрифужные фильтры, 10 кДа NMWL, UFC501024, Millipore) для удаления не включенного AzK. Лизаты промывают DPBS, содержащим HALT (3). Лизаты концентрируют до объема 20 мкл на заключительной стадии промывки. Всю суперфильтрацию проводят при 14000 об/мин в течение 10 мин при 4°C (5415C, Eppendorf).

Аффинная очистка EGFP. Клетки, собранные в экспериментах по трансфекции, лизируют с использованием M-PER с добавлением ингибитора протеазы HALT в соответствии с руководствами по реагентам. Концентрацию EGFP (флуоресценция е.а.) в образцах лизата определяют с использованием планшетного ридера Infinite M200 Pro и стандартной кривой EGFP. Лизат, содержащий 200 нг эквивалента EGFP, разводят в 200 мкл буфера W (50 mM HEPES pH 8, 150 mM NaCl, 1 mM ЭДТК) и смешивают с 10 мкл магнитных гранул Strep-Tactin (5% (об./об.) суспензия MagStrep 'Type 3' XT гранул, продукт № 2-4090-002, IBA Lifesciences; Goettingen, Germany). Очистку проводят в соответствии с руководством по реагентам с увеличенным временем связывания (2 ч при 4°C). EGFP не элюируют из гранул. Конъюгат гранула-EGFP используют непосредственно в следующих экспериментах.

Клик-реакция на EGFP. Клик-реакции проводят, как описано в предыдущей работе (см. Zhang et al., Nature 551:644-647 (2017)) с модификациями. Коротко, конъюгат гранула-EGFP со стадии аффинной очистки разводят в 20 мкл DPBS. Смесь инкубируют с 25 мкМ TAMRA-DBCO (продукт № A131, Click Chemistry Tools; Scottsdale, AZ) в течение 1 ч при 37°C в темноте. Альтернативно, конъюгат гранула-EGFP, полученный на стадии аффинной очистки, разводят в 20 мкл DPBS. Смесь инкубируют с 2 mM трис(3-гидроксипропилтриазилилметил)амин (THPTA) (CAS 760952-88-3, Sigma-Aldrich), 1 mM CuSO₄, 15 mM аскорбата натрия (CAS 134-03-2, Sigma-Aldrich) и 0,5 mM 4-фенил-1-бутина (CAS 16520-62-0, Sigma-Aldrich) в течение 1 ч при 37°C в темноте. Клик-реакцию

обработанного цельноклеточного лизата проводят путем инкубации 20 мкл суперфильтрованного клеточного лизата с 25 мкМ йодацетамида (CAS 144-48-9, Sigma-Aldrich) в течение 1 ч при 37°C с последующей инкубацией полученной смеси с 25 мкМ DBCO-TAMRA в течение 1 ч при 37°C в темноте.

Вестерн-блоттинг анализ белкового сдвига. Вестерн-блоттинг анализ белкового сдвига проводят, как описано в предыдущей работе 2, с некоторыми модификациями. Коротко, смесь для клик-реакции кипятят непосредственно в красителе, загружающем 1' белок (250 мМ Tris-HCl, 30% (об./об.) глицерин, 2% (масс./об.) SDS) при 95°C в течение 15 минут, и продукты разделяют с помощью SDS-PAGE (с использованием 5% концентрирующего геля (масс./об.) акриламида:бис-акриламида 29:1 (Fisher), 0,125 М TrisHCl и 0,1% SDS, pH 6,8 (ProtoGel Stacking Buffer, National Diagnostics)); и 15% разделяющего геля (масс./об.) акриламида:бис-акриламида 29:1 (Fisher), 0,375 М Tris-HCl и 0,1% SDS, pH 8,8 (ProtoGel Resolving Buffer, National Diagnostics); спейсера 1,5 мм Mini-PROTEAN Short Plates (Bio-Rad)) с набором маркерных белков (Color Prestained Protein Standard, Broad Range, NEB). Гели запускают при 60 В в течение 30 мин, и затем при 135 В в течение примерно 3 ч в буфере SDS-PAGE (25 мМ основание Tris, 200 мМ глицин, 0,1% (масс./об.) SDS). Затем полосы переносят на мембрану PVDF (0,2 мкм, Bio-Rad) полусухим переносом с буфером, содержащим 20% (об./об.) MeOH, 50 мМ основания Tris, 400 мМ глицина, 0,0373% (масс./об.) SDS, при 22 В в течение 21 мин. Мембраны блокируют 5% (масс./об.) нежирного молока в PBS-T (PBS pH 7,4, 0,01% (об./об.) Tween-20) в течение 1-2 ч при комнатной температуре с последующей инкубацией с анти-GFP антителом кролика (продукт № G1544, партия 046M4871V, Sigma-Aldrich; 1:3000 в PBS-T) при 4°C в течение ночи. Затем мембраны промывают в течение 2,5 мин PBS-T с последующей инкубацией с конъюгированным анти-кроличьим антителом козы Alexa Fluor 647 (продукт № A32733, партия #SD250298, Thermo Fisher Scientific; 1:20000 в PBS-T) в течение 1 ч при комнатной температуре. Мембраны промывают в течение 3,5 мин PBS-T и визуализируют с помощью формирования изображения на люминесцентном фосфорном покрытии (Typhoon 9410; Build S4 410 5.0.0409.0700, GE Healthcare Life Sciences) с разрешением 50 мкм; лазерным возбуждением 532 нм и фильтром эмиссии 580/30 нм с 400 В PMT для TAMRA; лазерным возбуждением 622 нм и фильтром эмиссии 670/30 нм с 500 В PMT для Alexa Fluor 647. Изображения псевдоокрашивают и накладывают с помощью ImageJ, полосы количественно определяют с помощью Image Lab (Bio-Rad).

Масс-спектрометрия. Конъюгаты гранул-EGFP, соединенные с 4-фенил-1-бутином, кипятят непосредственно с красителем, загружающим 1' белок, при 95°C в течение 15 минут, и подвергают SDS-PAGE, по существу, как для вестерн-блоттинга сдвига белка, описанного выше, с набором маркерных белков. Гели прогоняют при 60 В в течение 30 мин, и затем при 135 В в течение примерно 30 мин в буфере SDS-PAGE. Полосы геля между 25 кДа и 32 кДа вырезают и собирают с последующим восстановлением (10 мМ ДТТ), алкилированием (55 мМ йодацетамида) и переваром с использованием трипсина. Затем образцы анализируют с помощью нано-ЖХ-МС/МС, как описано ранее (см. Powers

et al., J. Bacteriol. 193:340-348 (2011)). Коротко, зависимые от данных MS/MS получают с помощью масс-спектрометра с линейной ионной ловушкой Thermo FiN-Nigan LTQ с использованием самодельного источника наноэлектрораспыления при 2 кВ на конце. После одного MS спектра проводят 4 сканирования MS/MS на наиболее распространенных ионах после применения списка динамического исключения. Tandемные масс-спектры экстрагируют с использованием программного обеспечения Xcalibur. Все образцы MS/MS анализируют с использованием Mascot (версия 2.1.04; Matrix Science, London, United Kingdom) с предоставленной последовательностью EGFP, принимая во внимание фермент перевара трипсин.

Количественная масс-спектрометрия с высоким разрешением интактных белков. Масс-спектрометрию интактных белков проводят, как описано ранее (см. Feldman et al., J. Am. Chem. Soc. 141:10644-10653 (2019)). Очищенный белок EGFP (5 мкг) разводят водой (со степенью очистки для масс спектрометрии) и обессоливают суперфилтрацией (Amicon Ultra - 0,5 мл центрифужные фильтры, 10 кДа NMWL, UFC501024, Millipore). Затем обессоленный белок вводят (6 мкл, ~250 нг) в Waters IClass LC, соединенный с Waters G2-XS TOF. Условия потока следующие: 0,4 мл/мин 50:50 смеси вода:ацетонитрил плюс 0,1% муравьиной кислоты. Ионизацию проводят ИЭР+, данные собирают между m/z 500 и m/z 2000. Спектральное объединение проводят по основной части пика, и объединенный спектр был подвергают деконволюции с использованием Waters MaxEnt1.

Анализ распада иРНК. Для каждой протестированной иРНК, 2 лунки из 12-луночного планшета с клетками CHO-KS1 трансфицируют 600 нг иРНК и 1800 нг соответствующей тРНК с последующим добавлением 1 мМ AzK к культуре клеток. После 4-ч инкубации, обе лунки клеток дважды промывают DPBS, и затем клетки в 1 лунке собирают с использованием реагента TRIzole (15596026, Thermo Fisher; 400 мкл TRIzole на каждую лунку). В то же время, клеточную культуру (содержащую реагенты для трансфекции) в другой лунке истощают и добавляют свежую клеточную среду. Еще через 4 ч (всего 8 ч), клетки из оставшейся лунки дважды промывают DPBS, и затем собирают с использованием TRIzole. Оба образца раствора TRIzole очищают с использованием набора для экстракции общей РНК (R1013, Zymo). Общую РНК (1000 нг) из каждого образца используют в качестве матрицы для РВ-кПЦР с праймерами AZ112 и AZ86 (подходящими как для CS2 UTR, так и для CYBA UTR), из значений C_q которых рассчитывают исходное количество иРНК в соответствующем образце общей РНК. Очищенную соответствующую природную иРНК, полученную в результате транскрипции *in vitro*, используют для построения стандартных кривых для количественной оценки. Долю распада иРНК от 4 ч (конец процесса трансфекции) до 8 ч рассчитывают путем деления количества разницы иРНК между 4 ч и 8 ч по количеству иРНК через 4 ч.

Таблица 1. Праймеры

SEQ ID NO	Праймер	Последовательность

1	AZ01	GACAAATTAATACGACTCACTATAGGAAAACCTGATCATG TAGATCGAAC
2	A338	CCCCAGGCTTTACACTTTTATG
3	AZ67	TmGGCGGAAAACCCCGGGAATCTAACCCGGCTGAACGGA TT
4	AZ86	TCCACGCCGAACCTCCCGATC
5	AZ87	TCCCGGCTTCGCTGCATTTATTGC
6	A3112	AAAATCACGGCAGACAAACAAAAGAATGG
7	YZ73	ATGGGTCTCACACAAACTCGAGTACAACTTTAACTCACAC
8	YZ74	ATGGGTCTCGATTCCATTCTTTTGTGTGTCTGC
9	ED101	TAATACGACTCACTATAGG

Таблица 2. Олигонуклеотиды

SEQ ID NO	Олигонуклеотид	Олигонуклеотидная последовательность
10	EGFP_Y151_TAC	CTCGAGTACAACCTTTAACTCACACAATGTATACAT CACGGCAGACAAACAAAAAGAATGGAATC
11	EGFP_Y151_TAG	CTCGAGTACAACCTTTAACTCACACAATGTAGTAAT CACGGCAGACAAACAAAAAGAATGGAATC
12	EGFP_Y151_AXC	CTCGAGTACAACCTTTAACTCACACAATGTAAXCAT CAGGCAGACAAACAAAAAGAATGGAATC
13	EGFP_Y151_AYC	CTCGAGTACAACCTTTAACTCACACAATGTAAAYCAT ACGGCAGACAAACAAAAAGAATGGAATC
14	EGFP_Y151_GXC	CTCGAGTACAACCTTTAACTCACACAATGTAGXCAT CAGGCAGACAAACAAAAAGAATGGAATC
15	EGFP_Y151_GYC	CTCGAGTACAACCTTTAACTCACACAATGTAGYCAT CAGGCAGACAAACAAAAAGAATGGAATC
16	EGFP_Y151_GXT	CTCGAGTACAACCTTTAACTCACACAATGTAGXTAT ACGGCAGACAAACAAAAAGAATGGAATC
17	EGFP_Y151_GYT	CTCGAGTACAACCTTTAACTCACACAATGTAGYTAT CACGGCAGACAAACAAAAAGAATGGAATC
18	EGFP_Y151_AGX	CTCGAGTACAACCTTTAACTCACACAATGTAAGXAT CACGGCAGACAAACAAAAAGAATGGAATC
19	EGFP_Y151_XCC	CTCGAGTACAACCTTTAACTCACACAATGTAXCCAT CACGGCAGACAAACAAAAAGAATGGAATC

20	EGFP_Y151_YCC	CTCGAGTACAACCTTTAACTCACACAATGTAYCCAT CACGGCAGACAAACAAAAGAATGGAATC
21	Mm_tRNA_GTA	CCTGATCATGTAGATCGAACGGACTGTAAATCCGT TCAGCCGGGTAGATTC
22	Mm_tRNA_CTA	CCTGATCATGTAGATCGAACGGACTCTAAATCCGT TCAGCCGGGTAGATTC
23	Mm_tRNA_GYT	CCTGATCATGTAGATCGAACGGACTGYTAATCCGT TCAGCCGGGTAGATTC
24	Mm_tRNA_GXT	CCTGATCATGTAGATCGAACGGACTGXTAATCCGT TCAGCCGGGTAGATTC
25	Mm_tRNA_GYC	CCTGATCATGTAGATCGAACGGACTGYCAATCCGT TCAGCCGGGTAGATTC
26	Mm_tRNA_GXC	CCTGATCATGTAGATCGAACGGACTGXCAATCCGT TCAGCCGGGTAGATTC
27	Mm_tRNA_AYC	CCTGATCATGTAGATCGAACGGACTAYCAATCCGT TCAGCCGGGTAGATTC
28	Mm_tRNA_AXC	CCTGATCATGTAGATCGAACGGACTAXCAATCCGT TCAGCCGGGTAGATTC
29	Mm_tRNA_YCT	CCTGATCATGTAGATCGAACGGACTYCTAATCCGT TCAGCCGGGTAGATTC
30	Mm_tRNA_XCT	CCTGATCATGTAGATCGAACGGACTXCTAATCCGT TCAGCCGGGTAGATTC
31	Mm_tRNA_GGY	CCTGATCATGTAGATCGAACGGACTGGYAATCCG TTCAGCCGGGTAGATTC
32	Mm_tRNA_GGX	CCTGATCATGTAGATCGAACGGACTGGXAATCCG TTCAGCCGGGTAGATTC

Другие последовательности

IRES (SEQ ID NO: 33):

CATCTAGGGCGGCCAATTCCGCCCCTCTCCCTCCCCCCCCCTAACGTTACTG
GCCGAAGCCGCTTGAATAAGGCCGGTGTGCGTTTGTCTATATGTGATTTTCCACCA
TATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGA
GCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTGC
TGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACC
CTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCC
ACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCAGTGCCACGTTGTGAGTT
GGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTG

AAGGATGCCCAGAAGGTACCCCATTTGTATGGGATCTGATCTGGGGCCTCGGTGCAC
 ATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCGAACCACGG
 GGACGTGGTTTTCTTTGAAAAACACGATGATAAGCTTGCCAC

mCherry (SEQ ID NO: 34)

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATG
 CGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGG
 CGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACC
 AAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGC
 TCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTC
 CCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGAC
 CGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGC
 GCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGG
 GAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAA
 GCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCT
 ACAAGGCCAAGAAGCCCGTGACGCTGCCCCGGCGCCTACAACGTCAACATCAAGTTG
 GACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGA
 GGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAA

ChPylRS_C211 (SEQ ID NO: 35)

ATGGATAAAAAACCGCTGGACGTTCTGATCTCCGCTACGGGTCTGTGGATGA
 GCCGCACGGGTACGCTGCATAAAATCAAGCACTATGAGATTTCTCGTTCTAAAATCT
 ACATCGAAATGGCGTGTGGTGACCATCTGGTTGTGAACAACTCTCGTTCTTGTCTC
 CGGCACGTGCATTCCGTTATCATAAATACCGTAAACCTGCAAACGTTGTCGTGTTT
 CTGACGAAGATATCAACAACCTTCCTGACCCGTTCTACCGAAGGCCAAAACCTCTGTTA
 AAGTTAAAGTTGTTTCTGAACCGAAAGTGAAAAAGCGATGCCGAAATCTGTTTCTC
 GTGCGCCGAAACCGCTGGAAAATCCGGTTTCTGCGAAAGCGTCTACCGACACCTCTC
 GTTCTGTTCCGTCTCCGGCGAAATCTACCCCGAACTCTCCGGTTCCGACCTCTGCAA
 GTGCCCCCGCACTTACGAAGAGCCAGACTGACAGGCTTGAAGTCCTGTAAACCCA
 AAAGATGAGATTTCCCTGAATTCCGGCAAGCCTTTCAGGGAGCTTGAGTCCGAATTG
 CTCTCTCGCAGAAAAAAGACCTGCAGCAGATCTACGCGGAAGAAAGGGAGAATTA
 TCTGGGGAAACTCGAGCGTGAAATTACCAGGTTCTTTGTGGACAGGGGTTTTCTGGA
 AATAAAATCCCCGATCCTGATCCCTCTTGAGTATATCGAAAGGATGGGCATTGATAA
 TGATAACCGAACTTTCAAAACAGATCTTCAGGGTTGACAAGAACTTCTGCCTGAGACC
 CATGCTTGCTCCAAACCTTTACAACCTACCTGCGCAAGCTTGACAGGGGCCCTGCCTGA
 TCCAATAAAAAATTTTTGAAATAGGCCCATGCTACAGAAAAGAGTCCGACGGCAAAG
 AACACCTCGAAGAGTTTACCATGCTGAACTTCTGCCAGATGGGATCGGGATGCACA
 CGGGAAAATCTTGAAAGCATAATTACGGACTTCCTGAACCACCTGGGAATTGATTTC
 AAGATCGTAGGCGATTCTTGCATGGTCTATGGGGATAACCCTTGATGTAATGCACGGA
 GACCTGGAACCTTCTCTGCAGTAGTCGGACCCATAACCGCTTGACCGGGAATGGGGT
 ATTGATAAACCTTGATAGGGGCAGGTTTCGGACTCGAACGCCTTCTAAAGGTAA
 ACACGACTTTAAAAATATCAAGAGAGCTGCACGCTCGGAATCGTATTACAACGGCA

TCTCAACCAATCTGTAA

CS2 5'UTR (SEQ ID NO: 36):

GAATACAAGCTACTTGTCTTTTTGCAGGATCCGCCACC

CS2 3'UTR (SEQ ID NO: 37):

AAGCTTAATTAGCTGAGCTTGGACTCCTAAGCATGCAAGCTTGGCGTAATCA
TGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAC
GAGCCGGAAGCATAAAGTGTAAGCCTGGGG

CYBA 5'UTR (SEQ ID NO: 38):

CGCGCCTAGCAGTGTCCCAGCCGGGTTCGTGTCGCC

CYBA 3'UTR (SEQ ID NO: 39):

CCTCGCCCCGGACCTGCCCTCCCGCCAGGTGCACCCACCTGCAATAAATGCA
GCGAAGCCGGGA

EGFP (вектор Golden Gate) (с 2xStrepTag) (SEQ ID NO: 40):

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGTGGTGCCCATCCTGGTTCG
AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCC
GTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGC
TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCAAGGCTAC
GTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA
GGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACT
TCAAGGAGGACGGCAACATCCTGGGGCACAAGAGACCCTCGAGAATATTCTCGAGG
GTCTCGGAATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGGCGACGGCCCCGTGCTGCT
GCCCCACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGA
AGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA
TGGACGAGCTGTACAAGAAGCTTTGGAGCCACCCGCAGTTCGAGAAAGGTGGAGGT
TCCGGAGGTGGATCGGGAGGTTCGGCGTGGAGCCACCCGCAGTTCGAAAAATAA

FLP (SEQ ID NO: 41)

ATGCCACAATTTGATATATTATGTAAAACACCACCTAAGGTGCTTGTTCGTCA
GTTTGTGGAAAGGTTTGAAAGACCTTCAGGTGAGAAAATAGCATTATGTGCTGCTGA
ACTAACCTATTTATGTTGGATGATTACACATAACGGAACAGCAATCAAGAGAGCCA
CATTCATGAGCTATAATACTATCATAAGCAATTCGCTGAGTTTGGATATTGTCAACA
AGTCACTGCAGTTTAAATACAAGACGCAAAAAGCAACAATTCTGGAAGCCTCATT
AAGAAATTGATTCCTGCTTGGGAATTTACAATTATTCCTTACTATGGACAAAAACAT
CAATCTGATATCACTGATATTGTAAGTAGTTTGAATTACAGTTCGAATCATCGGAA
GAAGCAGATAAGGGAAATAGCCACAGTAAAAAAATGCTTAAAGCACTTCTAAGTGA
GGGTGAAAGCATCTGGGAGATCACTGAGAAAATACTAAATTCTGTTTGAGTATACTTC
GAGATTTACAAAAACAAAACTTTATACCAATTCCTCTTCCTAGCTACTTTCATCAA
TTGTGGAAGATTCAGCGATATTAAGAACGTTGATCCGAAATCATTTAAATTAGTCCA
AAATAAGTATCTGGGAGTAATAATCCAGTGTTTAGTGACAGAGACAAAGACAAGCG

TTAGTAGGCACATATACTTCTTTAGCGCAAGGGGTAGGATCGATCCACTTGTATATT
 TGGATGAATTTTTGAGGAATTCTGAACCAGTCCTAAAACGAGTAAATAGGACCGGC
 AATTCTTCAAGCAACAAGCAGGAATACCAATTATTAAGATAACTTAGTCAGATC
 GTACAACAAAGCTTTGAAGAAAAATGCGCCTTATTCAATCTTTGCTATAAAAAATGG
 CCCAAAATCTCACATTGGAAGACATTTGATGACCTCATTTCTTTCAATGAAGGGCCT
 AACGGAGTTGACTAATGTTGTGGGAAATTGGAGCGATAAGCGTGCTTCTGCCGTGG
 CCAGGACAACGTATACTCATCAGATAACAGCAATACCTGATCACTACTTCGCACTAG
 TTTCTCGGTACTATGCATATGATCCAATATCAAAGGAAATGATAGCATTGAAGGATG
 AGACTAATCCAATTGAGGAGTGGCAGCATATAGAACAGCTAAAGGGTAGTGCTGAA
 GGAAGCATACGATACCCCGCATGGAATGGGATAATATCACAGGAGGTACTAGACTA
 CCTTTCATCCTACATAAATAGACGCATATAA

FRT (SEQ ID NO: 42)

GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC

Хотя в настоящем документе были показаны и описаны предпочтительные варианты осуществления настоящего изобретения, специалистам в данной области техники будет очевидно, что такие варианты осуществления, представлены только в качестве примера. Специалистам в данной области техники будут очевидны многочисленные вариации, изменения и замены без отступления от описания. Следует понимать, что различные альтернативы вариантам осуществления описания, описанным в настоящем документе, могут быть использованы при практическом применении описания. Подразумевается, что следующая формула изобретения определяет объем изобретения, и что способы и структуры в пределах объема этих пунктов формулы изобретения и их эквиваленты покрываются этим.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Эукариотическая клетка, содержащая:

(а) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; и

(b) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения,

где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (УВР) в эукариотической клетке, и где иРНК способна транслироваться в клетке с образованием полипептида, по меньшей мере, с одной аминокислотой не природного происхождения.

2. Эукариотическая клетка по п. 1, отличающаяся тем, что тРНК заряжена аминокислотой не природного происхождения.

3. Эукариотическая клетка по любому из предыдущих пп., дополнительно содержащая полипептид, транслированный с иРНК, где указанный полипептид содержит аминокислоту не природного происхождения, где полипептид необязательно содержит эукариотический профиль гликозилирования.

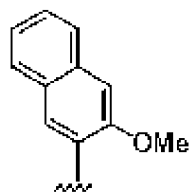
4. Эукариотическая клетка по любому из предыдущих пп., дополнительно содержащая тРНК синтетазу, где тРНК синтетаза предпочтительно аминоацилирует тРНК с аминокислотами не природного происхождения.

5. Эукариотическая клетка по любому из предыдущих пп., отличающаяся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

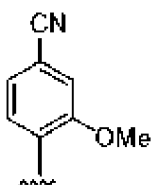
6. Эукариотическая клетка по любому из предыдущих пп., отличающаяся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

7. Эукариотическая клетка по любому из предыдущих пп., отличающаяся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

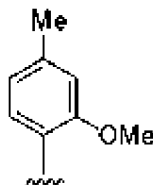
8. Эукариотическая клетка по любому из предшествующих пп., отличающаяся тем, что каждое основание не природного происхождения и второе основание не природного происхождения независимо выбраны из группы, состоящей из



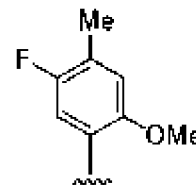
(NaM),

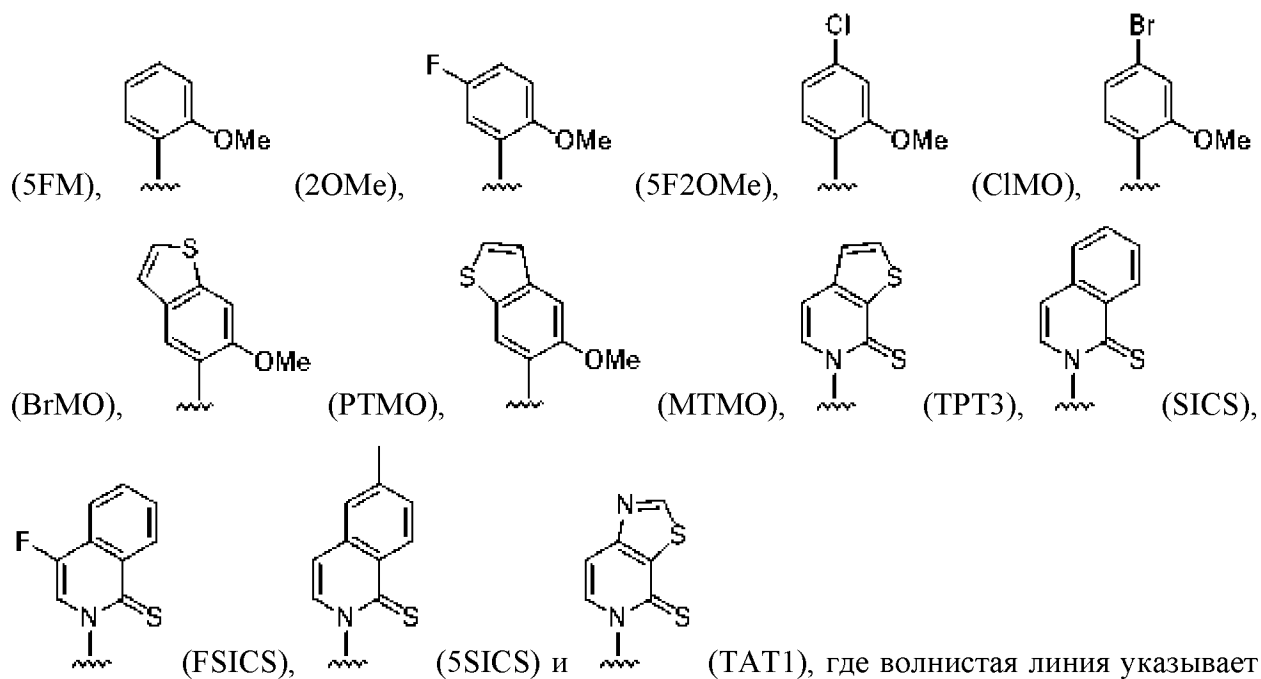


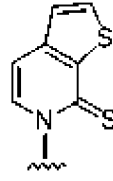
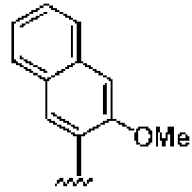
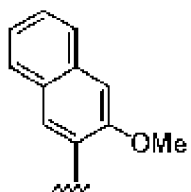
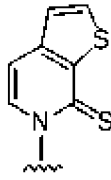
(CNMO),



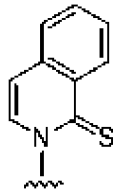
(MMO2),

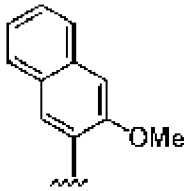


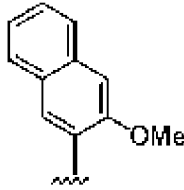


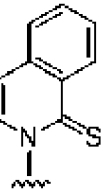
9. Эукариотическая клетка по любому из предыдущих пп., когда первое основание не природного происхождения представляет собой  (TPT3), второе основание не природного происхождения представляет собой  (NaM), и когда первое основание не природного происхождения представляет собой  (NaM), второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

10. Эукариотическая клетка по любому из предыдущих пп., когда первое основание

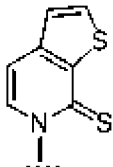
не природного происхождения представляет собой  (5SICS), второе основание не

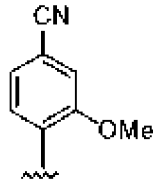
природного происхождения представляет собой  (NaM), и когда первое

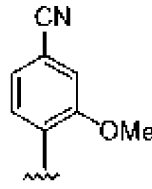
основание не природного происхождения представляет собой  (NaM), второе

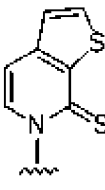
основание не природного происхождения представляет собой  (5SICS), где волнистая линия указывает на связь с рибозильной группой.

11. Эукариотическая клетка по любому из предыдущих пп., когда первое основание

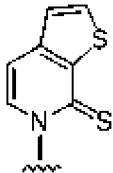
не природного происхождения представляет собой  (TPT3), второе основание не

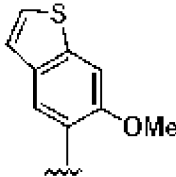
природного происхождения представляет собой  (CNMO), и когда первое

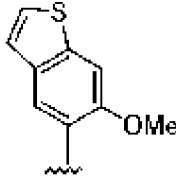
основание не природного происхождения представляет собой  (CNMO), второе

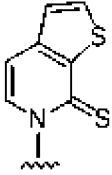
основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

12. Эукариотическая клетка по любому из предыдущих пп., когда первое основание

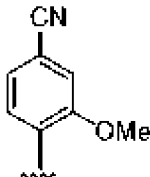
не природного происхождения представляет собой  (TPT3), второе основание не

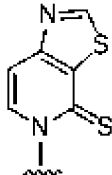
природного происхождения представляет собой  (PTMO), и когда первое

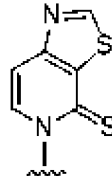
основание не природного происхождения представляет собой  (PTMO),

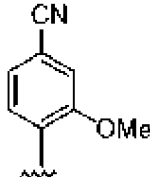
второе основание не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

13. Эукариотическая клетка по любому из предыдущих пп., когда первое основание

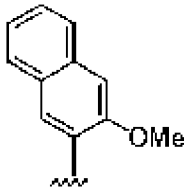
не природного происхождения представляет собой  (CNMO), второе основание

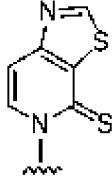
не природного происхождения представляет собой  (TAT1), и когда первое

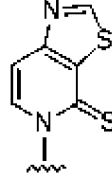
основание не природного происхождения представляет собой  (TAT1), второе

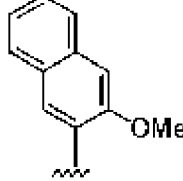
основание не природного происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

14. Эукариотическая клетка по любому из предыдущих пп., когда первое основание

не природного происхождения представляет собой  (NaM), второе основание

не природного происхождения представляет собой  (TAT1), и когда первое

основание не природного происхождения представляет собой  (TAT1), второе

основание не природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

15. Эукариотическая клетка по любому из пп. 3-14, отличающаяся тем, что, по меньшей мере, одна аминокислота не природного происхождения:

- представляет собой аналог лизина;
- содержит ароматическую боковую цепь;
- содержит азидогруппу;
- содержит алкиновую группу; или
- содержат альдегидную или кетонную группу.

16. Эукариотическая клетка по п. 15, отличающаяся тем, что по меньшей мере, одна аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненилизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононаноевой кислоты, п-пропаргилоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтирозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-

(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

17. Эукариотическая клетка по п. 16, отличающаяся тем, что, по меньшей мере, одна аминокислота не природного происхождения представляет собой N6-((азидозтокси)-карбонил)-L-лизин (AzK).

18. Эукариотическая клетка по любому из предыдущих пп., отличающаяся тем, что эукариотическая клетка представляет собой клетку человека.

19. Эукариотическая клетка по непосредственно предшествующему п., отличающаяся тем, что клетка человека представляет собой клетку НЕК293Т.

20. Эукариотическая клетка по любому из пп. 1-18, отличающаяся тем, что клетка представляет собой клетку млекопитающего, необязательно, где клетка представляет собой клетку хомяка.

21. Эукариотическая клетка по непосредственно предшествующему п., отличающаяся тем, что клетка млекопитающего представляет собой клетку яичника китайского хомяка (CHO).

22. Эукариотическая клетка по любому из пп. 18-21, дополнительно содержащая полипептид, транслированный из иРНК, где полипептид содержит аминокислоту не природного происхождения и профиль гликозилирования млекопитающих.

23. Эукариотическая клетка по любому из предыдущих пп., отличающаяся тем, что указанная клетка выделена.

24. Полусинтетический организм, содержащий эукариотическую клетку по любому из предыдущих пп.

25. Культура эукариотических клеток, содержащая множество эукариотических клеток по любому из пп. 1-24.

26. Способ доставки клетки в организм, включающий контакт организма с клеткой по любому из пп. 1-23.

27. Способ по п. 26, отличающийся тем, что организм представляет собой млекопитающее, необязательно, где млекопитающее представляет собой человека.

28. Способ получения полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения, в эукариотической клетке, включающий:

(а) введение в клетку:

(i) информационной РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения; и

(ii) транспортной РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения в эукариотической клетке, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке; и

(b) трансляцию полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения из иРНК с использованием тРНК.

29. Способ по предыдущему пункту, отличающийся тем, что тРНК заряжена аминокислотой не природного происхождения.

30. Способ получения полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения, в эукариотической клетке, включающий:

(а) получение эукариотической клетки, содержащей:

(i) информационную РНК (иРНК) с кодоном, содержащим первое основание не природного происхождения;

(ii) транспортную РНК (тРНК) с антикодоном, содержащим второе основание не природного происхождения, где первое и второе основания не природного происхождения способны образовывать пару оснований не природного происхождения (UBP) в эукариотической клетке; и

(b) трансляцию полипептида, содержащего, по меньшей мере, одну аминокислоту не природного происхождения, из иРНК с использованием тРНК с помощью рибосомы, которая является эндогенной для эукариотической клетки.

31. Способ получения полипептида в эукариотической клетке, где полипептид содержит, по меньшей мере, одну аминокислоту не природного происхождения, включающий:

(а) предоставление эукариотической клетки, где эукариотическая клетка содержит:

(i) иРНК, содержащую кодон, где кодон содержит первое основание не природного происхождения;

(ii) тРНК, содержащую антикодон, где антикодон содержит второе основание не природного происхождения, и где первое и второе основания не природного происхождения способны образовывать комплементарную пару оснований; и

(iii) тРНК синтетазу, где тРНК синтаза предпочтительно аминоацилирует тРНК с, по меньшей мере, одной аминокислотой не природного происхождения по сравнению с природной аминокислотой; и

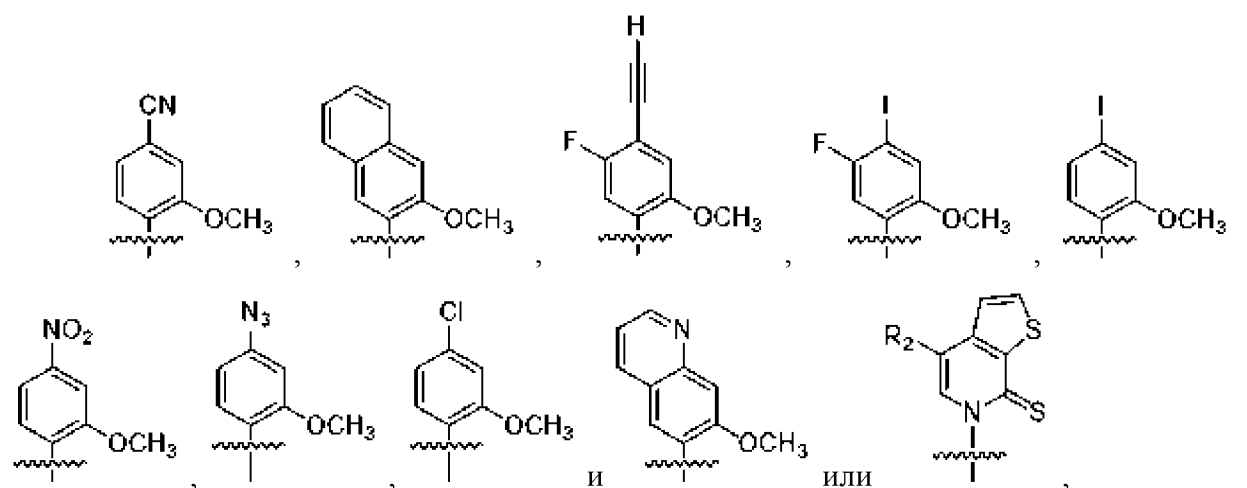
(b) предоставление еще одной аминокислоты не природного происхождения в эукариотическую клетку, где эукариотическая клетка продуцирует полипептид, содержащий, по меньшей мере, одну аминокислоту не природного происхождения.

32. Способ по любому из пп. 26-31, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК.

33. Способ по любому из пп. 26-31, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК.

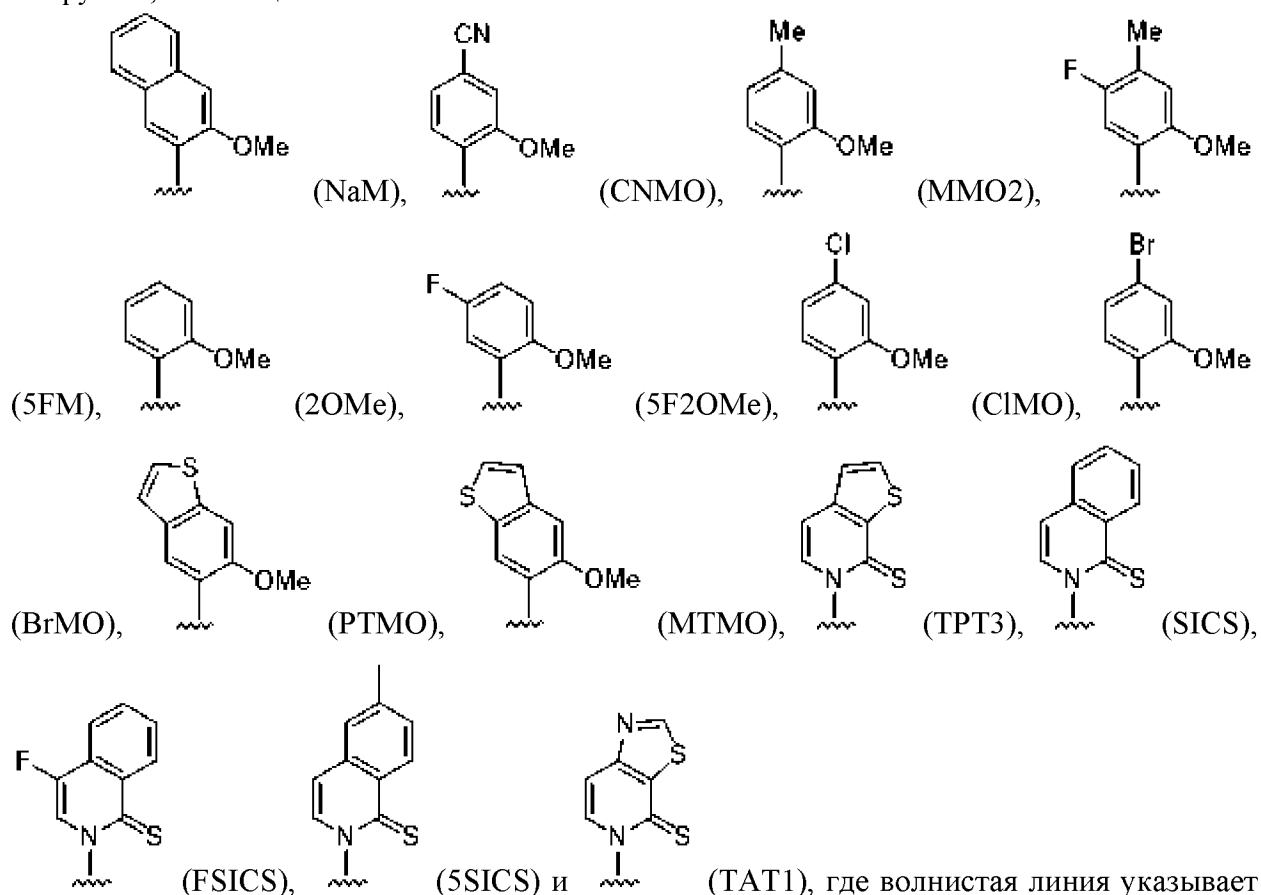
34. Способ по любому из пп. 26-31, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК.

35. Способ по любому из пп. 26-34, отличающийся тем, что одно или несколько оснований не природного происхождения, содержащих кодон иРНК, имеют формулу



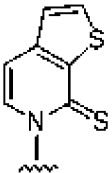
где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азидо, и волнистая линия указывает на связь с рибозильной группой.

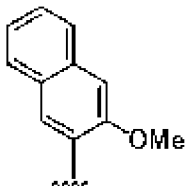
36. Способ по любому из пп. 26-35, отличающийся тем, что первое основание не природного происхождения или второе основание не природного происхождения выбрано из группы, состоящей из

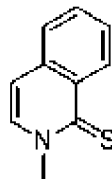


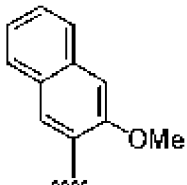
на связь с рибозильной группой.

37. Способ по п. 36, отличающийся тем, что первое основание не природного

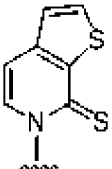
происхождения представляет собой  (TPT3) и второе основание не природного

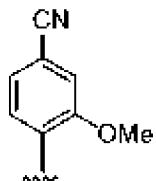
происхождения представляет собой  (NaM), или первое основание не

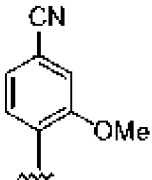
природного происхождения представляет собой  (5SICS), и второе основание не

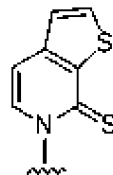
природного происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

38. Способ по п. 36, отличающийся тем, что первое основание не природного

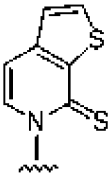
происхождения представляет собой  (TPT3) и второе основание не природного

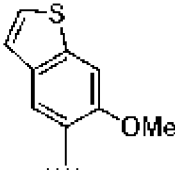
происхождения представляет собой  (CNMO), или первое основание не

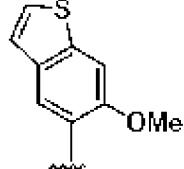
природного происхождения представляет собой  (CNMO), и второе основание

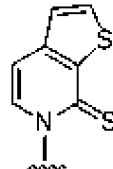
не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

39. Способ по п. 36, отличающийся тем, что первое основание не природного

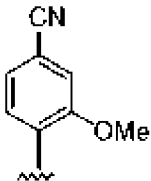
происхождения представляет собой  (TPT3) и второе основание не природного

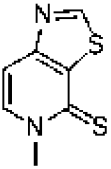
происхождения представляет собой  (PTMO), или первое основание не

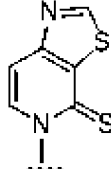
природного происхождения представляет собой  (PTMO), и второе основание

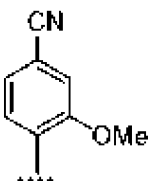
не природного происхождения представляет собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

40. Способ по п. 36, отличающийся тем, что первое основание не природного

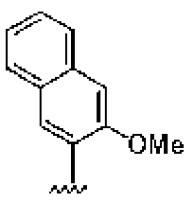
происхождения представляет собой  (CNMO) и второе основание не природного

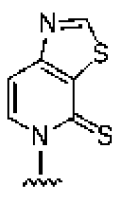
происхождения представляет собой  (TAT1), или первое основание не природного

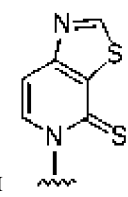
происхождения представляет собой  (TAT1) и второе основание не природного

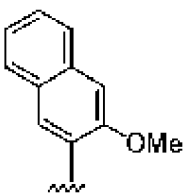
происхождения представляет собой  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

41. Способ по п. 36, отличающийся тем, что первое основание не природного

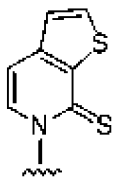
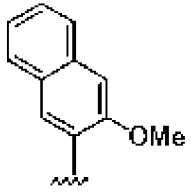
происхождения представляет собой  (NaM) и второе основание не природного

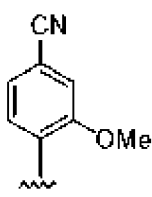
происхождения представляет собой  (TAT1), или первое основание не природного

происхождения представляет собой  (TAT1) и второе основание не природного

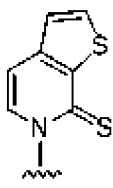
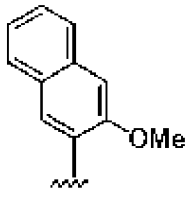
происхождения представляет собой  (NaM), где волнистая линия указывает на связь с рибозильной группой.

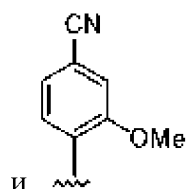
42. Способ по любому из пп. 26-36, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в первом положении (X-N-N) в кодоне иРНК, где первое

основание не природного происхождения выбрано из  (TPT3),  (NaM)

и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

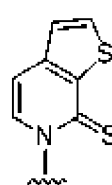
43. Способ по любому из пп. 26-36, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в среднем положении (N-X-N) в кодоне иРНК, где первое

основание не природного происхождения выбрано из  (TPT3),  (NaM)

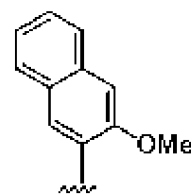


и (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

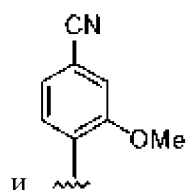
44. Способ по любому из пп. 26-36, отличающийся тем, что кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) расположено в последнем положении (N-N-X) в кодоне иРНК, где



(TPT3),

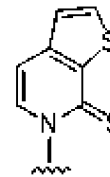


(NaM)

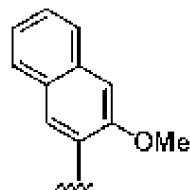


и (CNMO), где волнистая линия указывает на связь с рибозильной группой.

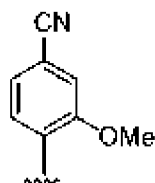
45. Способ по любому из пп. 26-36, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где второе основание не природного происхождения (X) расположено в первом положении (X-N-N) в антикодоне



(TPT3),



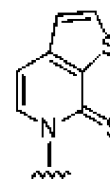
(NaM) и



(CNMO), и где волнистая линия указывает на связь с

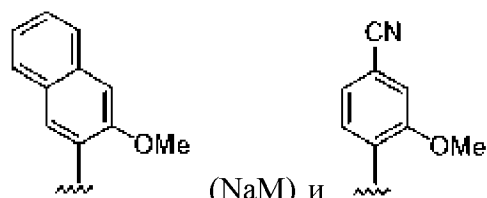
рибозильной группой.

46. Способ по любому из пп. 26-36, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где второе основание не природного происхождения (X) расположено в среднем положении (N-X-N) в антикодоне



(TPT3),

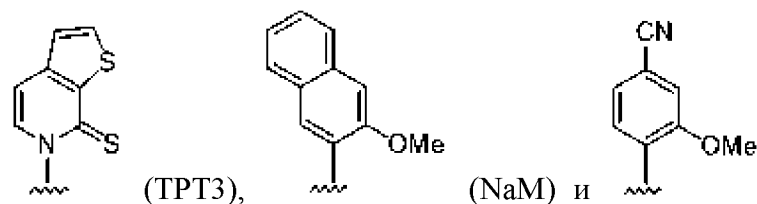
тРНК, где второе основание не природного происхождения выбрано из



(NaM) и (CNMO), и где волнистая линия указывает на связь с

рибозильной группой.

47. Способ по любому из пп. 26-36, отличающийся тем, что антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где второе основание не природного происхождения (X) расположено в последнем положении (N-N-X) в антикодоне тРНК, где второе основание не природного происхождения выбрано из



(TPT3), (NaM) и (CNMO), и где волнистая линия

указывает на связь с рибозильной группой.

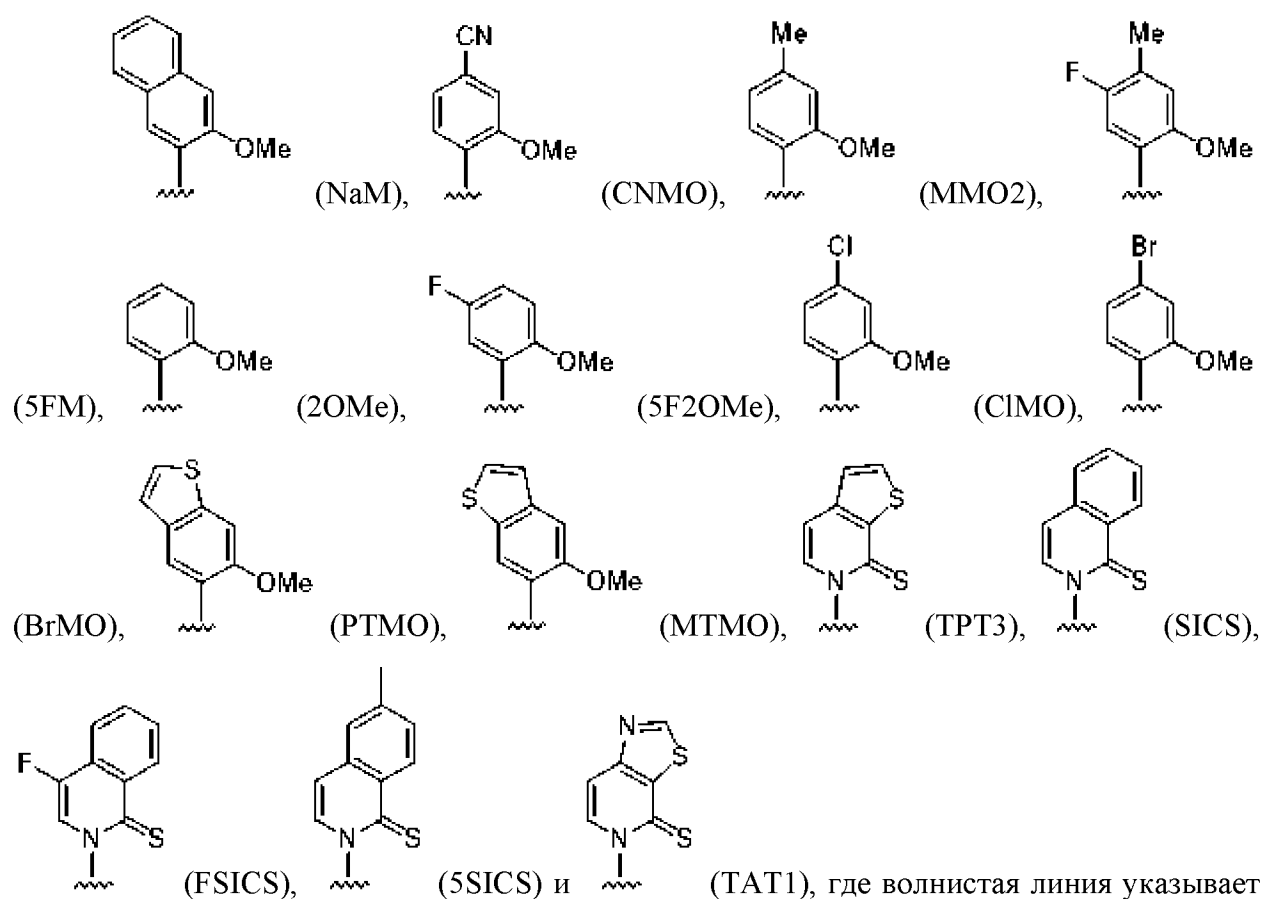
48. Способ по любому из пп. 26-36, отличающийся тем, что каждый кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где первое основание не природного происхождения (X) кодона в иРНК расположено в первом положении (X-N-N) кодона, и второе основание не природного происхождения (Y) антикодона тРНК расположено в последнем положении (N-N-Y) антикодона.

49. Способ по любому из пп. 26-36, отличающийся тем, что каждый кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в среднем положении (N-X-N) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в среднем положении (N-Y-N) антикодона.

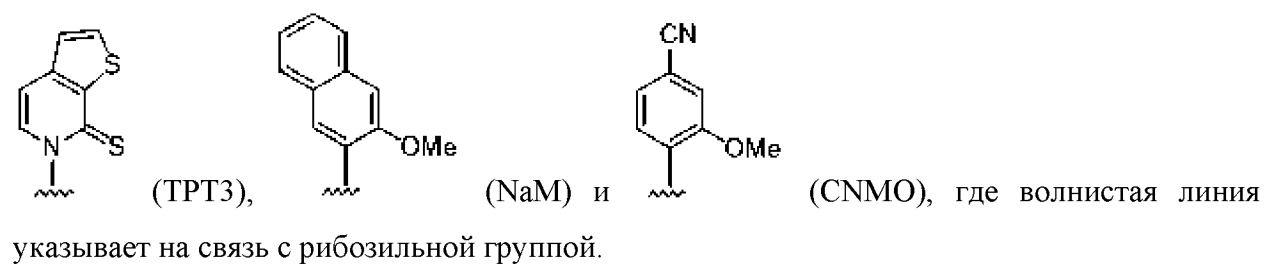
50. Способ по любому из пп. 26-36, отличающийся тем, что каждый кодон и антикодон содержат три непрерывных нуклеотидных основания (N-N-N), где кодон в иРНК содержит первое основание не природного происхождения (X), расположенное в последнем положении (N-N-X) кодона, и антикодон в тРНК содержит второе основание не природного происхождения (Y), расположенное в первом положении (Y-N-N) антикодона.

51. Способ по любому из пп. 48-50, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, являются одинаковыми или разными.

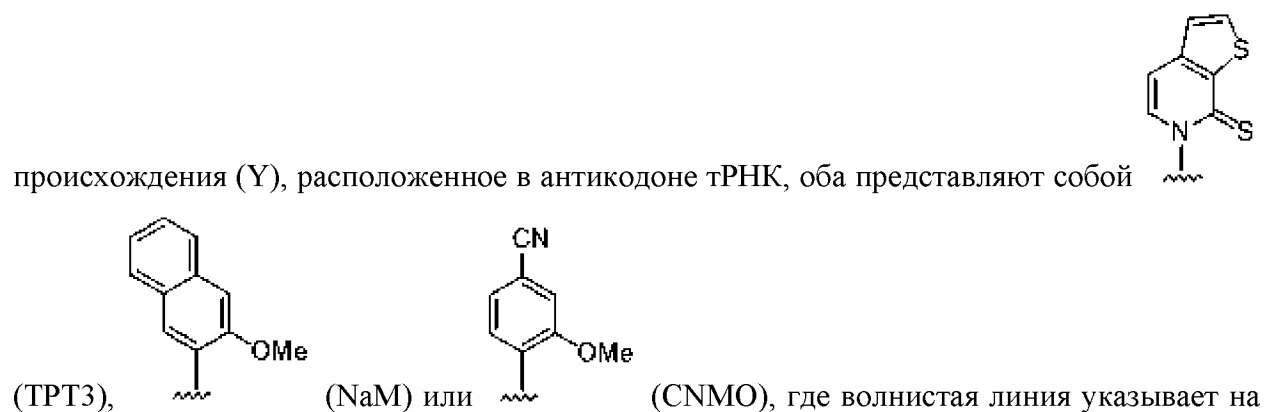
52. Способ по любому из пп. 48-51, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбраны из группы, состоящей из



53. Способ по п. 52, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, выбрано из группы, состоящей из

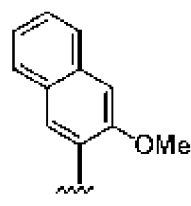


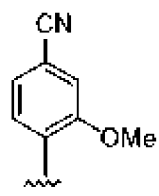
54. Способ по п. 53, отличающийся тем, что первое основание не природного происхождения (X), расположенное в кодоне иРНК, и второе основание не природного происхождения (Y), расположенное в антикодоне тРНК, оба представляют собой

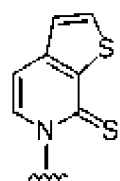


связь с рибозильной группой.

55. Способ по п. 53, отличающийся тем, что первое основание не природного

происхождения (X), расположенное в кодоне иРНК, выбрано из  (NaM) и

 (CNMO), и второе основание не природного происхождения (Y),

расположенное в антикодоне тРНК, представляет собой  (TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

56. Способ по любому из пп. 26-29, 31, 33, 35-41, 43, 46 и 49, отличающийся тем, что кодон в иРНК выбран из AXC, GXC или GXU, где X представляет собой первое основание не природного происхождения.

57. Способ по непосредственно предшествующему пункту, отличающийся тем, что антикодон в тРНК выбран из GYU, GYC и AYC, и Y представляет собой второе основание не природного происхождения.

58. Способ по п. 57, отличающийся тем, что кодон в иРНК представляет собой AXC, и антикодон в тРНК представляет собой GYU.

59. Способ по п. 57, отличающийся тем, что кодон в иРНК представляет собой GXC, и антикодон в тРНК представляет собой GYC.

60. Способ по п. 57, отличающийся тем, что кодон в иРНК представляет собой GXU, и антикодон представляет собой AYC.

61. Способ по любому из пп. 26-60, отличающийся тем, что первое основание не природного происхождения или второе основание не природного происхождения содержат модифицированную сахарную группу, выбранную из группы, состоящей из:

модификации в положении 2', включающей:

ОН, замещенный низший алкил, алкарил, аралкил, О-алкарил или О-аралкил, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂F, или их комбинацию;

О-алкил, S-алкил, N-алкил, или их комбинацию;

О-алкенил, S-алкенил, N-алкенил, или их комбинацию;

О-алкинил, S-алкинил, N-алкинил, или их комбинацию;

О-алкил-О-алкил, 2'-F, 2'-OCH₃, 2'-O(CH₂)₂OCH₃, или их комбинацию, где алкил,

алкенил и алкинил могут быть замещены или не замещены C_1 - C_{10} алкилом, C_2 - C_{10} алкенилом, C_2 - C_{10} алкинилом, $O[(CH_2)_nO]_mCH_3$, $-O(CH_2)_nOCH_3$, $-O(CH_2)_nNH_2$, $-O(CH_2)_nCH_3$, $-O(CH_2)_nNH_2$ и $-O(CH_2)_nON[(CH_2)_nCH_3]_2$, где n и m составляют от 1 до примерно 10;

модификации в положении 5', включающей:

5'-винил, 5'-метил (R или S) или их комбинацию;

модификации в положении 4', включающей:

4'-S, гетероциклоалкил, гетероциклоалкарил, аминокламино, полиалкламино, замещенный силл, группа, расщепляющая РНК, репортерная группа, интеркалятор, группа для улучшения фармакокинетических свойств олигонуклеотида или группа для улучшения фармакодинамических свойств олигонуклеотида, или их комбинацию;

или их комбинации.

62. Способ по любому из пп. 26-61, отличающийся тем, что, по меньшей мере, одна аминокислота не природного происхождения:

представляет собой аналог лизина;

содержит ароматическую боковую цепь;

содержит азидогруппу;

содержит алкиновую группу; или

содержат альдегидную или кетонную группу.

63. Способ по любому из пп. 26-61, отличающийся тем, что по меньшей мере, одна аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргилфенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Dopa, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацетил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, O-аллилтирозина, O-метил-L-тирозина, O-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-O-ацетил-GlcNAc-серина, L-фосфосерина, фосфосерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-((фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

64. Способ по п. 63, отличающийся тем, что аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK).

65. Способ по любому из пп. 26-64, отличающийся тем, что клетка представляет собой клетку человека.

66. Способ по п. 65, отличающийся тем, что клетка человека представляет собой

клетку HEK293T.

67. Способ по любому из пп. 26-64, отличающийся тем, что клетка представляет собой клетку хомяка.

68. Способ по п. 67, отличающийся тем, что клетка хомяка представляет собой клетку яичника китайского хомяка (CHO).

69. Способ по любому из пп. 26-68, отличающийся тем, что тРНК получают из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

70. Способ по любому из пп. 26-69, отличающийся тем, что клетка содержит тРНК-синтетазу, полученную из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

71. Система для экспрессии полипептида не природного происхождения, содержащая:

(а) по меньшей мере, одну аминокислоту не природного происхождения;

(b) иРНК, кодирующую не встречающийся в природе полипептид, где указанная иРНК содержит, по меньшей мере, один кодон, содержащий одно или несколько первых оснований не природного происхождения;

(с) тРНК, содержащую, по меньшей мере, один антикодон, содержащий одно или несколько вторых оснований не природного происхождения, где одно или несколько первых оснований не природного происхождения и одно или несколько вторых оснований не природного происхождения способны образовывать одну или несколько комплементарных пар оснований; и

(d) эукариотическую рибосому, способную транслировать иРНК в полипептид, содержащий аминокислоту не природного происхождения, с использованием тРНК и тРНК-синтетазы,

где тРНК заряжена аминокислотой не природного происхождения, или система дополнительно содержит тРНК синтетазу или одну или несколько конструкций нуклеиновой кислоты, содержащих последовательность нуклеиновой кислоты, кодирующую тРНК синтетазу, где тРНК синтетаза предпочтительно аминоацилирует тРНК, по меньшей мере, с одной аминокислотой не природного происхождения.

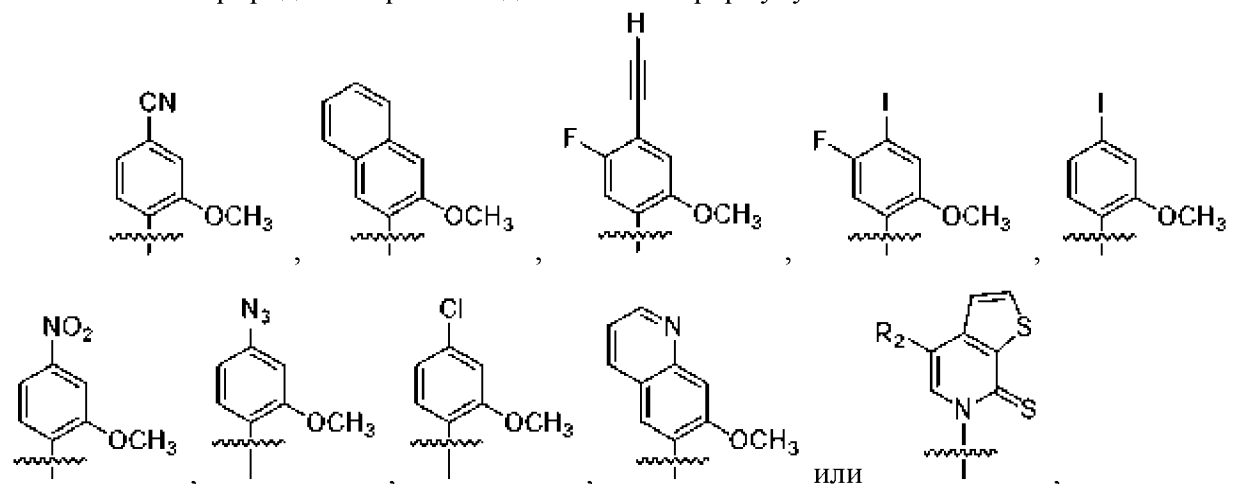
72. Система по п. 71, отличающаяся тем, что по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в, по меньшей мере, одном кодоне иРНК.

73. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК.

74. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько

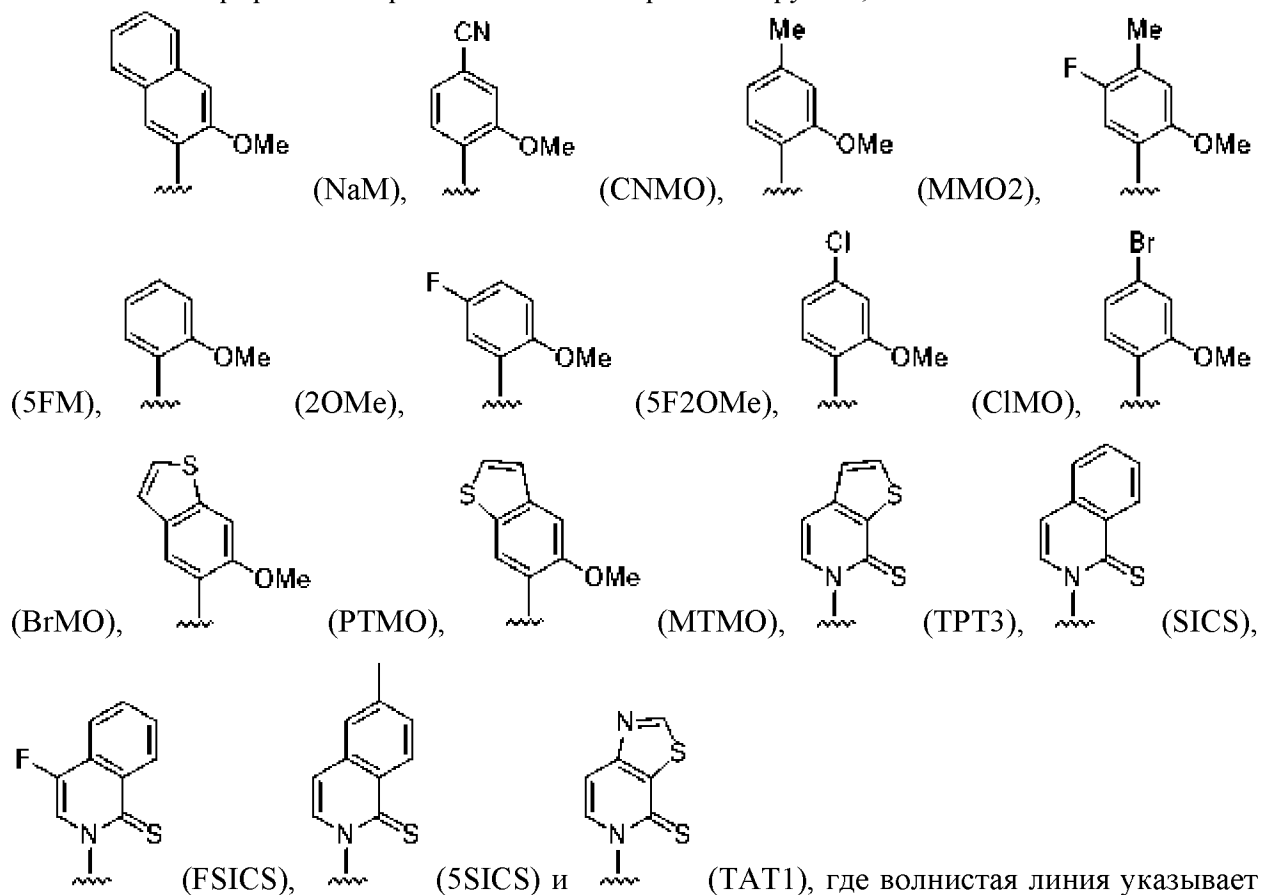
первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в, по меньшей мере, одном кодоне иРНК.

75. Система по любому из пп. 71-74, отличающаяся тем, что одно или несколько оснований не природного происхождения имеют формулу



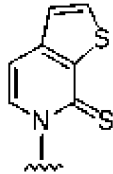
где R_2 выбран из группы, состоящей из водорода, алкила, алкенила, алкинила, метокси, метантиола, метанселено, галогена, циано и азида, и волнистая линия указывает на связь с рибозильной группой.

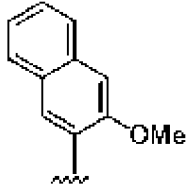
76. Система по любому из пп. 71-74, отличающаяся тем, что одно или несколько первых оснований не природного происхождения или одно или несколько вторых оснований не природного происхождения выбраны из группы, состоящей из

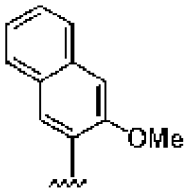


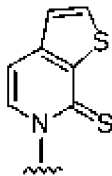
на связь с рибозильной группой.

77. Система по п. 76, когда одно или несколько первых оснований не природного

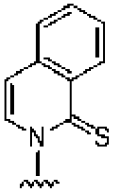
происхождения представляют собой  (TPT3), одно или несколько вторых

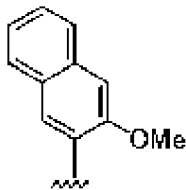
оснований не природного происхождения представляют собой  (NaM), и
когда одно или несколько первых оснований не природного происхождения представляют

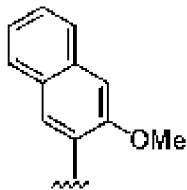
собой  (NaM), второе основание не природного происхождения представляет

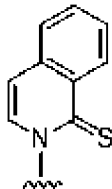
собой  (TPT3), где волнистая линия указывает на связь с рибозильной группой.

78. Система по п. 76, когда одно или несколько первых оснований не природного

происхождения представляют собой  (5SICS), одно или несколько вторых

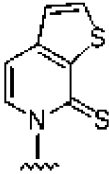
оснований не природного происхождения представляют собой  (NaM), и
когда одно или несколько первых оснований не природного происхождения представляют

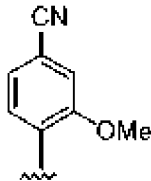
собой  (NaM), одно или несколько вторых оснований не природного

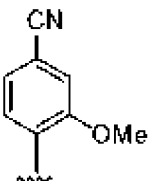
происхождения представляют собой  (5SICS), где волнистая линия указывает на

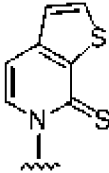
связь с рибозильной группой.

79. Система по п. 76, когда одно или несколько первых оснований не природного

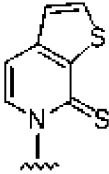
происхождения представляют собой  (TPT3), одно или несколько вторых

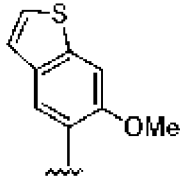
оснований не природного происхождения представляют собой  (CNMO), и
когда одно или несколько первых оснований не природного происхождения представляют

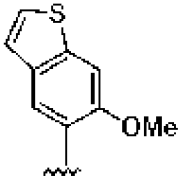
собой  (CNMO), одно или несколько вторых оснований не природного

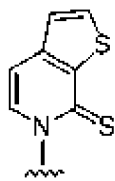
происхождения представляют собой  (TPT3), где волнистая линия указывает на
связь с рибозильной группой.

80. Система по п. 76, когда одно или несколько первых оснований не природного

происхождения представляют собой  (TPT3), одно или несколько вторых

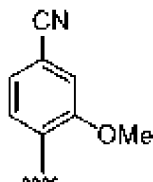
оснований не природного происхождения представляют собой  (PTMO), и
когда одно или несколько первых оснований не природного происхождения представляют

собой  (PTMO), одно или несколько вторых оснований не природного

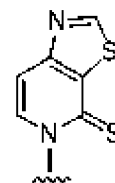


происхождения представляют собой (TPT3), где волнистая линия указывает на связь с рибозильной группой.

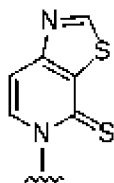
81. Система по п. 76, когда одно или несколько первых оснований не природного



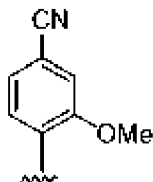
происхождения представляют собой (CNMO), одно или несколько вторых



оснований не природного происхождения представляют собой (TAT1), и когда одно или несколько первых оснований не природного происхождения представляют собой

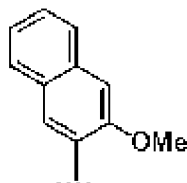


(TAT1), одно или несколько вторых оснований не природного происхождения

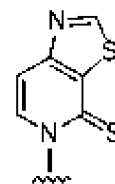


представляют собой (CNMO), где волнистая линия указывает на связь с рибозильной группой.

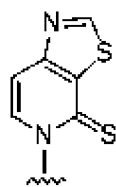
82. Система по п. 76, когда одно или несколько первых оснований не природного



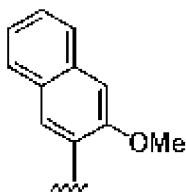
происхождения представляют собой (NaM), одно или несколько вторых



оснований не природного происхождения представляют собой (TAT1), и когда одно или несколько первых оснований не природного происхождения представляют собой

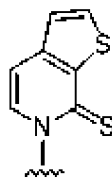


(TAT1), одно или несколько вторых оснований не природного происхождения

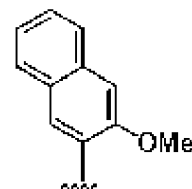


представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

83. Система по п. 76, когда одно или несколько первых оснований не природного

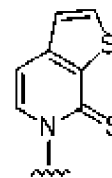


происхождения представляют собой (TPT3) и одно или несколько вторых

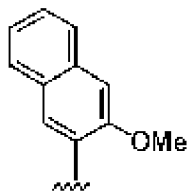


оснований не природного происхождения представляют собой (NaM), где волнистая линия указывает на связь с рибозильной группой.

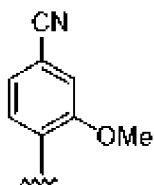
84. Система по любому из пп. 71-74, отличающаяся тем, что одно или несколько



первых оснований не природного происхождения выбраны из (TPT3),



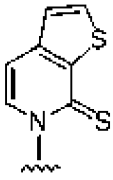
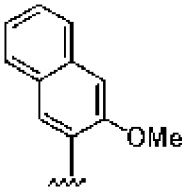
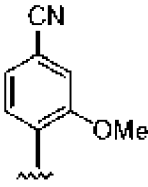
(NaM) и



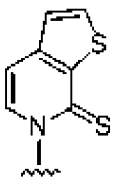
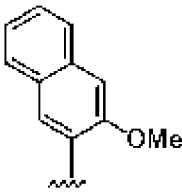
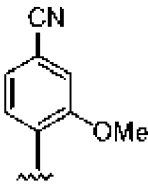
(CNMO), где волнистая линия указывает на связь с

рибозильной группой.

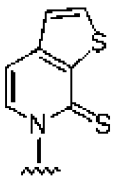
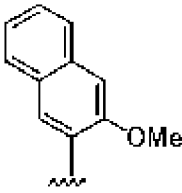
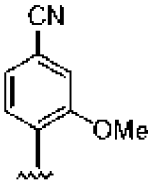
85. Система по п. 71, отличающаяся тем, что по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в кодоне иРНК, где одно или несколько первых оснований не природного

происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

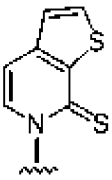
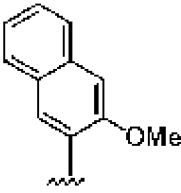
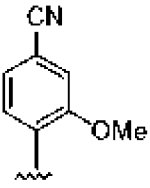
86. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в кодоне иРНК, где одно или несколько первых оснований не природного

происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

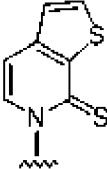
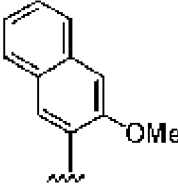
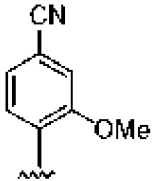
87. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон иРНК содержит три непрерывных нуклеотидных основания (N-N-N), где одно или несколько первых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в кодоне иРНК, где одно или несколько первых оснований не природного

происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

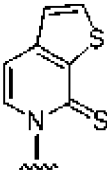
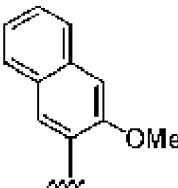
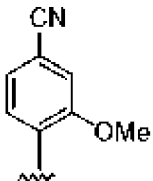
88. Система по п. 71, отличающаяся тем, что, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько вторых оснований не природного происхождения (X) расположены в первом положении (X-N-N) в антикодоне тРНК, где одно или несколько вторых оснований не

природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

89. Система по п. 71, отличающаяся тем, что, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько вторых оснований не природного происхождения (X) расположены в среднем положении (N-X-N) в антикодоне тРНК, где одно или несколько вторых оснований не

природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

90. Система по п. 71, отличающаяся тем, что, по меньшей мере, один антикодон тРНК содержит три непрерывных нуклеотидных основания (N-N-N); и где одно или несколько вторых оснований не природного происхождения (X) расположены в последнем положении (N-N-X) в антикодоне тРНК, где одно или несколько вторых оснований не

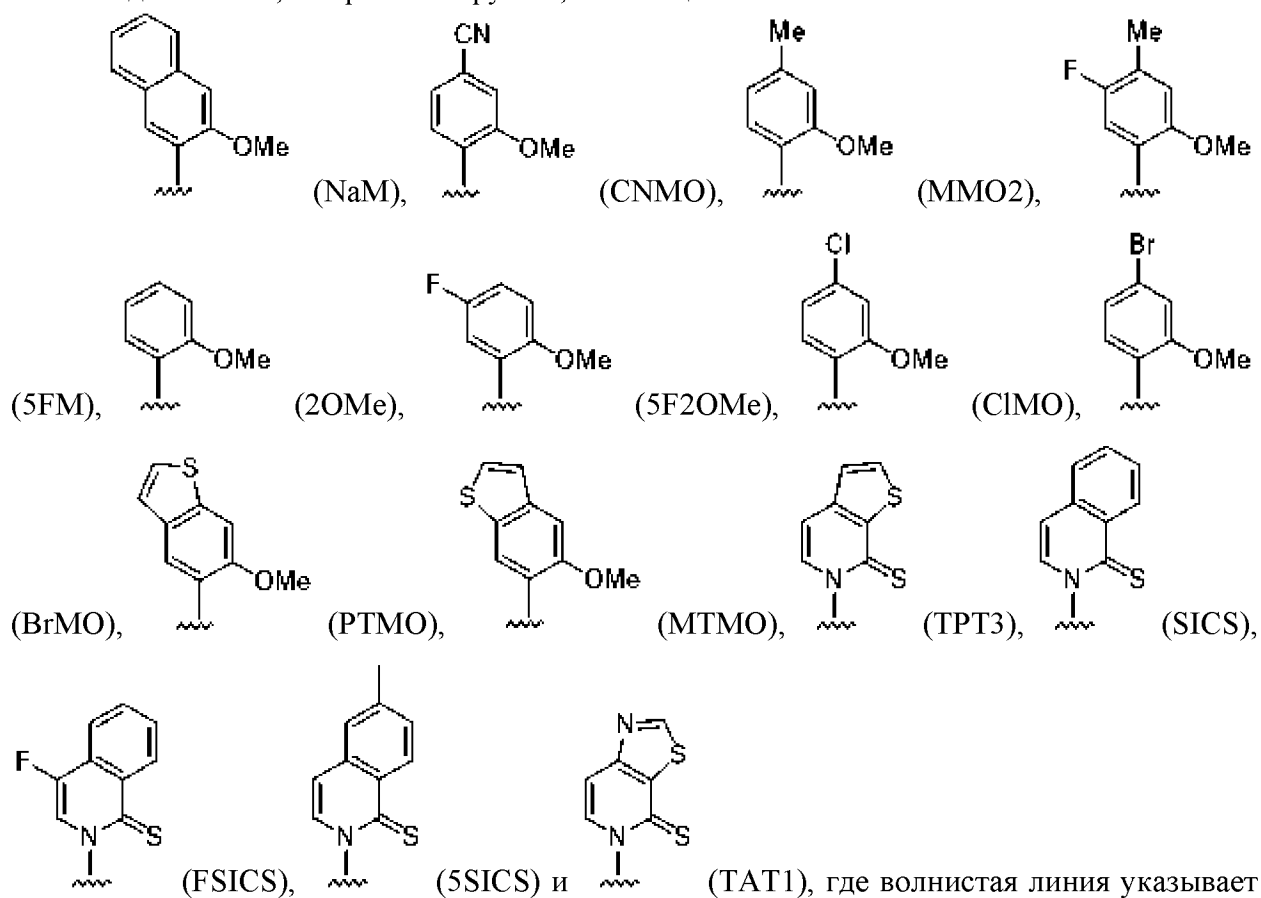
природного происхождения выбраны из  (TPT3),  (NaM) и  (CNMO), и где волнистая линия указывает на связь с рибозильной группой.

91. Система по п. 71, отличающаяся тем, что каждый из, по меньшей мере, одного кодона и, по меньшей мере, одного антикодона независимо содержит три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон содержит одно или несколько первых оснований не природного происхождения (X), расположенных в первом положении (X-N-N) кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в последнем положении (N-N-Y) антикодона.

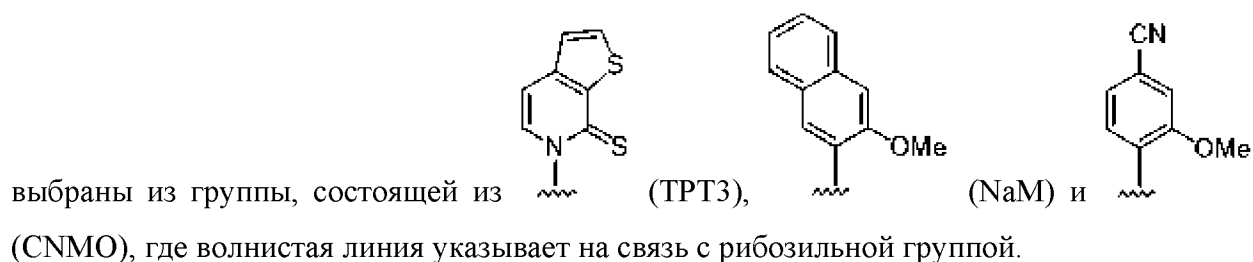
92. Система по п. 91, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

93. Система по любому из пп. 91-92, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и

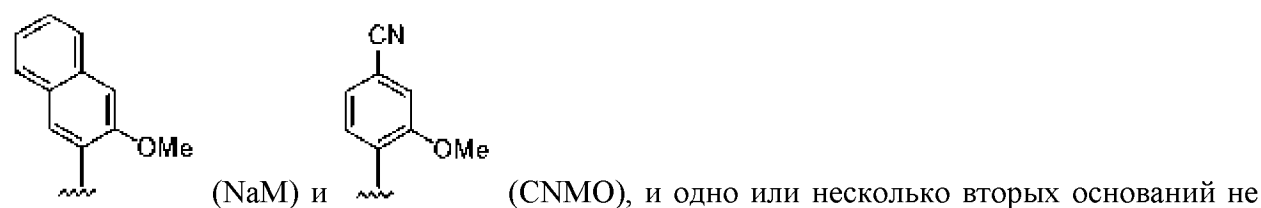
одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

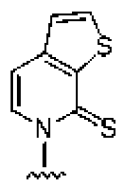


94. Система по п. 93, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК,



95. Система по п. 94, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, выбраны из



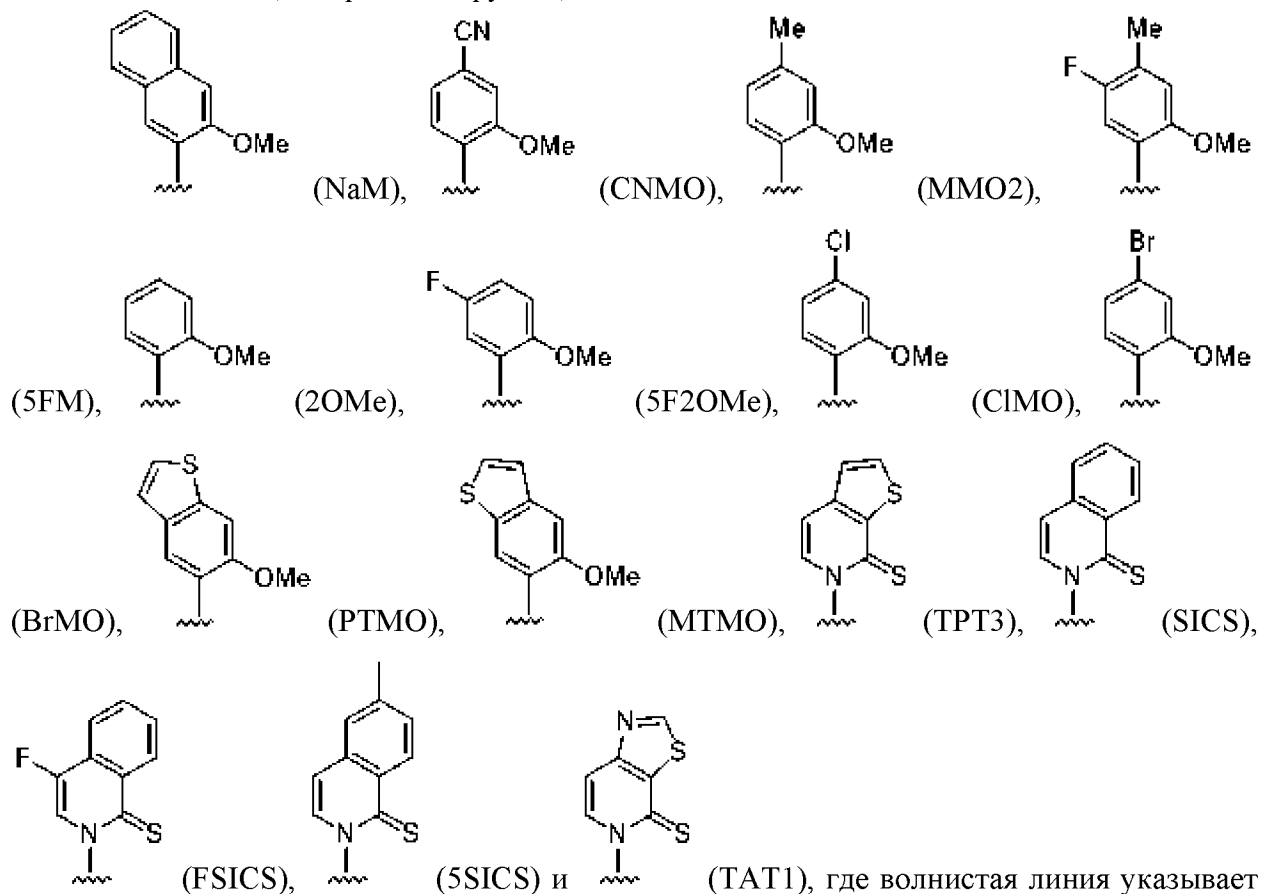


(TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

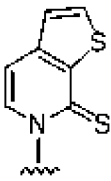
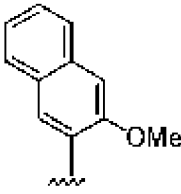
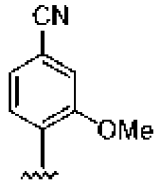
96. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон и, по меньшей мере, один антикодон каждый независимо содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенный в среднем положении (N-X-N), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в среднем положении (N-Y-N) антикодона.

97. Система по п. 96, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

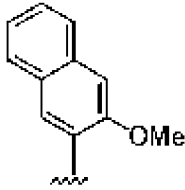
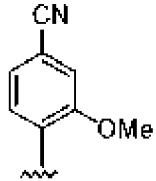
98. Система по любому из пп. 96-97, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

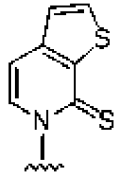


99. Система по п. 98, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК,

выбраны из группы, состоящей из  (TPT3),  (NaM) и  (CNMO), где волнистая линия указывает на связь с рибозильной группой.

100. Система по п. 99, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, выбраны из

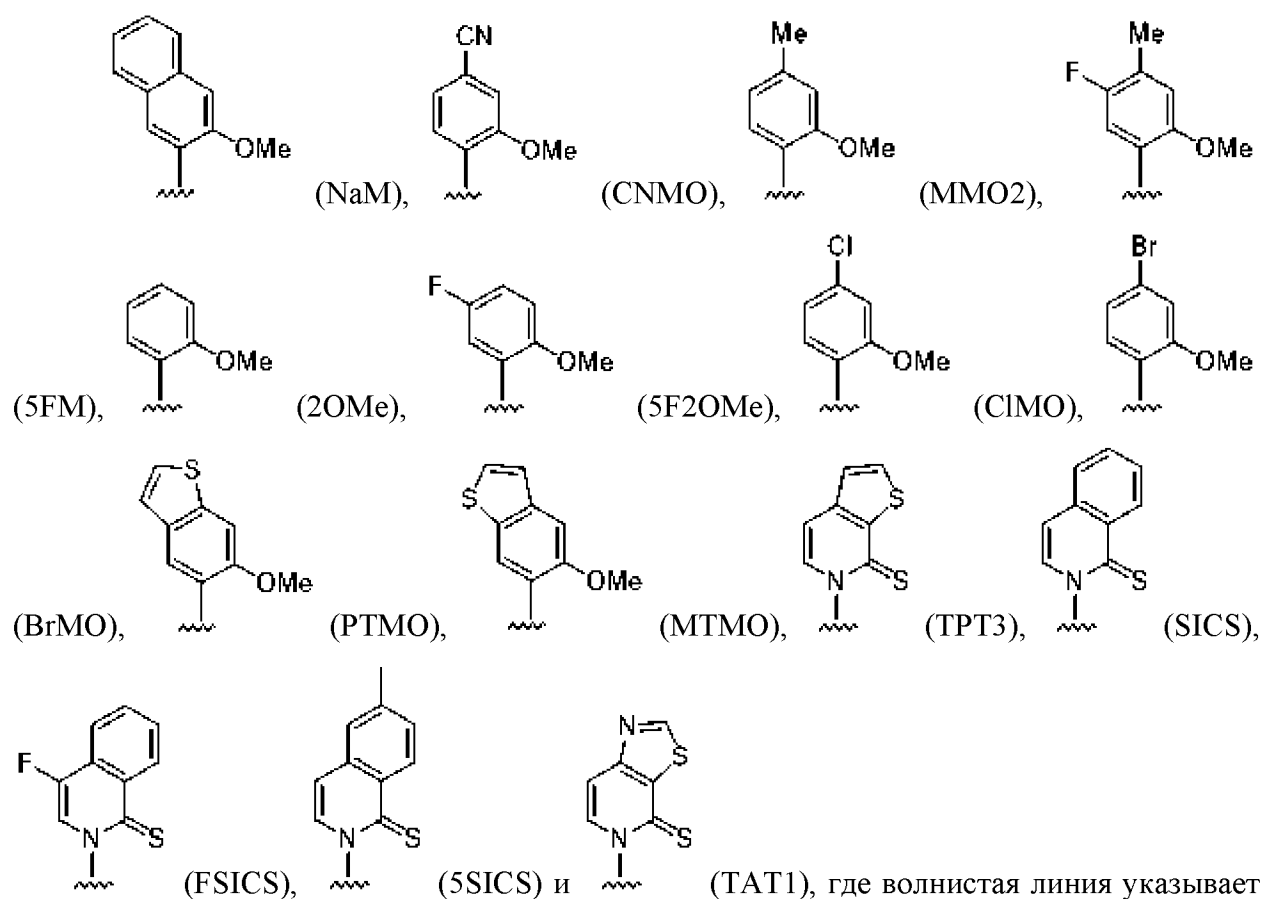
 (NaM) и  (CNMO), и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК представляет собой

 (TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

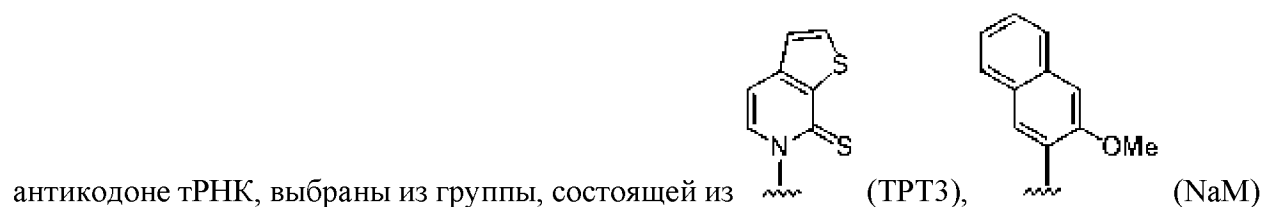
101. Система по п. 71, отличающаяся тем, что, по меньшей мере, один кодон и, по меньшей мере, один антикодон каждый независимо содержат три непрерывных нуклеотидных основания (N-N-N), и где, по меньшей мере, один кодон в иРНК содержит одно или несколько первых оснований не природного происхождения (X), расположенный в последнем положении (N-N-X), по меньшей мере, одного кодона, и, по меньшей мере, один антикодон в тРНК содержит одно или несколько вторых оснований не природного происхождения (Y), расположенных в первом положении (Y-N-N) антикодона.

102. Система по п. 101, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, являются одинаковыми или разными.

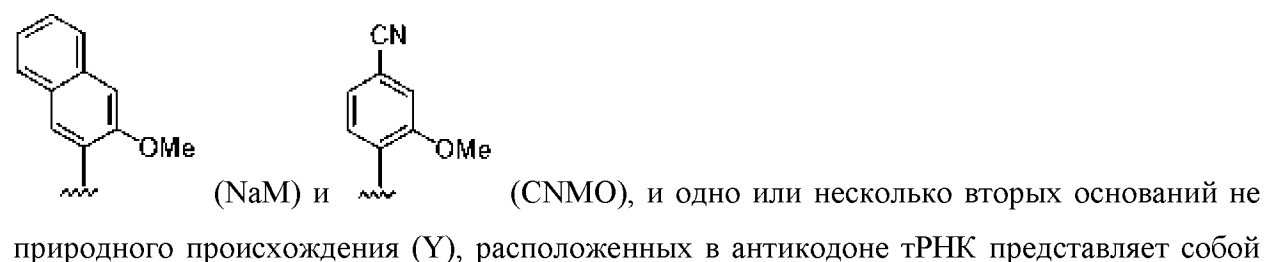
103. Система по любому из пп. 101-102, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в антикодоне тРНК, выбраны из группы, состоящей из

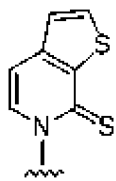


104. Система по п. 103, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, и одно или несколько вторых оснований не природного происхождения (Y), расположенных в



105. Система по п. 104, отличающаяся тем, что одно или несколько первых оснований не природного происхождения (X), расположенных в кодоне иРНК, выбраны из





(TPT3), где в каждом случае волнистая линия указывает на связь с рибозильной группой.

106. Система по любому из пп. 71-105, отличающаяся тем, что, по меньшей мере, один кодон в иРНК выбран из АХС, GXC или GXU, где X представляет собой одно или несколько первых оснований не природного происхождения.

107. Система по непосредственно предыдущему пункту, отличающаяся тем, что, по меньшей мере, один антикодон в тРНК выбран из GYU, GYC и AYC, и Y представляет собой одно или несколько вторых оснований не природного происхождения.

108. Система по п. 107, отличающаяся тем, что, по меньшей мере, один кодон в иРНК представляет собой АХС, и, по меньшей мере, один антикодон в тРНК представляет собой GYU.

109. Система по п. 107, отличающаяся тем, что, по меньшей мере, один кодон в иРНК представляет собой GXC, и, по меньшей мере, один антикодон в тРНК представляет собой GYC.

110. Система по п. 107, отличающаяся тем, что, по меньшей мере, один кодон в иРНК представляет собой GXU, и, по меньшей мере, один антикодон представляет собой AYC.

111. Система по любому из пп. 71-110, отличающаяся тем, что тРНК получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

112. Система по любому из пп. 71-111, отличающаяся тем, что тРНК синтетаза получена из *Methanococcus jannaschii*, *Methanosarcina barkeri*, *Methanosarcina mazei* или *Methanosarcina acetivorans*.

113. Система по любому из пп. 71-112, полученная in vitro или бесклеточная.

114. Система по любому из пп. 71-113, содержащая клеточный лизат.

115. Система по любому из пп. 71-113, которая представляет собой реконструированную систему из очищенных компонентов.

116. Система по любому из пп. 71-112, которая находится в эукариотической клетке.

117. Система по п. 116, отличающаяся тем, что эукариотическая клетка представляет собой клетку человека.

118. Система по п. 116, отличающаяся тем, что эукариотическая клетка представляет собой клетку НЕК293Т.

119. Система по п. 116, отличающаяся тем, что эукариотическая клетка представляет собой клетку хомяка.

120. Система по п. 119, отличающаяся тем, что клетка хомяка представляет собой клетку яичника китайского хомяка (СНО).

121. Система по любому из пп. 71-120, отличающаяся тем, что аминокислота не природного происхождения:

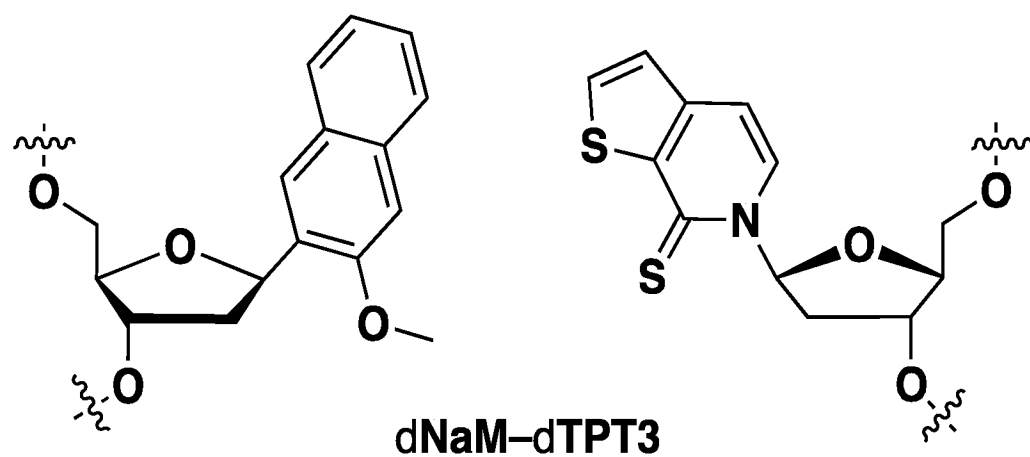
представляет собой аналог лизина;
 содержит ароматическую боковую цепь;
 содержит азидогруппу;
 содержит алкиновую группу; или
 содержат альдегидную или кетонную группу.

122. Система по любому из пп. 71-121, отличающаяся тем, что аминокислота не природного происхождения выбрана из группы, состоящей из N6-((азидоэтокси)-карбонил)-L-лизина (AzK), N6-((пропаргилэтокси)-карбонил)-L-лизина (PraK), BCN-L-лизина, норборненлизина, TCO-лизина, метилтетразинлизина, аллилоксикарбониллизина, 2-амино-8-оксононановой кислоты, 2-амино-8-оксооктановой кислоты, п-ацетил-L-фенилаланина, п-азидометил-L-фенилаланина (pAMF), п-йодо-L-фенилаланина, м-ацетилфенилаланина, 2-амино-8-оксононановой кислоты, п-пропаргиллоксифенилаланина, п-пропаргилфенилаланина, 3-метилфенилаланина, L-Дора, фторированного фенилаланина, изопропил-L-фенилаланина, п-азидо-L-фенилаланина, п-ацил-L-фенилаланина, п-бензоил-L-фенилаланина, п-бромфенилаланина, п-амино-L-фенилаланина, изопропил-L-фенилаланина, О-аллилтирозина, О-метил-L-тирозина, О-4-аллил-L-тирозина, 4-пропил-L-тирозина, фосфонтитрозина, три-О-ацетил-GlcNAc-серина, L-фосфосерина, фосфоносерина, L-3-(2-нафтил)аланина, 2-амино-3-((2-((3-(бензилокси)-3-оксопропил)амино)этил)селанил)пропановой кислоты, 2-амино-3-(фенилселанил)пропановой кислоты, селеноцистеина, N6-(((2-азидобензил)окси)карбонил)-L-лизина, N6-(((3-азидобензил)окси)карбонил)-L-лизина и N6-(((4-азидобензил)окси)карбонил)-L-лизина.

123. Система по любому из пп. 71-122, отличающаяся тем, что аминокислота не природного происхождения представляет собой N6-((азидоэтокси)-карбонил)-L-лизин (AzK).

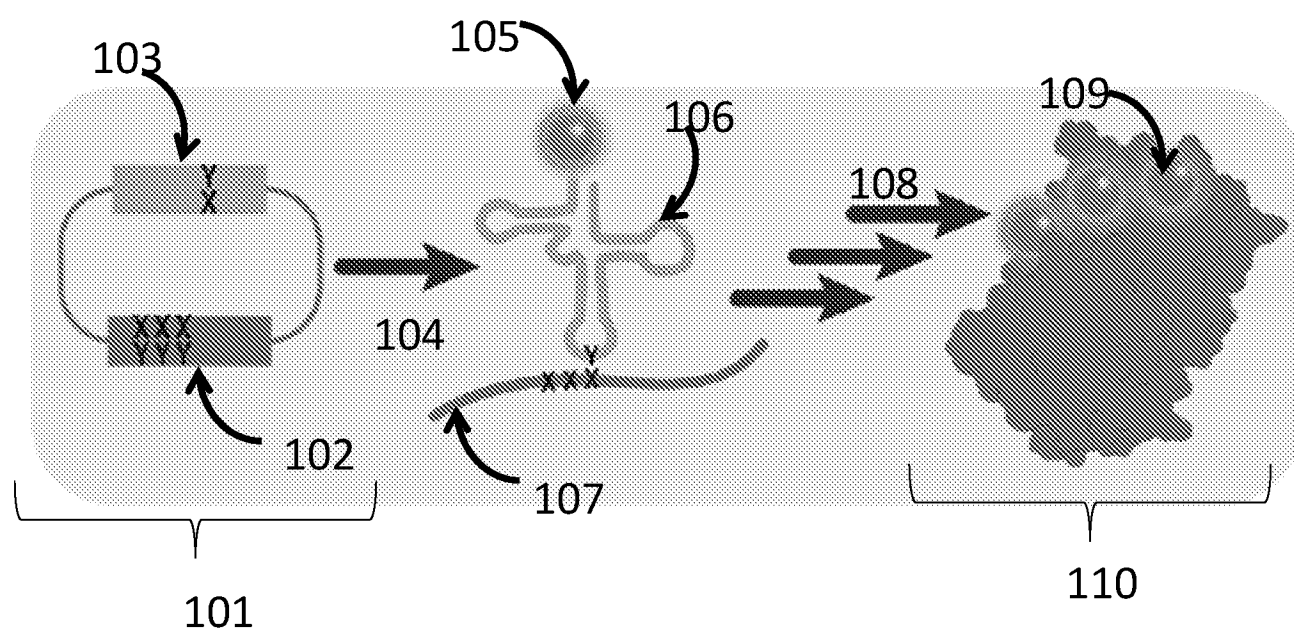
124. Система по любому из пп. 71-123, отличающаяся тем, что тРНК заряжена аминокислотой не природного происхождения.

ФИГ.1А

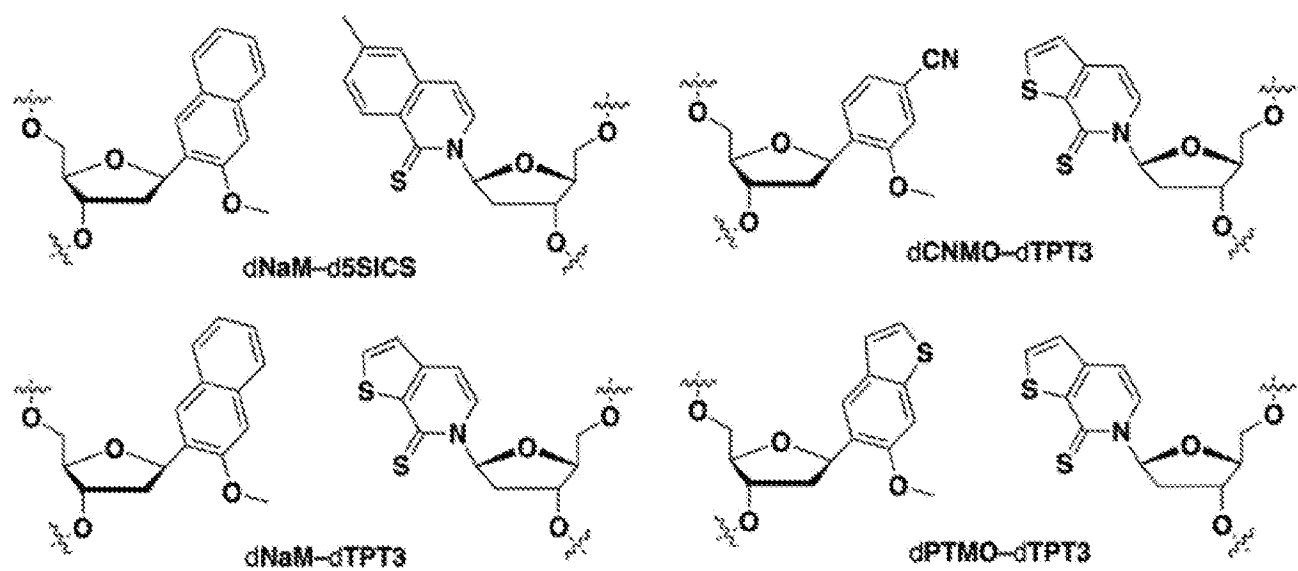


ФИГ.1В

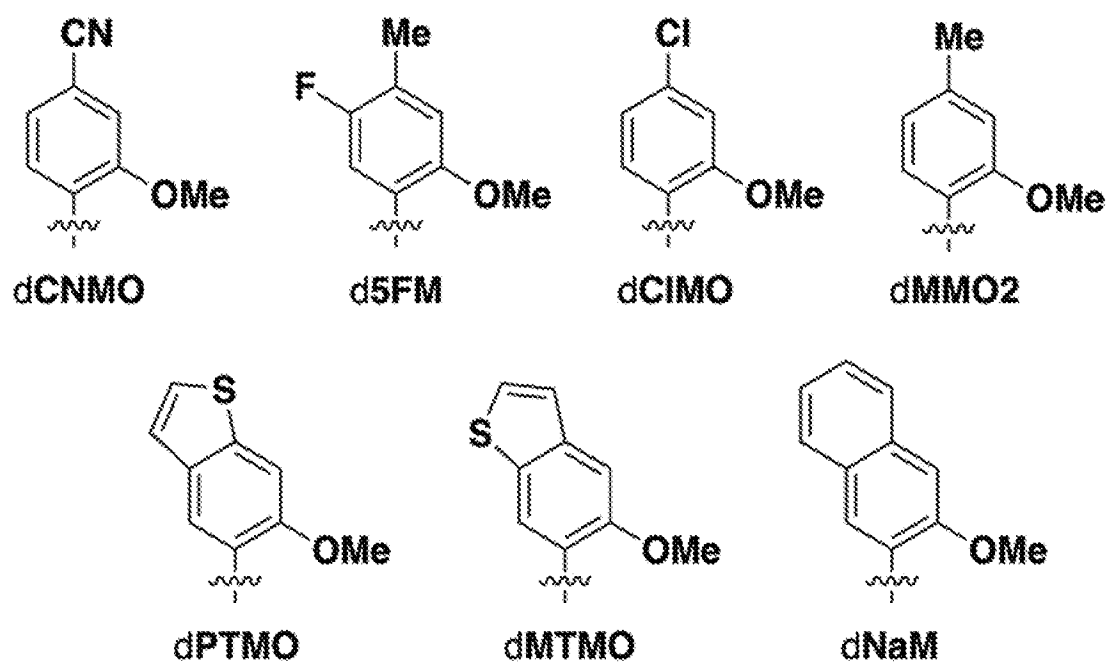
100



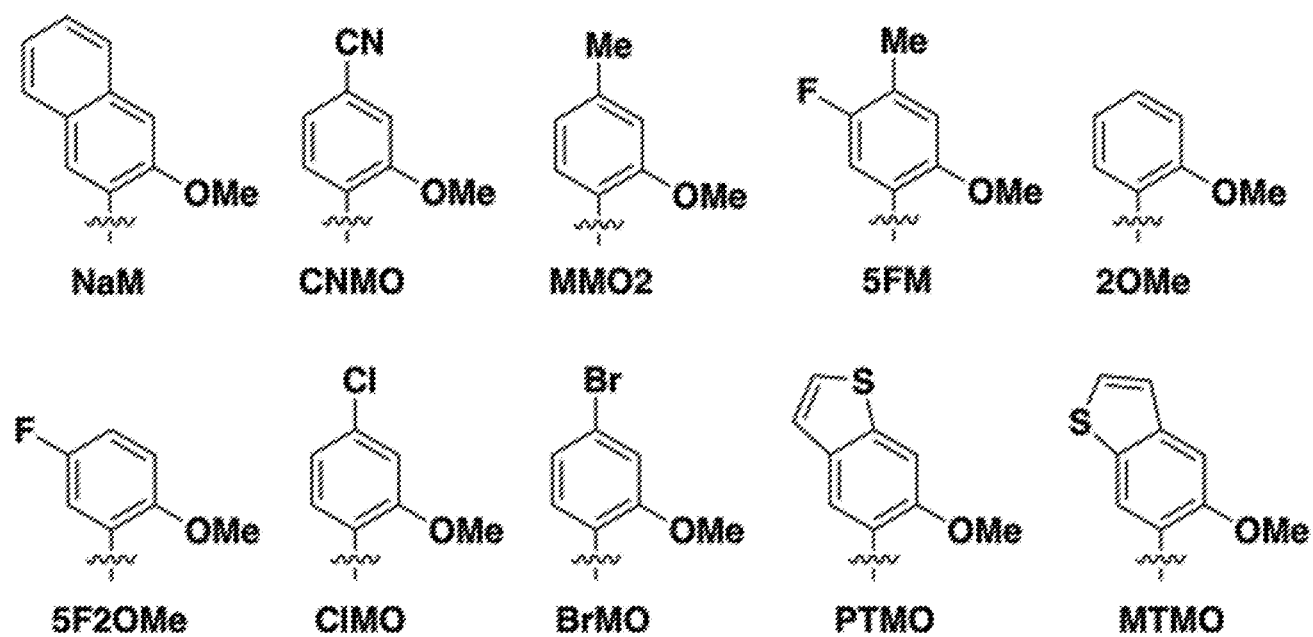
ФИГ.1С



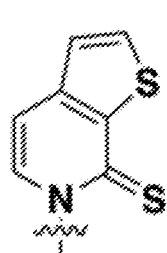
ФИГ.2



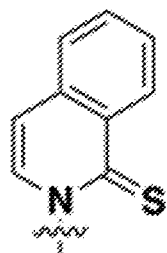
ФИГ.3А



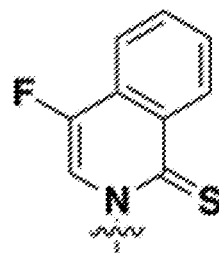
ФИГ.3В



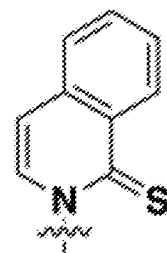
TPT3



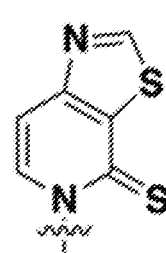
SICS



FSICS

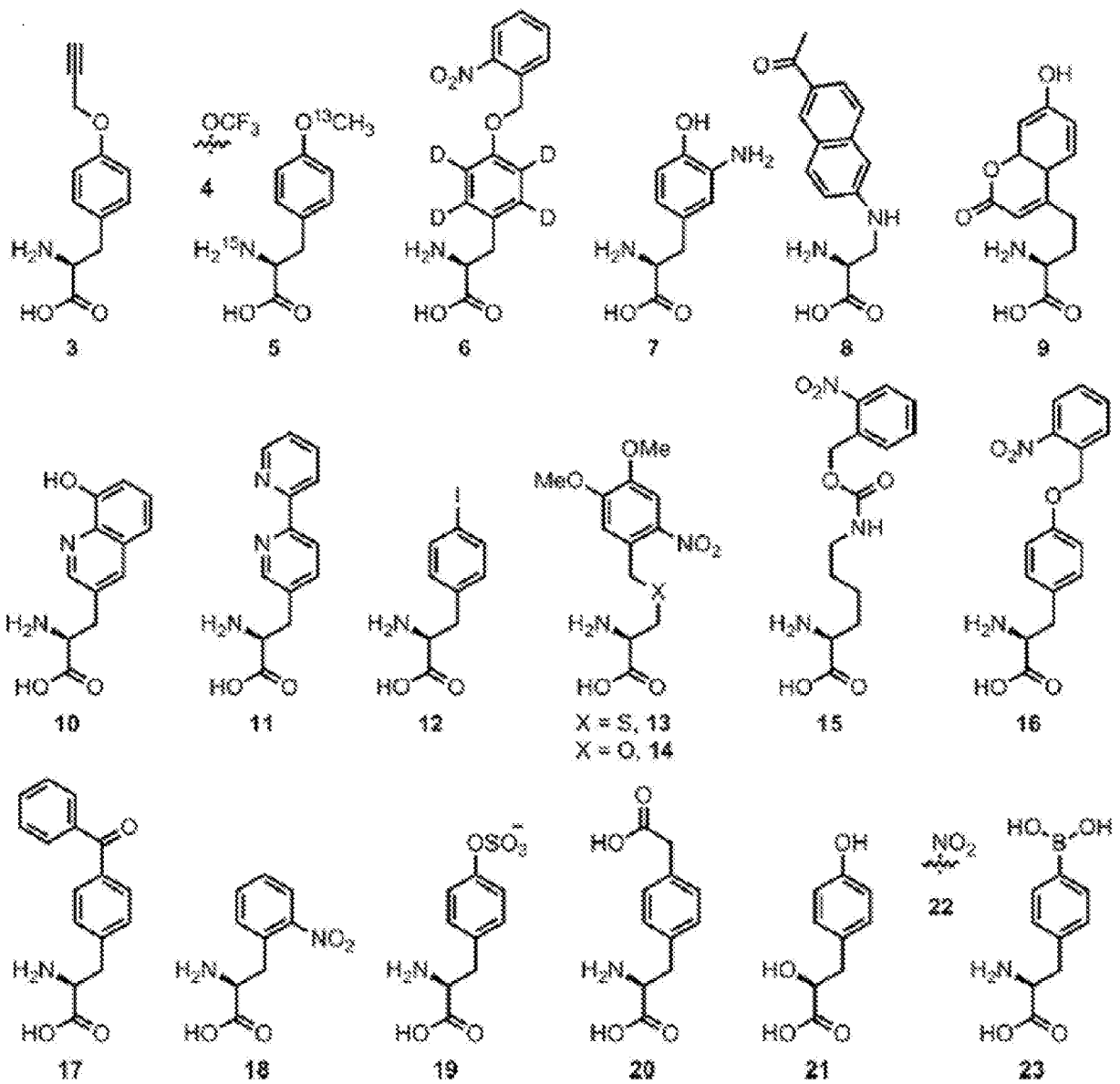


5SICS

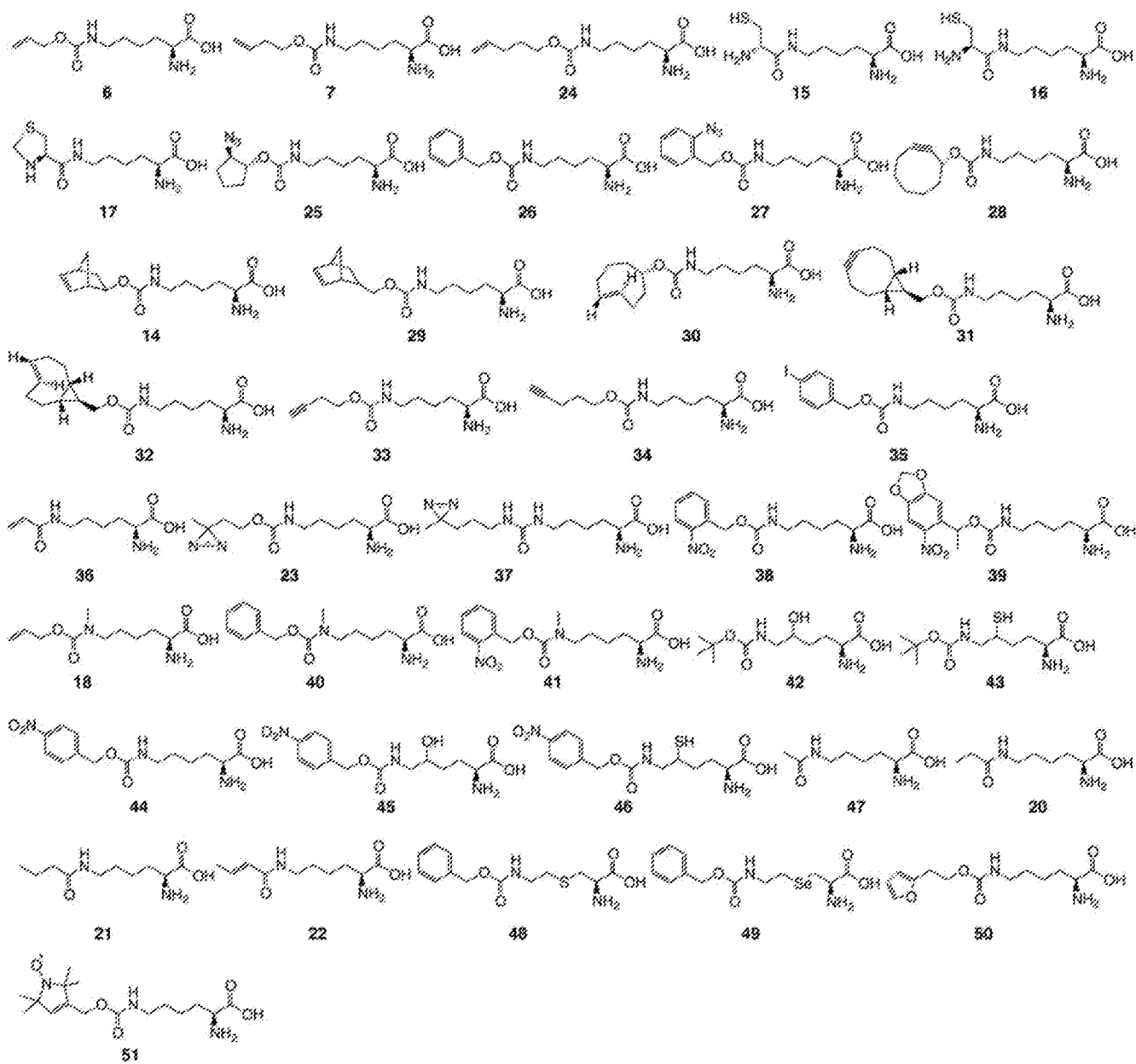


TAT1

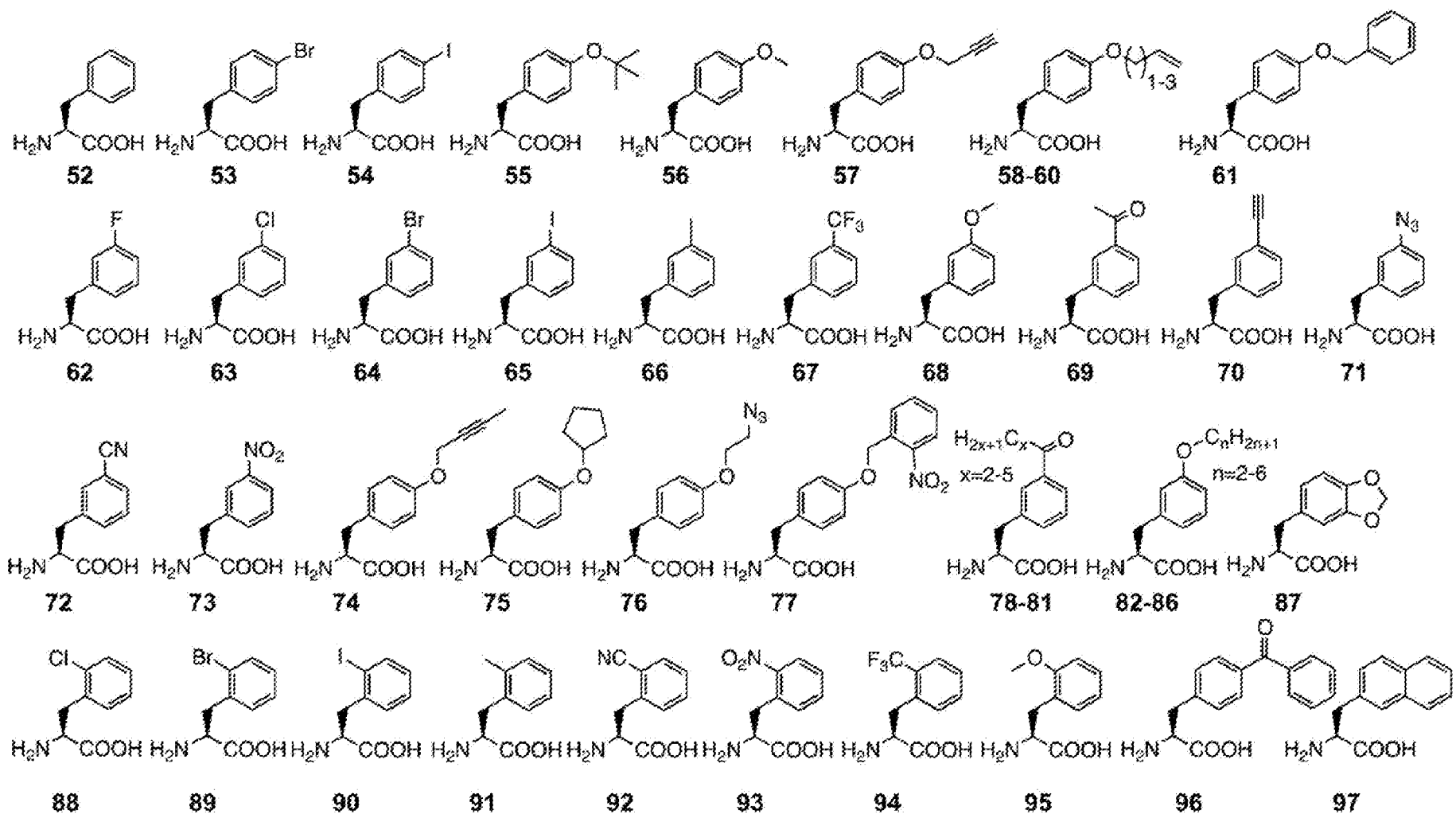
ФИГ.4А



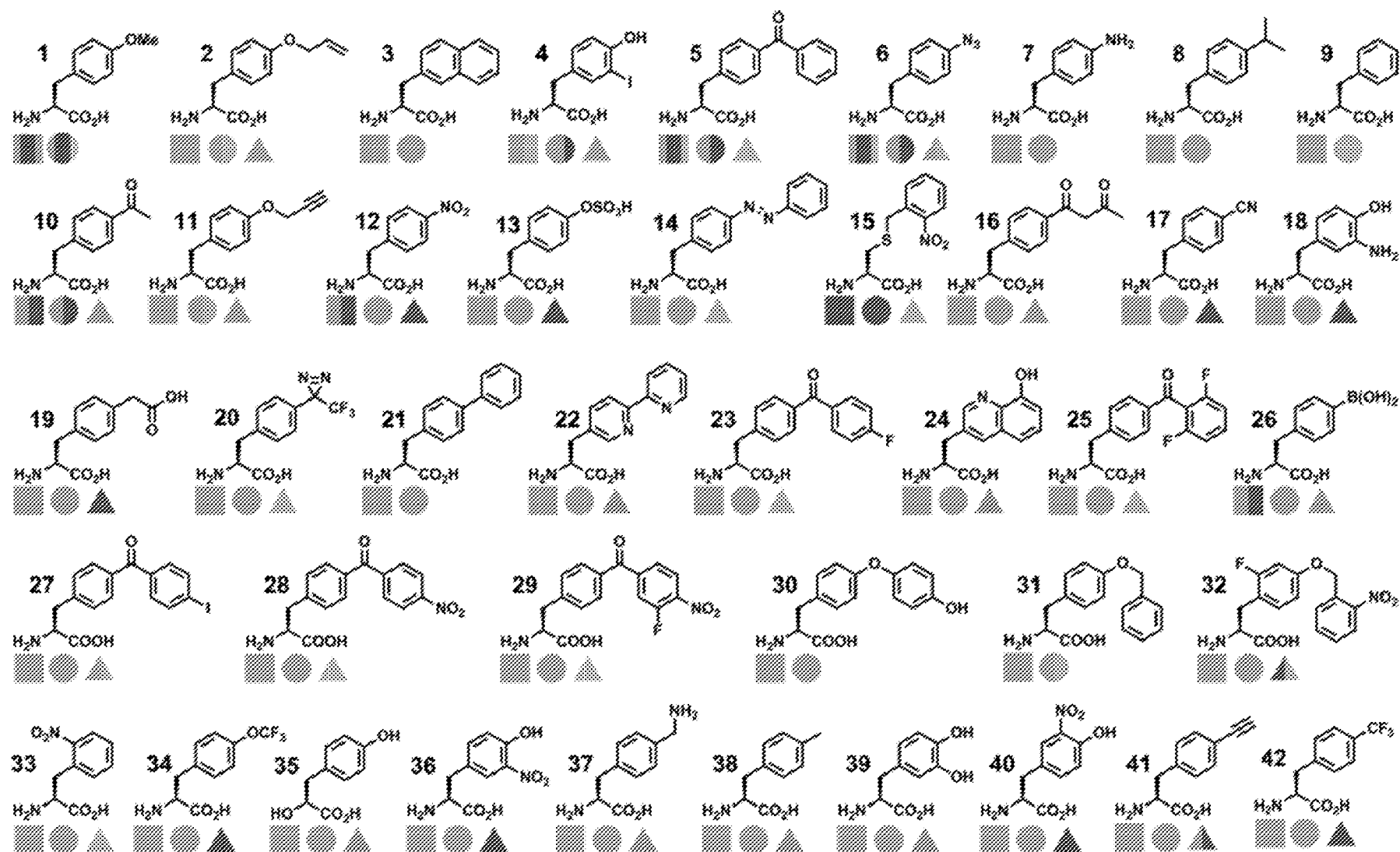
ФИГ.4В



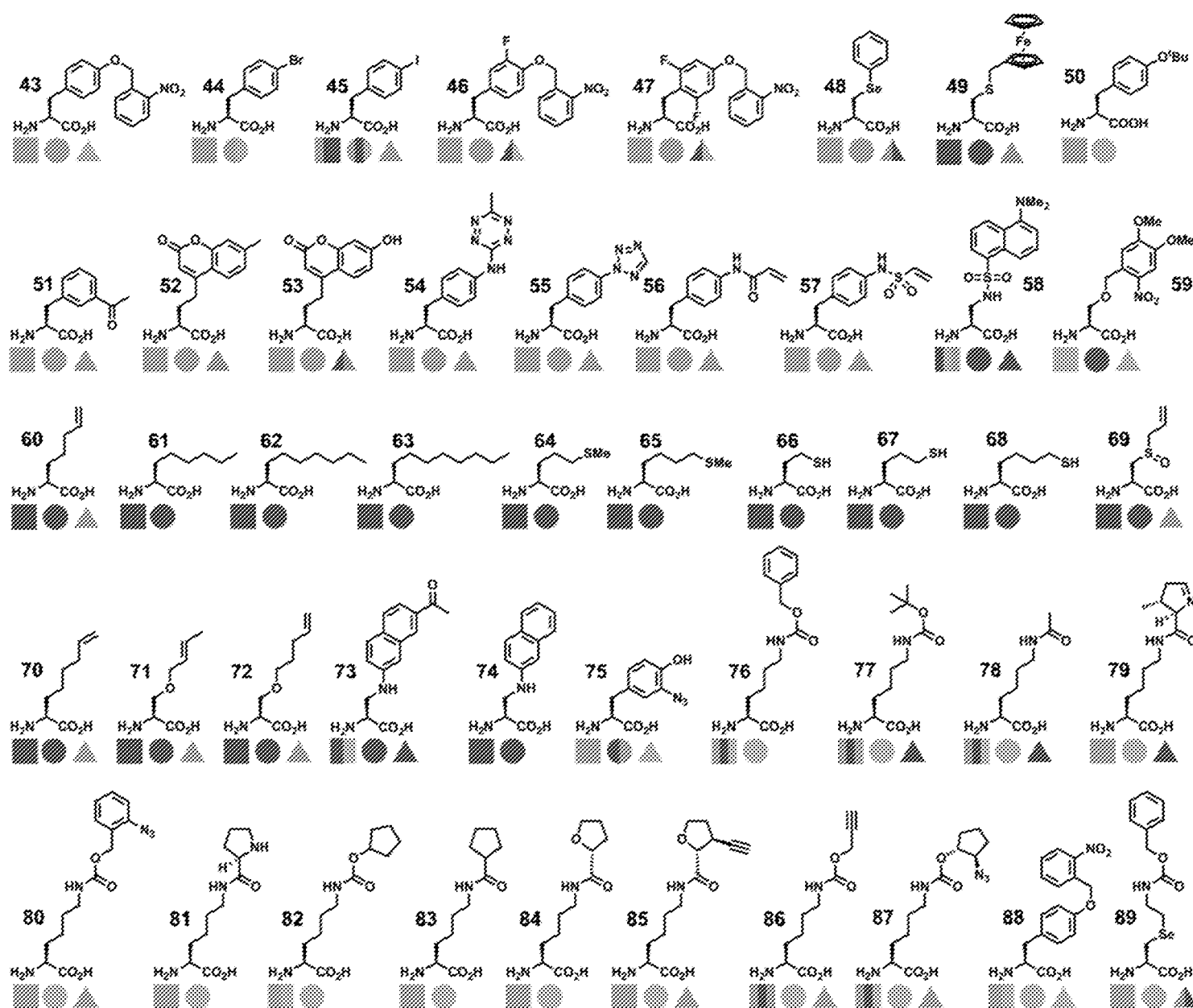
ФИГ.4С



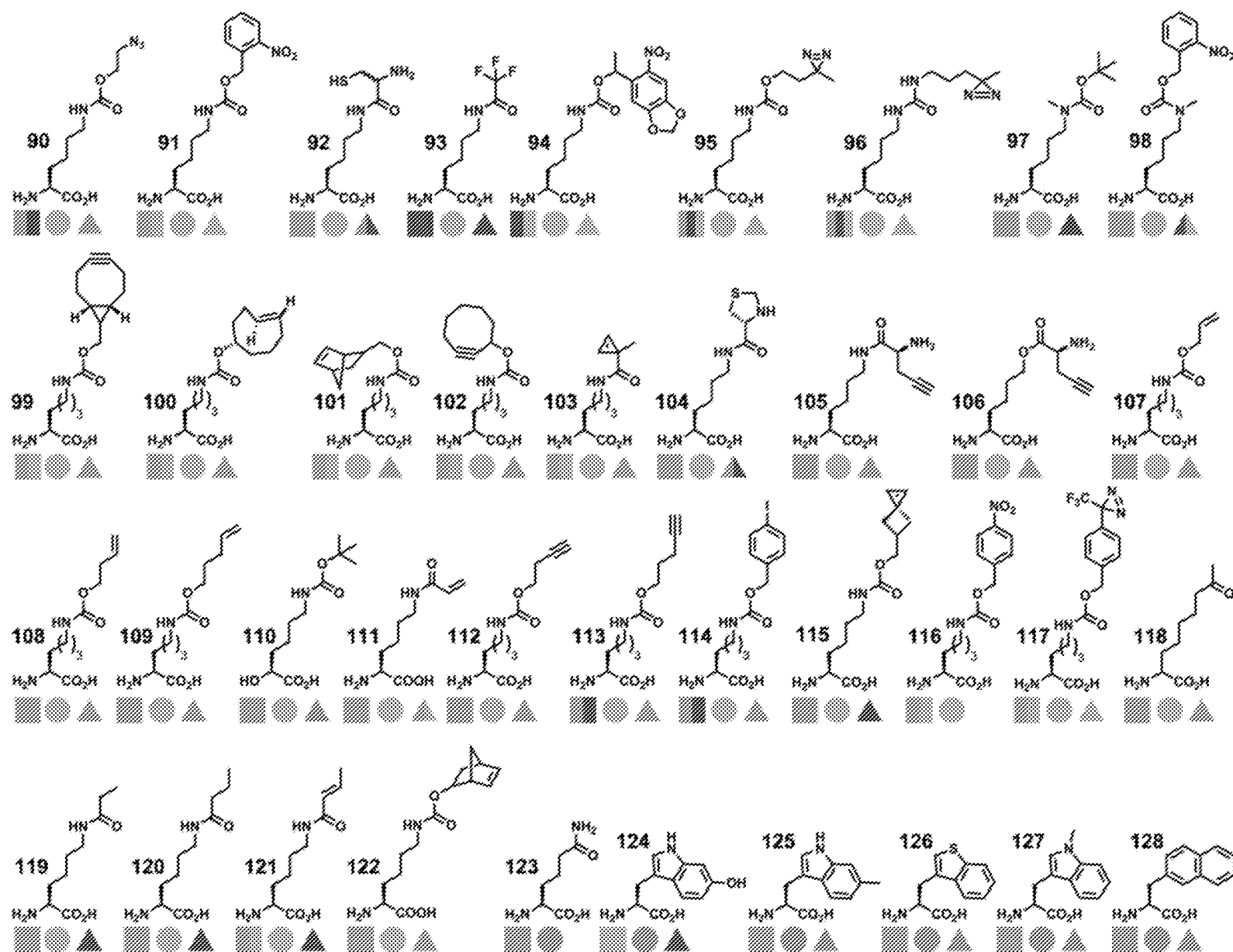
ФИГ.4D



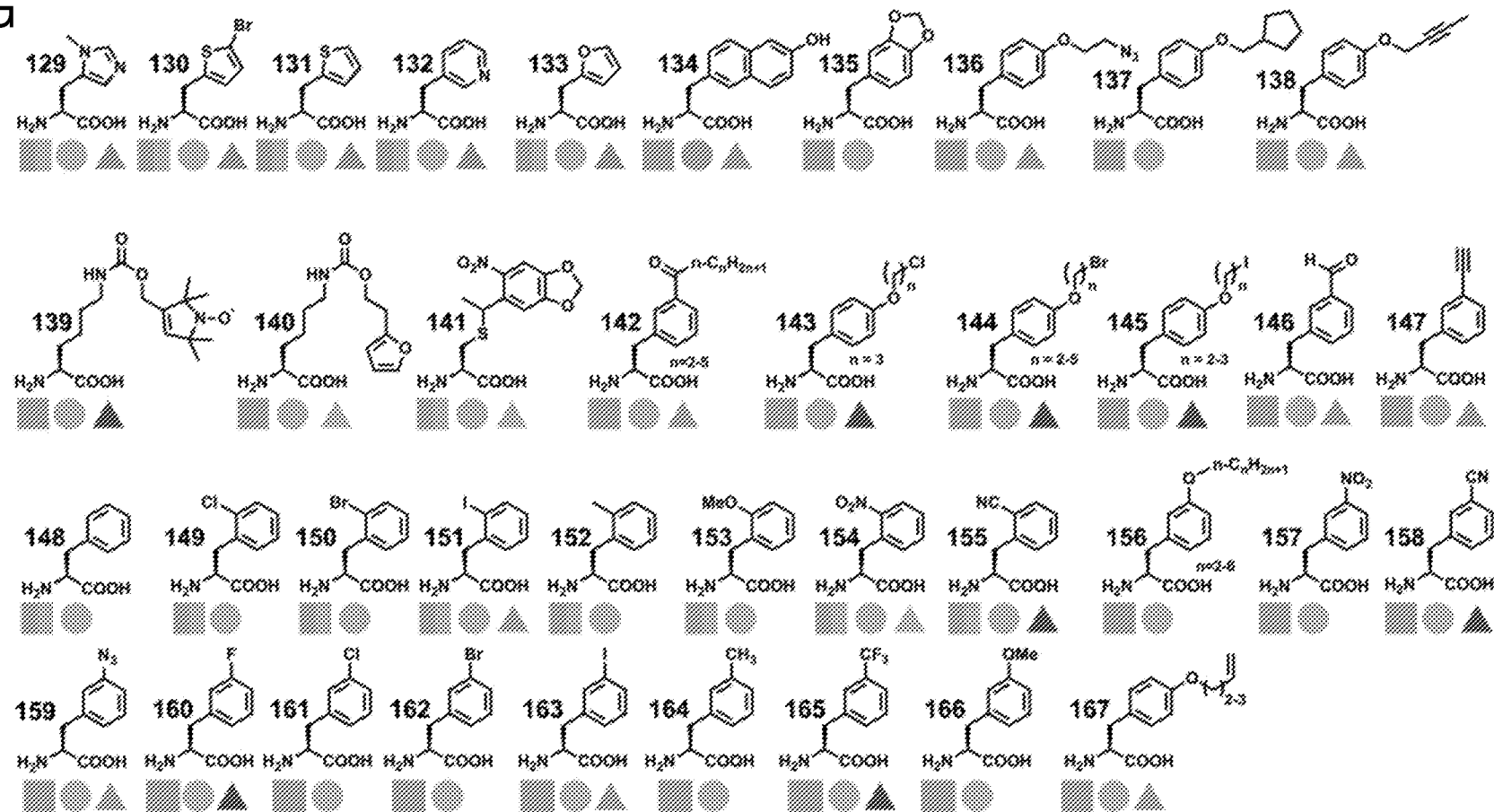
ФИГ.4Е



ФИГ.4F



ФИГ.4G



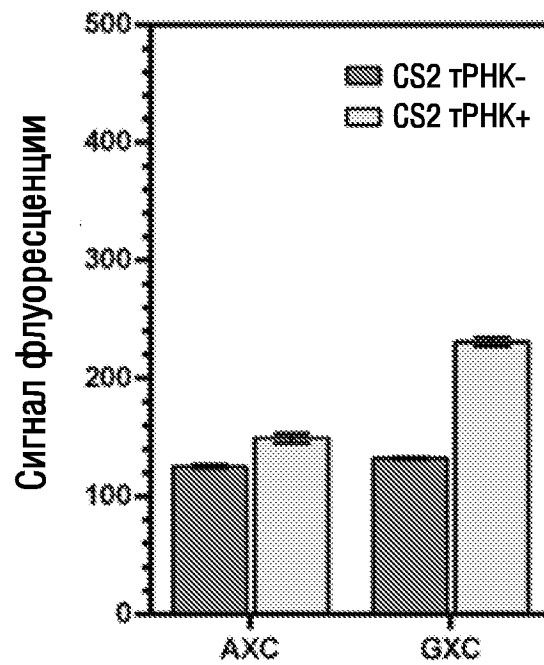
Организм

Пара аатрансфераза/тРНК

Функция

- | | | |
|--------------------------------------|-----------|---|
| E. Coli | Mj Tyr | Биоконъюгация |
| Другие прокариоты | Ec Leuc | Репортер |
| Дрожжи | Ec Tyr | Природный PTM (или предшественник) |
| Эукариотические (клеточная культура) | Mm/Mb ryl | Фотореактивный (поперечное сшивание/заклучение) |
| Эукариотические (организм) | другие | Новая активность (хелатирующая, окислительно-восстановительная) |

ФИГ.5А

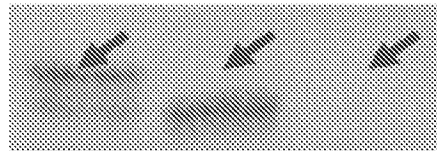


Кодоны неприродного происхождения

ФИГ.5В

EGFP иРНК	GXC	GXC	-
GXC тРНК	+	-	-
TAMRA	+	+	+

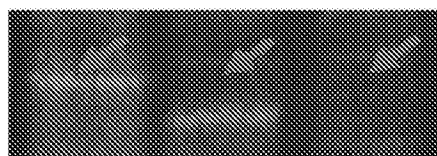
Анти-GFP



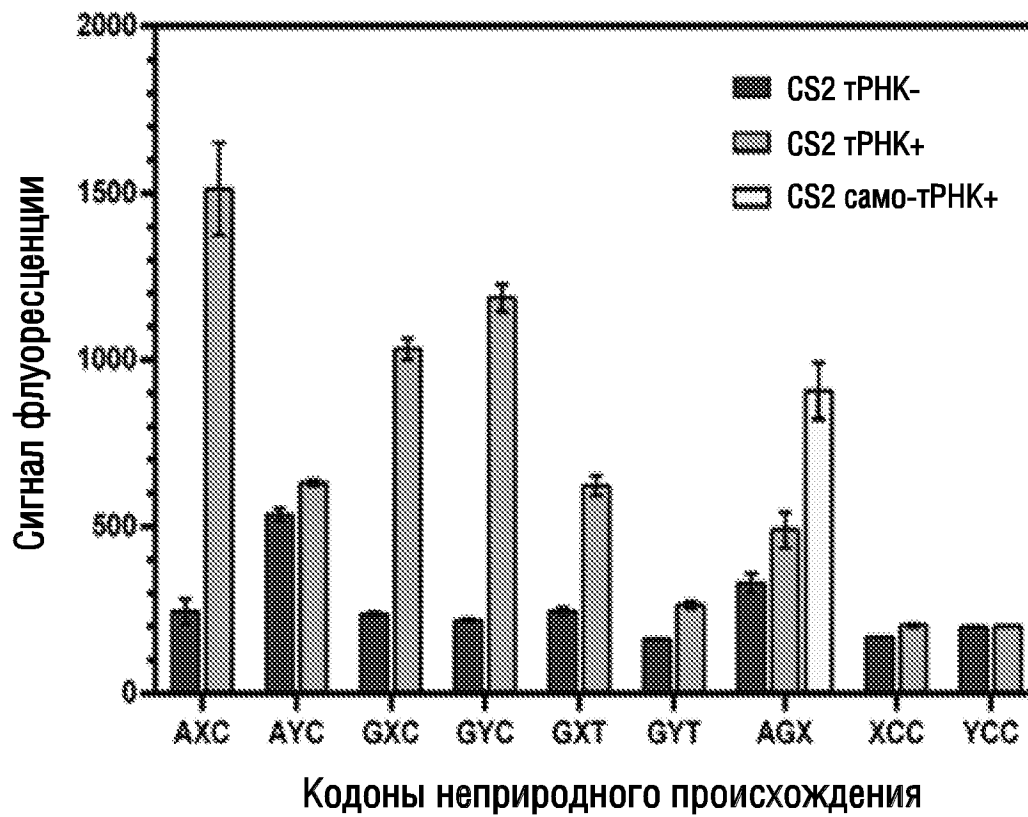
TAMRA



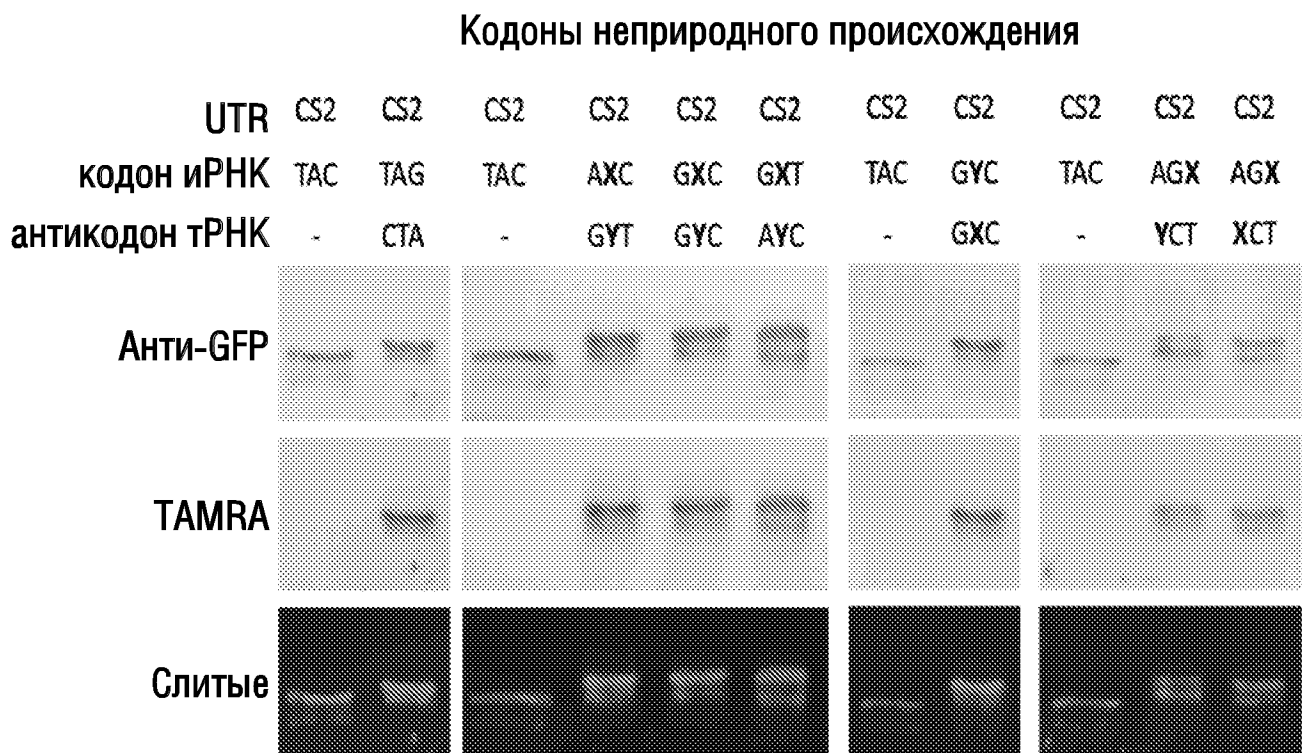
Слитые



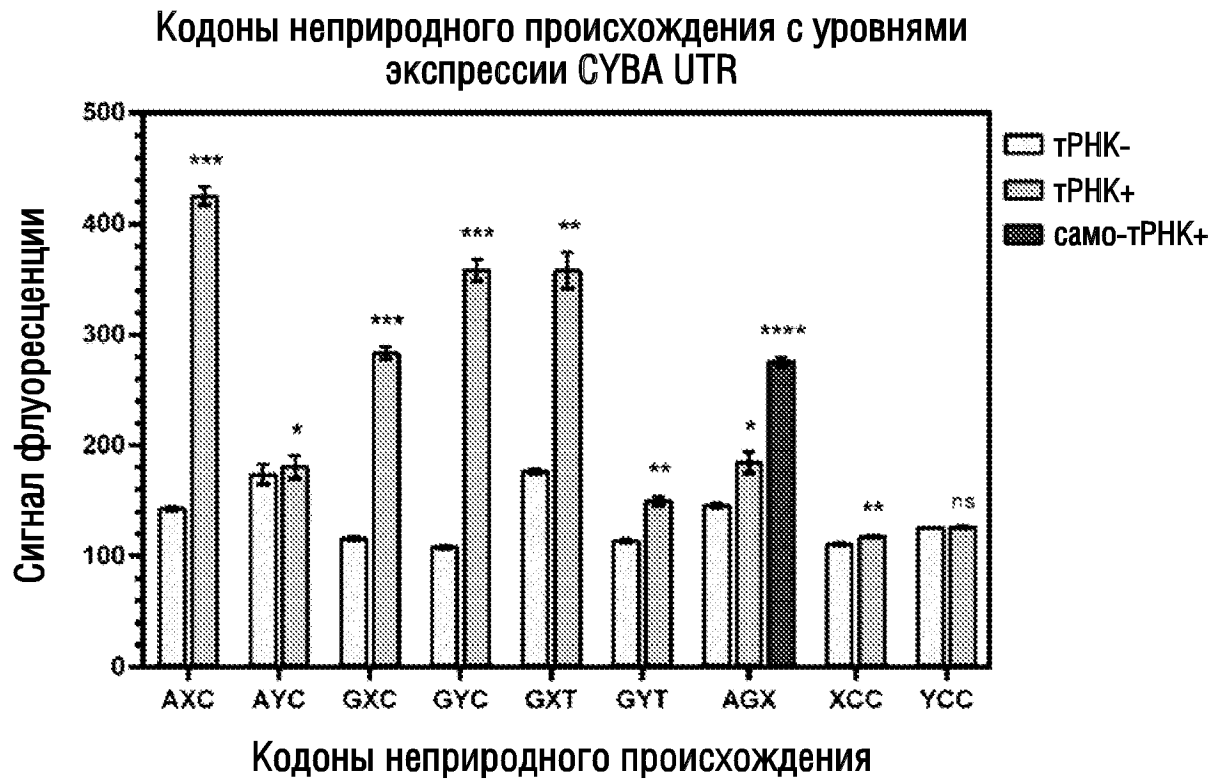
ФИГ.6А



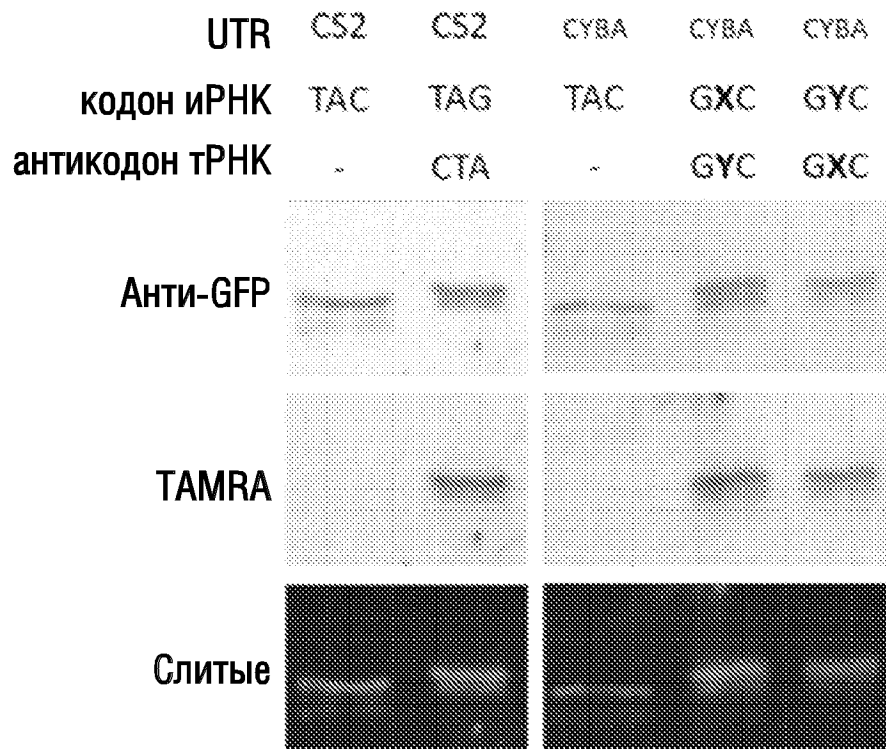
ФИГ.6В



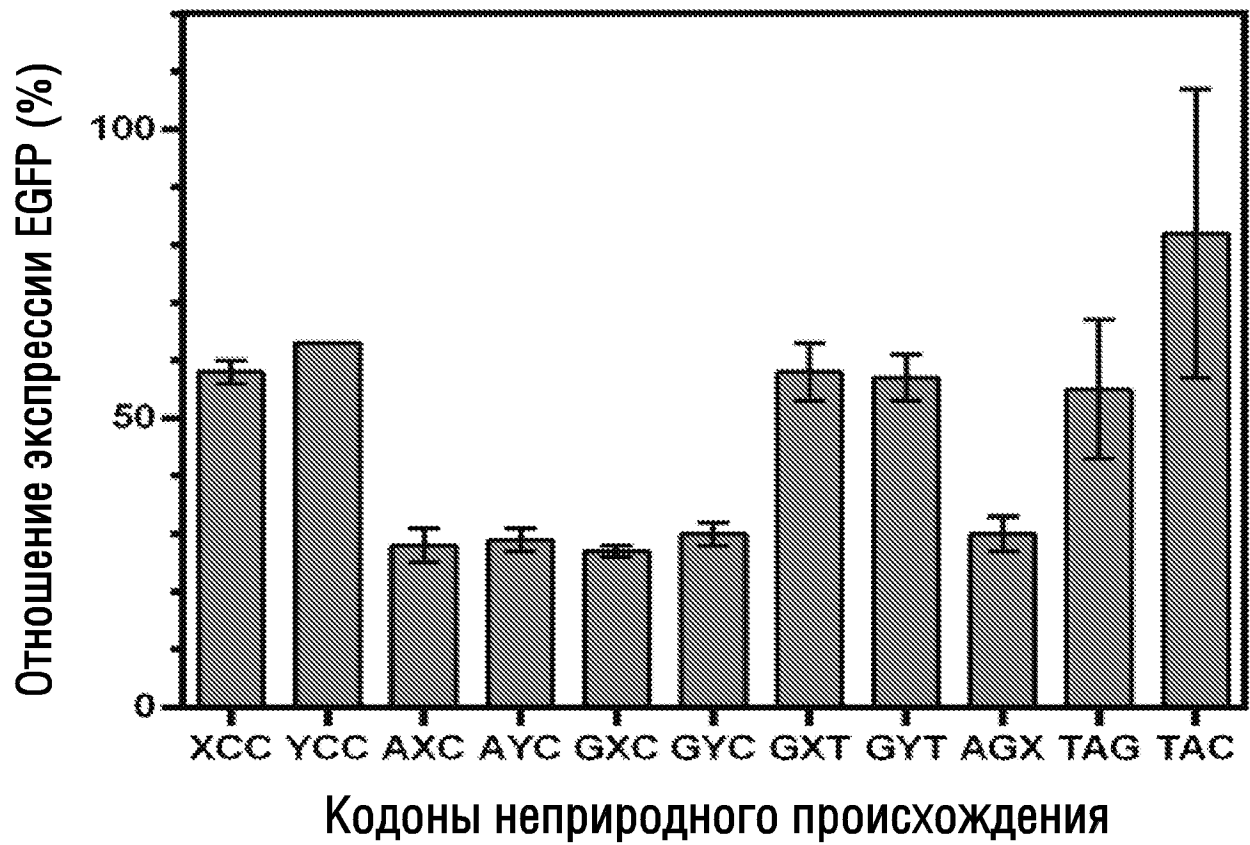
ФИГ.7А



ФИГ.7В



ФИГ.7С



ФИГ.7D

