

(19)



Евразийское  
патентное  
ведомство

(11) 038576

(13) B1

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента  
2021.09.16

(51) Int. Cl. C07H 19/00 (2006.01)  
C12N 15/09 (2006.01)

(21) Номер заявки  
201690274

(22) Дата подачи заявки  
2014.08.08

(54) СПОСОБ САЙТ-СПЕЦИФИЧЕСКОГО ФЕРМЕНТАТИВНОГО МЕЧЕНИЯ  
НУКЛЕИНОВЫХ КИСЛОТ IN VITRO ВВЕДЕНИЕМ НЕ ВСТРЕЧАЮЩИХСЯ В  
ПРИРОДЕ НУКЛЕОТИДОВ

(31) 61/863,649

(56) WO-A1-0023456  
WO-A1-2008067825  
WO-A2-2007085485  
WO-A1-03055898  
EP-A1-2130835

(32) 2013.08.08

(33) US

(43) 2016.07.29

(86) PCT/US2014/050423

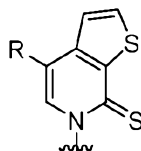
(87) WO 2015/021432 2015.02.12

(71)(73) Заявитель и патентовладелец:  
ДЗЕ СКРИППС РИСЁЧ  
ИНСТИТЮТ (US)

(72) Изобретатель:  
Роумсберг Флойд Е., Мальшев  
Денис А., Ли Линцзюнь, Лаверн  
Томас, Ли Чжэнтао (US)

(74) Представитель:  
Строкова О.В. (RU)

(57) Изобретение относится к соединениям, содержащим аналоги не встречающихся в природе нуклеотидов указанной формулы, которые образуют не встречающиеся в природе пары оснований во время опосредованной ДНК-полимеразой репликации ДНК или опосредованной РНК-полимеразой транскрипции РНК. Также раскрыты композиция для образования не встречающейся в природе пары оснований, двухцепочечный олигонуклеотидный дуплекс, способ проведения сайт-специфической функционализации двухцепочечного олигонуклеотидного дуплекса, а также олигонуклеотид.



038576 B1

038576 B1

### Заявление государственной поддержки

Настоящее изобретение было сделано при поддержке правительства в виде гранта под номером GM060005, предоставленного национальными институтами Здоровья США. Правительство США имеет определенные права на данное изобретение.

### Перекрестная ссылка на родственные заявки

Настоящая заявка испрашивает преимущество по предварительной заявке США № 61/863649, поданной 8 августа 2013 г., полное содержание которой включено в настоящий документ путем отсылки.

### Уровень техники

Олигонуклеотиды и их применения в корне изменили биотехнологию. Тем не менее, олигонуклеотиды, состоящие как из ДНК, так и из РНК, содержат только четыре природных нуклеотида, представляющих собой аденозин (А), гуанозин (G), цитозин (C), тимин (T) в случае ДНК, и четыре природных нуклеотида, представляющих собой аденозин (A), гуанозин (G), цитозин (c) и уридин (U) в случае РНК, что существенно ограничивает потенциальные функции и применения олигонуклеотидов.

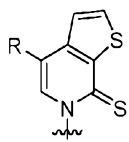
Возможность синтезировать/амплифицировать олигонуклеотиды (ДНК или РНК) определенной последовательности с помощью полимераз, например, применяя полимеразную цепную реакцию (ПЦР) или системы изотермической амплификации (например, транскрипция с T7 РНК-полимеразой), произвела революцию в биотехнологии. В дополнение ко всем потенциальным применениям в области нанотехнологии, это позволило освоить широкий спектр новых технологий, таких как *in vitro* эволюция посредством SELEX (Систематическая эволюция лигандов экспоненциальным обогащением (Systematic Evolution of Ligands by Exponential Enrichment)) РНК- и ДНК-аптамеров и ферментов. См., например, Oliphant A.R., Brandl C.J. & Struhl K. (1989), Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 proteins, *Mol. Cell. Biol.*, 9:2944-2949; Tuerk C. & Gold L. (1990), Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, 249:505-510; Ellington A.D. & Szostak J.W. (1990), *In vitro* selection of RNA molecules that bind specific ligands, *Nature*, 346:818-822.

К сожалению, данные применения ограничены лимитированным химическим/физическим разнообразием, присутствующим в природном генетическом алфавите (четыре природных нуклеотида А, С, G и T в ДНК и четыре природных нуклеотида А, С, G и U в РНК). Соответственно, большой интерес вызывают способы, которые позволили бы проводить ферментативный синтез/амплификацию олигонуклеотидов, сайт-специфически помеченных функциональными группами, которых нет у нуклеотидов природного генетического алфавита. В настоящее время доступные возможности получения сайт-специфических производных нуклеиновых кислот включают химический синтез на твердой подложке, комбинированный химический/ферментативный синтез и способы концевое мечения. Способы концевое мечения применимы только к концам олигонуклеотидов, в то время как химический синтез позволяет получать только короткие олигонуклеотиды (<200 нуклеотидов для ДНК и <70 нуклеотидов для РНК). Ферментативная функционализация зависит от ферментативного узнавания модификации, представляющей интерес, и, что более проблематично, не является сайт-специфической.

### Сущность изобретения

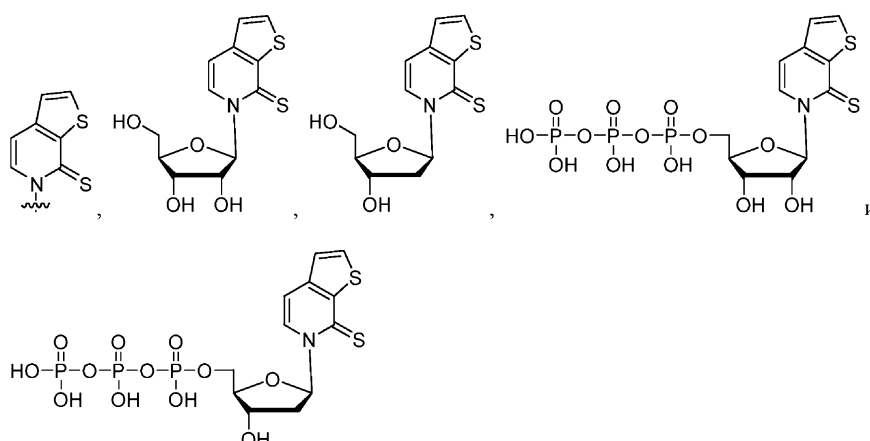
Композиции и способы, описанные в настоящем изобретении, основаны на расширении генетического алфавита *in vitro* и обеспечивают сайт-специфическое включение не встречающихся в природе нуклеотидов (unnatural nucleotides), как описано здесь, несущих реакционноспособные линкеры, приспособленные к взаимодействию с карго(cargo)-реагентами, содержащими группы, обладающие комплементарной реакционной способностью, или линкеры, несущие карго, присоединенные к ним, в любом положении любой ДНК- или РНК-последовательности, например, с помощью стандартных способов ПНР или изотермической транскрипции.

В различных вариантах осуществления настоящего изобретения обеспечивается соединение, содержащее аналог нуклеинового основания формулы



где R представляет собой водород, C<sub>2</sub>-C<sub>3</sub> алкенил, C<sub>2</sub>-C<sub>3</sub> алкинил, метокси, метантиол, метанселено, циано или азидо; или R представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу; и волнистая линия показывает точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку, причем рибозильный, дезоксирибозильный или дидезоксирибозильный остаток находится в свободной форме, присоединен к монофосфатной, дифосфатной, трифосфатной, α-тиотрифосфатной, β-тиотрифосфатной или γ-тиотрифосфатной группе или включен в РНК или ДНК. В одном варианте осуществления R представляет собой H. В одном варианте осуществления рибозильный или дезоксирибозильный остаток содержит трифосфатную или α-тиотрифосфатную группу, связанную с его 5'-

гидроксильной группой. В одном варианте осуществления рибозильный или дезоксирибозильный остаток включен в олигонуклеотидную цепь РНК или ДНК соответственно. В одном варианте осуществления R содержит группу  $-C\equiv C-CH_2NHR_3$ , где  $R_3$  представляет собой водород или представляет собой аминозащитную группу; ацетиленовую группу, подходящую для применения в клик-реакции с азидом; тиольную группу, подходящую для применения в реакции с группой, способной реагировать с тиольной группой; альдегидную группу, подходящую для применения в реакции с группой, способной реагировать с альдегидной группой; азидную группу, подходящую для применения в реакции с алкином; группу  $-C\equiv C-(CH_2)_n-C\equiv CH$ , где n равно 1, 2, 3, 4, 5, или 6; или группу  $-C\equiv C-(CH_2)_{n1}-O-(CH_2)_{n2}-C\equiv CH$ , где  $n1$  и  $n2$ , каждый независимо, равны 1, 2 или 3. В одном варианте осуществления соединение имеет формулу, выбранную из

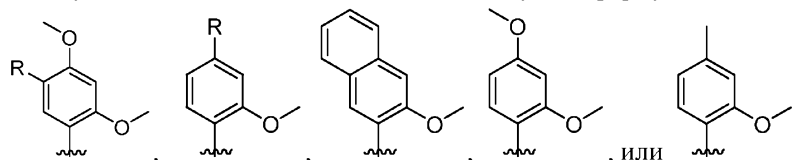


В различных вариантах осуществления настоящего изобретения обеспечивается композиция для образования не встречающейся в природе пары оснований, содержащая первый аналог нуклеинового

основания, имеющий формулу

или

И второй аналог нуклеинового основания, имеющий любую из формул



где R независимо представляет собой водород,  $C_2-C_3$  алкил,  $C_2-C_3$  алкинил, метокси, метантиол, метансело, циано или азидо; или R представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу; и волнистая линия показывает точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку, причем рибозильный, дезоксирибозильный или дидезоксирибозильный остаток находится в свободной форме, присоединен к монофосфатной, дифосфатной, трифосфатной,  $\alpha$ -тиотрифосфатной,  $\beta$ -тиотрифосфатной или  $\gamma$ -тиотрифосфатной группе или включен в РНК или ДНК.

В различных вариантах осуществления настоящего изобретения обеспечивается двухцепочечный олигонуклеотидный дуплекс, в котором первая олигонуклеотидная цепь содержит описанное в настоящей заявке соединение и вторая комплементарная олигонуклеотидная цепь содержит комплементарное, способное к образованию пары нуклеиновое основание в ее сайте спаривания комплементарных оснований. В одном варианте осуществления изобретения первая цепь содержит нуклеиновое основание, име-

ющее формулу

и вторая цепь содержит комплементарное, способное к образованию пары нуклеиновое основание,

выбранное из

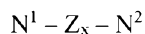
и

в ее сайте спаривания комплементарных оснований.

В различных вариантах осуществления настоящего изобретения обеспечивается способ проведения

сайт-специфической функционализации двухцепочечного олигонуклеотидного дуплекса, включающий включение не встречающегося в природе нуклеинового основания, выбранного из описанного в настоящей заявке соединения, в первую олигонуклеотидную цепь; затем синтез второй олигонуклеотидной цепи, которая комплементарна первой олигонуклеотидной цепи, причем вторая олигонуклеотидная цепь содержит нуклеиновое основание, комплементарное не встречающемуся в природе нуклеиновому основанию в сайт-специфической комплементарной позиции в нем, в таких условиях, что первая олигонуклеотидная цепь и вторая олигонуклеотидная цепь образуют двухцепочечный олигонуклеотидный дуплекс; затем контактирование двухцепочечного олигонуклеотидного дуплекса, включающего не встречающееся в природе нуклеиновое основание, с группой, обладающей реакционной способностью, комплементарной заместителю R не встречающегося в природе нуклеинового основания, в условиях, подходящих для осуществления реакции взаимодействия между заместителем R нуклеинового основания и группой, обладающей комплементарной реакционной способностью, с получением сопряженного линкера, чтобы обеспечить функционализированный двухцепочечный олигонуклеотидный дуплекс, где заместитель R не встречающегося в природе нуклеинового основания представляет собой C<sub>2</sub>-C<sub>3</sub> алкенил, C<sub>2</sub>-C<sub>3</sub> алкинил, метокси, метантиол, метанселено, циано или азидо; или заместитель R не встречающегося в природе нуклеинового основания представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу.

В различных вариантах осуществления настоящего изобретения обеспечивается олигонуклеотид, имеющий формулу



где N<sup>1</sup> представляет собой нуклеотид или концевую фосфатную группу; N<sup>2</sup> представляет собой нуклеотид или концевую гидроксильную группу; Z представляет собой соединение по п.1, и x означает целое число от 1 до 20.

#### Краткое описание фигур

На фиг. 1 проиллюстрировано образование пар dTPT3-dNaM, d5SICS-dNaM, d5SICS-dMMO2, d5SICS-dDMO в ДНК или РНК.

На фиг. 2 проиллюстрированы дериватизированные с помощью линкера нуклеотиды dTPT3R, d5SICSR, dMMO2R, dMMO2pR, dDMOR, dNaMpR, dNaMpR, dFEMO и dEMO, где R = 3-аминопропин-1-ил (обозначенный как A, например, dTPT3A); R = дихлорацетил-3-аминопропин-1-ил (обозначенный как PA); R = 4-оксагепта-1,6-диин-1-ил (обозначенный как CO); R = гепта-1,6-диин-1-ил (обозначенный как CC).

На фиг. 3 представлена общая схема стратегии пост-синтетического сайт-специфического мечения на основе фосфороатного способа.

На фиг. 4 представлена общая схема стратегии пост-синтетического сайт-специфического, основанного на аминировании мечения. Модифицированные линкером нуклеотиды могут быть также непосредственно включены в матричную ДНК с использованием стандартного твердофазного синтеза олигонуклеотидов и соответствующих фосфорамидитов.

На фиг. 5 представлена общая схема стратегии пост-синтетического сайт-специфического мечения на основе клик-химии. Модифицированные линкером нуклеотиды могут быть также непосредственно включены в матричную ДНК с использованием стандартного твердофазного синтеза олигонуклеотидов и соответствующих фосфорамидитов.

На фиг. 6 приведены репрезентативные данные, иллюстрирующие пост-амплификационное мечение ДНК, проанализированное по влиянию стрептавидина (SA) на электрофоретический сдвиг в геле. Быстрее мигрирующая полоса соответствует диДНК, в то время как медленнее мигрирующая полоса соответствует комплексу 1:1 между диДНК и стрептавидином. (A) Эффективность мечения составляет 72% с d5SICSPA-dNaM и 80% с dTPT3PA-dNaM. (B) Эффективность мечения составляет 6%, 84% и 56% с d5SICSPA-dNaM в первом (праймер 1: не встречающаяся в природе пара оснований в положении 1), девятом (праймер 2: не встречающаяся в природе пара оснований в положении одиннадцать) и одиннадцатом (праймер 3: не встречающаяся в природе пара оснований в позиции девять) положении. Соответствующие эффективности мечения с dTPT3PA-dNaM составляют 72%, 94% и 81%.

На фиг. 7 представлены данные гель-электрофореза, подтверждающие полноразмерную транскрипцию РНК, содержащей дериватизированные с помощью линкера аналоги 5SICS или MMO2.

Фиг. 8 иллюстрирует 12 групп α-аналогов нуклеиновых оснований, α1-α12.

Фиг. 9 иллюстрирует 6 групп β-аналогов нуклеиновых оснований, β1-β6.

Фиг. 10 иллюстрирует 2 группы α-аналогов нуклеиновых оснований, α13 и α14; где каждый X независимо представляет собой углерод или азот; где каждый R<sub>1</sub> независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азид-

ную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый R является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород.

Фиг. 11 иллюстрирует примеры  $\alpha$ 14a-аналогов нуклеиновых оснований, включая дериватизированные с помощью линкера аналоги нуклеиновых оснований, MMSPCO и MMSPA.

Фиг. 12 иллюстрирует 2 группы  $\beta$ -аналогов нуклеиновых оснований,  $\beta$ 7 и  $\beta$ 8; где каждый X независимо представляет собой углерод или азот; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород; и где R является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго.

Фиг. 13 иллюстрирует примеры дериватизированных с помощью линкера P8-аналогов нуклеиновых оснований.

На фиг. 14 показано процентное содержание не встречающихся в природе пар оснований, сохранившихся в ДНК после амплификации в течение 6 раундов скринингов.

Фиг. 15 иллюстрирует  $\alpha$ -аналоги нуклеиновых оснований,  $\alpha$ 15a,  $\alpha$ 15b,  $\alpha$ 16a и  $\alpha$ 16b; где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, азид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой кислород, серу или селен; и где каждый E2 независимо представляет собой серу или селен.

Фиг. 16 иллюстрирует  $\beta$ -аналоги нуклеиновых оснований,  $\beta$ 9a и  $\beta$ 9b; где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, азид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой кислород, серу или селен.

#### Подробное описание

Настоящее изобретение относится к аналогам не встречающихся в природе нуклеотидов, содержащим преимущественно гидрофобные аналоги нуклеотидных оснований, которые образуют не встречающиеся в природе пары оснований во время опосредованной ДНК-полимеразой репликации ДНК или опосредованной РНК-полимеразой транскрипции РНК. Таким образом не встречающиеся в природе нуклеотидные основания могут быть введены сайт-специфически в олигонуклеотиды (одно- или двухцепочечные ДНК или РНК), где они могут обеспечивать сайт-специфическое расщепление или могут предостав-

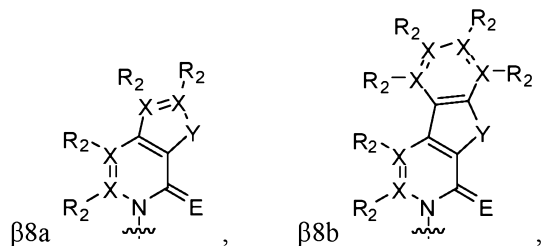
лять реакционноспособный линкер, который может подвергаться функционализации с карго-несущим реагентом посредством реакции с первичной аминогруппой или с помощью клик-химии с алкиновой группой линкера не встречающегося в природе нуклеотидного основания.

В различных вариантах осуществления линкеры присоединяются к карго на стадии нуклеозидтрифосфата, тем самым позволяя непосредственное получение нужного сайт-специфически меченого олигонуклеотида, например, с помощью автоматизированных устройств синтеза полинуклеотидов, таких как устройства синтеза полинуклеотидов фосфорамидитным способом.

В других вариантах осуществления линкеры включают реакционноспособный центр (например, первичный амин, алкин, тиол, альдегид или азид), обеспечивая реакционноспособный линкер, позволяющий выполнить сайт-специфическую модификацию ДНК- или РНК-олигонуклеотида после его синтеза. Это может быть достигнуто с помощью карго-реагента, содержащего карго (например, молекулу, липосому, наночастицу и т.д.) и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру реакционноспособной группы линкера. В некоторых вариантах осуществления реакционноспособный центр линкера защищен стандартной защитной группой. Реакция взаимодействия нуклеинового основания, раскрытого в настоящем изобретении, несущего реакционноспособный линкер (после снятия защиты, если это необходимо), и карго-реагента, включающего карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному линкеру, служит для обеспечения связывания нуклеинового основания с карго через сопряженную группу линкера.

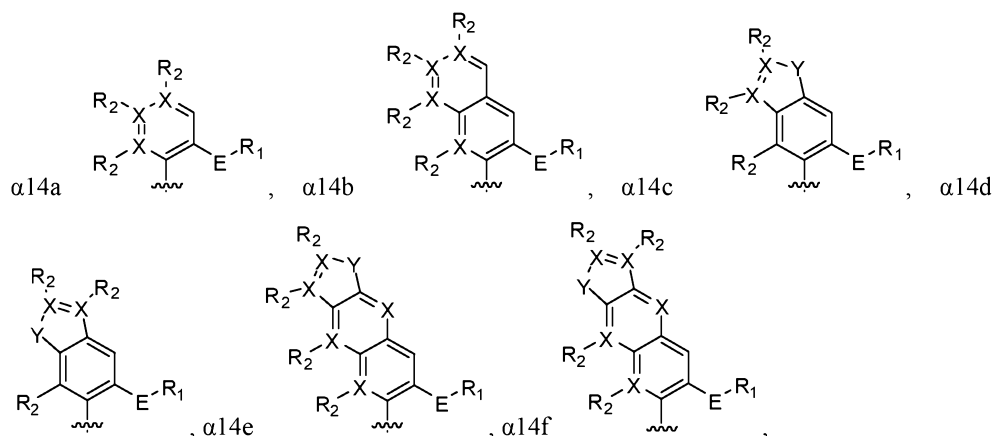
Композиции данного раскрытия в различных вариантах осуществления позволяют расширить ограниченный репертуар функциональных групп четырех природных ДНК-нуклеотидов и четырех природных РНК-нуклеотидов для того, чтобы включать практически любые функциональные группы, представляющие интерес, сайт-специфически встроенные в ДНК- или РНК-олигонуклеотид или в ДНК-аналог или РНК-аналог, такой как PNA или LNA или т.п. Карго необязательно включает функциональную группу, делающую возможным изменение молекулярного распознавания (т.е. для создания аптамеров), изменение реакционной способности (в том числе для катализа), и/или зонды для визуализации и/или характеристики (т.е. для развития диагностики).

В различных вариантах осуществления настоящего изобретения обеспечивается соединение, содержащее аналог нуклеинового основания любой из следующих формул  $\beta 8a$  или  $\beta 8b$ :



где каждый X независимо представляет собой углерод или азот; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород; и где аналог нуклеинового основания не является 4TFP или 7TFP или их линкерным производным.

В различных вариантах осуществления обеспечивается соединение, содержащее аналог нуклеинового основания любой из следующих формул:



где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород; и где аналог нуклеинового основания не является FIMO, MIMO, FEMO, PrMO, EMO, MEMO, IMO, MMO2, DMO, NMO, 5FM, 20Me, TMO, FDMO, VMO, ZMO, CIMO, TfMO, CNMO, NaM или QMO.

Здесь и далее волнистая линия указывает на точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку или к аналогу рибозильного, дезоксирибозильного или дидезоксирибозильного остатка, такому как "замкнутый" аналог рибозы (locked ribose analog), пептидная группа или т.п. В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный остаток или его аналог находится в свободной форме, присоединен к монофосфатной, дифосфатной или трифосфатной группе; необязательно содержащей  $\alpha$ -тиотрифосфатную,  $\beta$ -тиофосфатную или  $\gamma$ -тиофосфатную группу; или включен в РНК или ДНК или в РНК-аналог или ДНК-аналог.

В некоторых вариантах осуществления, когда речь идет либо о рибозильном остатке, либо о дезоксирибозильном остатке не встречающегося в природе нуклеинового основания, обеспеченного в настоящем изобретении, используется обозначение X, dX или (d)X, например, dTPT3 или (d)TPT3 относится к нуклеиновому основанию TPT3, связанному с любым из вариантов в положении, указанном волнистой линией. Таким образом, общее наименование dX относится к соединениям, имеющим рибозные или дезоксирибозные аналоги, присоединенные к ним, как обозначено волнистой линией. Когда конкретно речь идет о нуклеозиде с рибозильным остатком, приставка "d" отбрасывается, т.е. TPT3 относится к рибозильной форме. При включении нуклеотида или дериватизированных с помощью линкера вариантов в трифосфатный субстрат полимеразы (т.е. TPT3TP, dTPT3TP) считается, что они включаются в РНК- или ДНК-олигонуклеотид с использованием РНК- или ДНК-полимеразы соответственно.

В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный аналог обеспеченного нуклеозидного аналога (например,  $\beta 8a$ ,  $\beta 8b$ ,  $\alpha 14a$ ,  $\alpha 14b$ ,  $\alpha 14c$ ,  $\alpha 14d$ ,  $\alpha 14e$ ,  $\alpha 14f$ ) содержит 2'-функциональную группу. Примеры функциональных групп включают, без ограничения, метоксигруппу, галогеновую, -О-аллильную, -О-метоксиэтильную, первичный амин, -О-динитрофенольную, -О-динитрофенилэфирную, алкильную, -О-алкильную, тиольную, аминоэтоксиметильную, аминпропоксиметильную, аминоэтильную, цианоэтильную и гуанидиноэтильную группы. В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный аналог содержит 4'-тио-замену (например, атом кислорода остатка сахара заменяется атомом серы).

В некоторых вариантах осуществления алкильная группа аналога нуклеинового основания включает, без ограничения, метильную, этильную, пропильную и изопропильную группы. В некоторых вариантах осуществления галогеновая группа аналога нуклеинового основания включает, без ограничения, фтор, хлор, бром и йод.

В некоторых вариантах осуществления реакционноспособный линкер аналога нуклеинового основания содержит функциональную группу, включая, но не ограничиваясь ими, алкил, алкенил, алкинил, фенил, бензил, галоген, гидроксил, карбонил, альдегид, галоформил, сложный карбонатный эфир, карбоксилат, карбоксил, сложный эфир, метокси, гидроперокси, перокси, простой эфир, полуацеталь, полуацеталь, ортоэфир, метилendiокси, ортокарбонатный эфир, карбоксамид, первичный

амин, вторичный амин, имид, азид, азо, цианат, изоцианат, нитрат, нитрил, изонитрил, нитрозоокси, нитро, нитрозо, пиридил, сульфгидрил, сульфид, дисульфид, сульфинил, сульфо, тиоцианат, изотиоцианат, карбонотиоил, фосфино, фосфоно, фосфат, бороно, боронат, бороино, боринат и их комбинации. В некоторых вариантах осуществления реакционноспособный линкер аналога нуклеинового основания включает аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу.

В некоторых вариантах осуществления рибозильный или дезоксирибозильный остаток включает трифосфатную или  $\alpha$ -тиотрифосфатную группу, связанную с его 5'-гидроксильной группой. В некоторых вариантах осуществления рибозильный или дезоксирибозильный остаток включен в РНК- или ДНК-олигонуклеотидную цепь, соответственно, или рибозильный или дезоксирибозильный остаток или его аналог включен в РНК- или ДНК-аналог. В некоторых вариантах осуществления РНК-аналог или ДНК-аналог представляет собой пептидонуклеиновую кислоту (peptide nucleic acid (PNA)) или "замкнутую" нуклеиновую кислоту (locked nucleic acid (LNA)). В некоторых вариантах осуществления РНК-аналог или ДНК-аналог представляет собой бициклическое производное. Бициклические производные включают, без ограничения, нуклеиновую кислоту с 2'-O,4'-C-этиленовым мостиком (2'-O,4'-C-ethylene-bridged nucleic acid (ENA)), карбоциклическую "замкнутую" нуклеиновую кислоту (CLNA), циклогексен-нуклеиновую кислоту (cyclohexene nucleic acid (CENA)) и 2'-дезоксид-2'-N,4'-C-этилен-"замкнутую" нуклеиновую кислоту (2'-deoxy-2'-N,4'-C-ethylene-locked nucleic acid (AENA)). В некоторых вариантах осуществления РНК-аналог или ДНК-аналог представляет собой ациклическое производное. В некоторых вариантах осуществления РНК-аналог или ДНК-аналог представляет собой "разомкнутую" нуклеиновую кислоту (unlocked nucleic acid (UNA)). В некоторых вариантах осуществления РНК-аналог или ДНК-аналог содержит пиранозное кольцо вместо рибозы. В некоторых вариантах осуществления РНК-аналог или ДНК-аналог представляет собой арабино-нуклеиновую кислоту (arabino nucleic acid (ANA)) или гекситол-нуклеиновую кислоту (hexitol nucleic acid (HNA)).

В некоторых вариантах осуществления рибозильный или дезоксирибозильный остаток или его аналог замещен защитными и активирующими группами, подходящими для применения в автоматизированном устройстве для химического синтеза олигонуклеотидов. Примером автоматизированного устройства для химического синтеза олигонуклеотидов является устройство для фосфорамидитного синтеза.

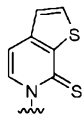
В некоторых вариантах осуществления по меньшей мере один  $R_2$  из аналога нуклеинового основания независимо содержит группу  $-C\equiv C-CH_2NHR_3$ , в которой  $R_3$  представляет собой водород или аминозащитную группу. Примером аминозащитной группы является дихлорацетильная группа. В некоторых вариантах осуществления по меньшей мере один  $R_2$  из аналога нуклеинового основания независимо содержит ацетиленовую группу, подходящую для применения в клик-реакции с карго-реагентом, содержащим карго и группу, способную реагировать с ацетиленовой группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  аналога нуклеинового основания независимо содержит тиольную группу, подходящую для применения в реакции с карго-реагентом, содержащим карго и группу, способную реагировать с тиольной группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  аналога нуклеинового основания независимо содержит альдегидную группу, подходящую для применения в реакции с карго-реагентом, содержащим карго и группу, способную реагировать с альдегидной группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  аналога нуклеинового основания независимо содержит азидную группу, подходящую для применения в реакции с карго-реагентом, содержащим карго и группу, способную реагировать с азидной группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  аналога нуклеинового основания независимо содержит  $-C\equiv C-(CH_2)_n-C\equiv CH$ , где  $n$  равно 1, 2, 3, 4, 5, или 6; или  $R_2$  представляет собой  $-C\equiv C-(CH_2)_{n1}-O-(CH_2)_{n2}-C\equiv CH$ , где  $n1$  и  $n2$ , каждый независимо друг от друга, равны 1, 2 или 3. В некоторых вариантах осуществления по меньшей мере один  $R_2$  независимо представляет собой сопряженный линкер, присоединенный к карго реакцией между аминогруппой и группой, способной реагировать с аминогруппой. Примером группы, способной реагировать с аминогруппой, является ацилирующая группа или алкилирующая группа, или сложный эфир N-гидроксисукцинимид. В некоторых вариантах осуществления по меньшей мере один  $R_2$  независимо представляет собой связанный линкер, присоединенный к карго реакцией между ацетиленовой группой и группой, способной реагировать с ацетиленовой группой. Примером группы, способной реагировать с ацетиленовой группой, является азидная группа. В некоторых вариантах осуществления ацетиленовая группа и азидная группа соединяются в медь-катализируемой клик-реакции. В некоторых вариантах осуществления по меньшей мере один  $R_2$  независимо представляет собой связанный линкер, присоединенный к карго реакцией между тиольной группой и группой, способной реагировать с тиольной группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  независимо представляет собой связанный линкер, присоединенный к карго реакцией между альдегидной группой и группой, способной реагировать с альдегидной группой. В некоторых вариантах осуществления по меньшей мере один  $R_2$  независимо представляет собой связанный линкер, присоединенный к карго реакцией между азидной группой и группой, способной реагировать с азидной группой. Примером группы, способной реагировать с азидной группой, является терминальный алкин или напряженный циклооктин. В некоторых вариантах осуществления по меньшей мере один  $R_2$  представляет собой независимо водород, соеди-



нение, содержащее  $\alpha$ -тиотрифосфатную группу, причем карго-реагент содержит  $\gamma$ -бром- $\alpha,\beta$ -ненасыщенную карбонильную группу, иод, бромацетил, или к нему присоединена азиридилиссульфонамидная группа.

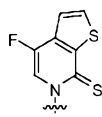
В некоторых вариантах осуществления карго аналога нуклеинового основания включает, без ограничения, белки, пептиды, аминокислоты, олигонуклеотиды, низкомолекулярные лекарственные препараты, алифатические группы, соединения, содержащие фотореакционноспособные группы, соединения, содержащие химически реакционноспособные группы, соединения, содержащие каталитические группы, соединения, содержащие хлорметилкетоны, липиды, биотин, флуоресцентные соединения, соединения-тушители флуоресценции, липосомы и наночастицы.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



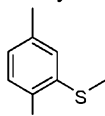
леинового основания TPT3 и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



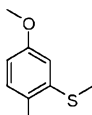
леинового основания FTPT3 и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



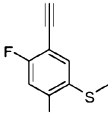
леинового основания MMS и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



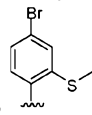
леинового основания DMS и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



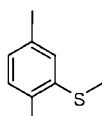
леинового основания FEMS и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



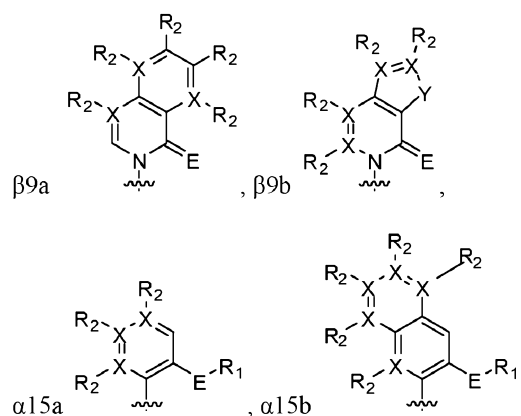
леинового основания BrMS и его производные и аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечивается аналог нук-



леинового основания IMS и его производные и аналоги.

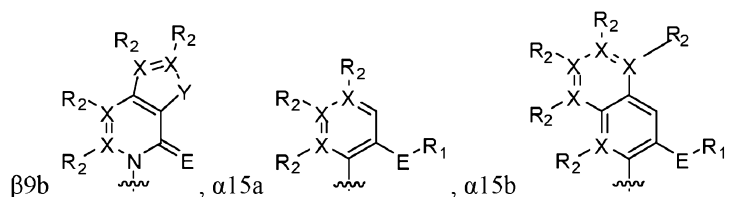
В некоторых вариантах осуществления настоящего изобретения обеспечиваются пары нуклеиновых оснований, содержащие аналог первого нуклеинового основания, имеющий любую из формул  $\beta 9a$  или  $\beta 9b$ ; и аналог второго нуклеинового основания, имеющий любую из формул  $\alpha 15a$  или  $\alpha 15b$



где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой кислород, серу или селен; и причем пара нуклеиновых оснований не является

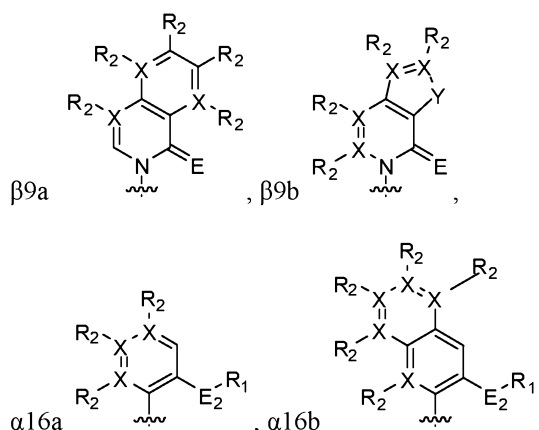
dICS-dMMO2, dICS-2OMe, dSICS-dMMO2, dSICS-d2OMe, dSNICS-dMMO2, dSNICS-d2OMe, d4SICS-dMMO2, d4SICS-d2OMe, d5SICS-dFIMO, d5SICS-dMIMO, d5SICS-dFEMO, d5SICS-dPrMO, d5SICS-dEMO, d5SICS-dMEMO, d5SICS-dIMO, d5SICS-dMMO2, d5SICS-dDMO, d5SICS-dNMO, d5SICS-d5FM, d5SICS-d2OMe, d5SICS-dTMO, d5SICS-dFDMO, d5SICS-dVMO, d5SICS-dZMO, d5SICS-dCIMO, d5SICS-dTfMO и d5SICS-dCNMO.

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий формулу  $\beta 9b$ , и аналог второго нуклеинового основания, имеющий любую из формул  $\alpha 15a$  или  $\alpha 15b$



где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, амид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; и где каждый E независимо представляет собой кислород, серу или селен.

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий любую из формул  $\beta 9a$  или  $\beta 9b$ , и аналог второго нуклеинового основания, имеющий любую из формул  $\alpha 16a$  или  $\alpha 16b$



где каждый X независимо представляет собой углерод или азот; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, азид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой кислород, серу или селен; и где каждый  $E_2$  независимо представляет собой серу или селен.

Волнистая линия указывает на точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку; или к аналогу рибозильного, дезоксирибозильного или дидезоксирибозильного остатка, такому как "замкнутый" аналог рибозы (locked ribose analog), пептидная группа или т.п. В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный остаток или его аналог находится в свободной форме, присоединен к монофосфатной, дифосфатной или трифосфатной группе; необязательно содержащей  $\alpha$ -тиотрифосфатную,  $\beta$ -тиофосфатную или  $\gamma$ -тиофосфатную группу; или включен в РНК или ДНК или в РНК-аналог или ДНК-аналог.

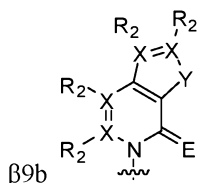
В некоторых вариантах осуществления алкинильная группа представляет собой этинильную или пропинильную группу. В некоторых вариантах осуществления алкильная группа представляет собой метильную, этильную, пропильную или изопропильную группу. В некоторых вариантах осуществления галоген представляет собой фтор, хлор, бром или йод.

В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный аналог пары нуклеиновых оснований содержит 2'-функциональную группу. Примеры функциональных групп включают, без ограничения, метоксигруппу, галогеновую, -О-аллильную, -О-метоксиэтильную, первичный амин, -О-динитрофенольную, -О-динитрофенилэфирную, алкильную, -О-алкильную, тиольную, аминоэтоксиметильную, аминопроксиметильную, аминоэтильную, цианоэтильную и гуанидиноэтильную группы. В некоторых вариантах осуществления рибозильный, дезоксирибозильный или дидезоксирибозильный аналог содержит 4'-тио-замену.

В некоторых вариантах осуществления пара нуклеиновых оснований содержит нуклеиновое основание, имеющее формулу  $\alpha 15b$ . В некоторых вариантах осуществления пара нуклеиновых оснований содержит нуклеиновое основание, имеющее формулу  $\alpha 15a$ . В некоторых вариантах осуществления пара нуклеиновых оснований содержит нуклеиновое основание, имеющее формулу  $\beta 9b$ . В некоторых вариантах осуществления пара нуклеиновых оснований содержит нуклеиновое основание, имеющее формулу  $\beta 9a$ , где каждый X представляет собой углерод, Y представляет собой серу, каждый  $R_2$  представляет собой водород и E представляет собой серу. В некоторых вариантах осуществления пара нуклеиновых оснований содержит нуклеиновое основание, имеющее формулу  $\alpha 15b$ , где каждый X представляет собой углерод, каждый  $R_2$  представляет собой водород,  $R_1$  представляет собой метильную группу и E представляет собой кислород. В некоторых вариантах осуществления пара нуклеиновых оснований содержит аналог первого нуклеинового основания, имеющий формулу  $\beta 9a$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16a$ . В некоторых вариантах осуществления пара нуклеиновых оснований содержит аналог первого нуклеинового основания, имеющий формулу  $\beta 9a$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16b$ . В некоторых вариантах осуществления пара нуклеиновых оснований содержит аналог первого нуклеинового основания, имеющий формулу  $\beta 9b$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16a$ . В некоторых вариантах осуществления пара нуклеино-

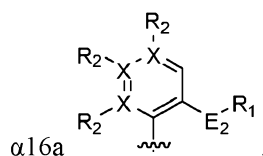
вых оснований содержит аналог первого нуклеинового основания, имеющий формулу  $\beta 9b$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16b$ .

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий формулу 9b, и аналог второго нуклеинового основания, имеющий формулу 9b



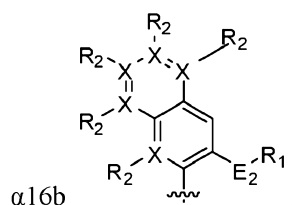
где каждый X независимо представляет собой углерод или азот; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, амид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; и где каждый E независимо представляет собой кислород, серу или селен. В некоторых вариантах осуществления пара нуклеиновых оснований представляет собой пару гомонуклеиновых оснований.

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий формулу  $\alpha 16a$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16a$



где каждый X независимо представляет собой углерод или азот; где каждый R<sub>1</sub> независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, амид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; и где каждый E<sub>2</sub> независимо представляет собой серу или селен. В некоторых вариантах осуществления пара нуклеиновых оснований представляет собой пару гомонуклеиновых оснований.

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий формулу  $\alpha 16b$ , и аналог второго нуклеинового основания, имеющий формулу  $\alpha 16b$



где каждый X независимо представляет собой углерод или азот; где каждый R<sub>1</sub> независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метокси, метантиол, метанселен, галоген, циано, азид, нитрогруппу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер,

к которому присоединен карго; и где каждый  $E_2$  независимо представляет собой серу или селен. В некоторых вариантах осуществления пара нуклеиновых оснований представляет собой пару гомонуклеиновых оснований.

В некоторых вариантах осуществления настоящего изобретения обеспечивается пара нуклеиновых оснований, содержащая аналог первого нуклеинового основания, имеющий любую из формул  $\beta 9a$ ,  $\beta 9b$ ,  $\alpha 15a$ ,  $\alpha 15b$ ,  $\alpha 16a$  или  $\alpha 16b$ ; и второе нуклеиновое основание, выбранное из группы, состоящей из цитозина, гуанина, аденина, тимина, урацила, 2-аминоаденин-9-ила, 2-аминоаденина, 2-Ф-аденина, 2-тиоурацила, 2-тиотимина, 2-тиоцитозина, 2-пропил- и алкильных производных аденина и гуанина, 2-аминоаденина, 2-аминопропил-аденина, 2-аминопиридина, 2-пиридола, 2'-дезоксидуридина, 2-амино-2'-дезоксиаденозина, 3-деазагуанина, 3-дезааденина, 4-тиоурацила, 4-тиотимина, урацил-5-ила, гипоксантин-9-ила (I), 5-метил-цитозина, 5-гидроксиметил-цитозина, ксантина, гипоксантина, 5-бром- и 5-трифторметилурацилов и цитозинов; 5-галоурацила, 5-галоцитозина, 5-пропинил-урацила, 5-пропинил-цитозина, 5-урацила, 5-замещенных, 5-галоген-, 5-замещенных пиримидинов, 5-гидроксицитозина, 5-бромцитозина, 5-бромуррацила, 5-хлорцитозина, хлорированного цитозина, циклоцитозина, цитозина-рабинозида, 5-фторцитозина, фторпиримидина, фторурацила, 5,6-дигидроцитозина, 5-иодоцитозина, гидроксимочевина, иодоурацила, 5-нитроцитозина, 5-бромуррацила, 5-хлорурацила, 5-фторурацила и 5-иодоурацила, 6-алкильных производных аденина и гуанина, 6-азапиримидинов, 6-азоурацила, 6-азоцитозина, азацитозина, 6-азотимина, 6-тиогуанина, 7-метилгуанина, 7-метиладенина, 7-деазагуанина, 7-деазагуанозина, 7-дезааденина, 7-деза-8-азагуанина, 8-азагуанина, 8-азааденина, 8-галоген-, 8-амино-, 8-тиол-, 8-тиоалкил- и 8-гидроксилзамещенных аденинов и гуанинов; N4-этилцитозина, N-2 замещенных пуринов, N-6 замещенных пуринов, O-6 замещенных пуринов тех соединений, которые повышают стабильность образования дуплекса, универсальных нуклеиновых кислот, гидрофобных нуклеиновых кислот, смешанных (promiscuous) нуклеиновых кислот, size-expanded нуклеиновых кислот, фторированных нуклеиновых кислот, трициклических пиримидинов, феноксазинцитидина ([5,4-b][1,4]бензоксазин-2(3H)-она), фенотиазинцитидина (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-она), "G-зажимов" (G-clamps), феноксазинцитидина (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-она), карбазолцитидина (2H-пиримидо[4,5-b]индол-2-она), пиридоиндолцитидина (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-она), 5-фторурацила, 5-бромуррацила, 5-хлорурацила, 5-иодоурацила, гипоксантина, ксантина, 4-ацетилцитозина, 5-(карбоксигидроксиметил)урацила, 5-карбоксиметиламинометил-2-тиоуридина, 5-карбоксиметиламинометилурацила, дигидроурацила,  $\beta$ -D-галактозилквеозина, инозина, N6-изопентиладенина, 1-метилгуанина, 1-метилюнозина, 2,2-диметилгуанина, 2-метиладенина, 2-метилгуанина, 3-метилцитозина, 5-метилцитозина, N6-аденина, 7-метилгуанина, 5-метиламинометилурацила, 5-метоксиаминометил-2-тиоурацила,  $\beta$ -D-маннозилквеозина, 5'-метоксикарбоксиметилурацила, 5-метоксиурацила, 2-метилтио-N6-изопентиладенина, урацил-5-оксиуксусной кислоты, wybutoxosine, псевдоурацила, квеозина, 2-тиоцитозина, 5-метил-2-тиоурацила, 2-тиоурацила, 4-тиоурацила, 5-метилурацила, урацил-5-метилового эфира оксиуксусной кислоты, урацил-5-оксиуксусной кислоты, 5-метил-2-тиоурацила, 3-(3-амино-3-N-2-карбоксипропил)урацила, (аср3)w и 2,6-диаминопурина и тех соединений, в которых пуриновое или пиримидиновое основание заменено гетероциклом.

Примером пары оснований, содержащей одно или более не встречающихся в природе нуклеиновых оснований, может служить пара не встречающихся в природе оснований dTPT3PA-dNaM (т.е. пара, образованная между dTPT3PA и dNaM; фиг. 1A и 1B). Кроме того, ортогональная реакционная способность различных разработанных реакционноспособных центров/линкеров (т.е. форсфотиоаты, амины и алкины) позволяет реализовывать селективное упорядоченное взаимодействие различных остатков с одним и тем же олигонуклеотидом (ДНК или РНК). Другая композиция дополнительно иллюстрируется парой не встречающихся в природе оснований dTPT3PA-dMMO2pCO, где в различных вариантах осуществления алкильная группа основания dMMO2pCO используется для присоединения одной функциональной группы через катализируемое медью (I) азид-алкиновое циклоприсоединение (Copper(I)-catalyzed Azide-Alkyne Cycloaddition (CuAAC)) и после снятия защиты с dTPT3PA свободный амин используется для присоединения другой функциональной группы через связывание N-гидроксисукцинимидом (NHS).

В то время как несколько других ранее описанных пар не встречающихся в природе оснований были дериватизированы с помощью линкеров, описанные здесь дериватизированные с помощью линкера пары, dTPT3PA-dNaM, d5SICSCO-dNaM, dTPT3PA-dMMO2pCO (фиг. 2), являются более проверенными, а также лучше реплицируются в ДНК и лучше транскрибируются в РНК. В частности, dTPT3PA-dNaM хорошо реплицируется и, таким образом, подходит для применения в практическом осуществлении способов, раскрытых и заявленных в настоящем изобретении.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dTPT3, причем dTPT3 в некоторых случаях дериватизируется линкером. Не встречающиеся в природе пары оснований, содержащие dTPT3 или дериватизированное с помощью линкера dTPT3 (например, dTPT3PA), включают, без ограничения,

dTPT3-MMS,

dTPT3-DMS, dTPT3-FEMS, dTPT3-BrMS, dTPT3-IMS, dTPT3-dDMN, dTPT3-d4OMe, dTPT3-dIQ, dTPT3-d2MN, dTPT3-d3OMe, dTPT3-dQL, dTPT3-d2Np, dTPT3-dDM4, dTPT3-dDM, dTPT3-dBEN, dTPT3-d3FB, dTPT3-dMM1, dTPT3-dMMO1, dTPT3-dDM2, dTPT3-dDM5, dTPT3-d2Py, dTPT3-d5MPy, dTPT3-dEPy, dTPT3-d3MPy, dTPT3-d34DMPy, dTPT3-d45DMPy, dTPT3-d4MPy, dTPT3-d35DMPy, dTPT3-dBP, dTPT3-dBTp, dTPT3-dBF, dTPT3-dIN, dTPT3-dTp, dTPT3-dBTz, dTPT3-dMTp, dTPT3-dAM, dTPT3-dMAN, dTPT3-dMMAN, dTPT3-dADM, dTPT3-dMMAN, dTPT3-dTOK588, dTPT3-dTOK576, dTPT3-dTOK587, dTPT3-dTOK586, dTPT3-dTOK580, dTPT3-dPhMO, dTPT3-dPyMO1, dTPT3-PyMO2, dTPT3-dPMO1, dTPT3-dPMO2, dTPT3-dPMO3, dTPT3-dFuMO1, dTPT3-dFuMO2, dTPT3-TpMO1, dTPT3-dTpMO2, dTPT3-dFIMO, dTPT3-dIMO, dTPT3-dMIMO, dTPT3-dMEMO, dTPT3-dFEMO, dTPT3-dPrMO, dTPT3-dMMO2, dTPT3-d2OMe, dTPT3-dDMO, dTPT3-dTMO, dTPT3-dNMO, dTPT3-dNOPy, dTPT3-d5FM, dTPT3-dNAM, dTPT3-dAMO1, dTPT3-dAPy, dTPT3-dAMO2, dTPT3-dMAPy, dTPT3-dAMO3, dTPT3-dDMAPy, dTPT3-dFDMO, dTPT3-dVMO, dTPT3-dQMO, dTPT3-dZMO, dTPT3-dCIMO, dTPT3-dTfMO и dTPT3-CNMO,

причем основание, комплементарное dTPT3, дериватизировано или не дериватизировано с помощью линкера (например, dMMO2pCO). dTPT3 проиллюстрировано на фиг. 9 в качестве β6-аналога. Пример dTPT3, дериватизированного с помощью линкера, продемонстрирован на фиг. 2, причем в некоторых случаях R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dTPT3, включают, без ограничения, α-аналоги или дериватизированные с помощью линкера α-аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления dTPT3 или dTPT3, дериватизированное с помощью линкера, является основанием, спаренным с dTPT3 или dTPT3, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dTPT3 или dTPT3, дериватизированное с помощью линкера, является основанием, спаренным с β-нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой, например dTPT3PA. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, например dTPT3A, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dMMS, где dMMS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dMMS или dMMS, дериватизированное с помощью линкера (например, dMMSPA), включают, без ограничения,

d7AI-dMMS, dM7AI-dMMS, dImPy-dMMS, dP7AI-dMMS, dPPP-dMMS, d8Q-dMMS, dICS-dMMS, dPICS-dMMS, dMICS-dMMS, d4MICS-dMMS, d5MICS-dMMS, dNICS-dMMS, dONICS-dMMS, d7OFP-dMMS, d7OTP-dMMS, d4OTP-dMMS, dPYR-dMMS, d4MP-dMMS, d3MP-dMMS, dPPYR-dMMS, dMOP-dMMS, d4MOP-dMMS, dSICS-dMMS, dSNICS-dMMS, d5SICS-dMMS, d4SICS-dMMS, dTPT1-dMMS, dTPT2-dMMS, dFPT1-dMMS и dFTPT3-dMMS,

причем основание, комплементарное dMMS, дериватизировано или не дериватизировано с помощью линкера (например, pFTPT3pA). dMMS проиллюстрировано на фиг. 11 в качестве α14a-аналога. В некоторых вариантах осуществления dMMS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dMMS, включают, без ограничения, β-аналоги или β-аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 9, 12 и 13. В некоторых вариантах осуществления dMMS или dMMS, дериватизированное с помощью линкера, является основанием, спаренным с dMMS или dMMS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых

оснований. В некоторых вариантах осуществления dMMS или dMMS, дериватизированное с помощью линкера, является основанием, спаренным с  $\alpha$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 8, 10, 11 и 15, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dDMS, причем dDMS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dDMS или dDMS, дериватизированное с помощью линкера (например, dDMSPA), включают, без ограничения,

d7AI-dDMS, dM7AI-dDMS, dImPy-dDMS, dP7AI-dDMS, dPPP-dDMS, d8Q-dDMS, dICS-dDMS, dPICS-dDMS, dMICS-dDMS, d4MICS-dDMS, d5MICS-dDMS, dNICS-dDMS, dONICS-dDMS, d7OFP-dDMS, d7OTP-dDMS, d4OTP-dDMS, dPYR-dDMS, d4MP-dDMS, d3MP-dDMS, dPPYR-dDMS, dMOP-dDMS, d4MOP-dDMS, dSICS-dDMS, dSNICS-dDMS, d5SICS-dDMS, d4SICS-dDMS, dTPT1-dDMS, dTPT2-dDMS, dFPT1-dDMS, dFTPT3-dDMS,

причем основание, комплементарное dDMS, дериватизировано или не дериватизировано с помощью линкера (например, rFTPT3pA). dDMS проиллюстрировано на фиг. 11 в качестве  $\alpha$ 14a-аналога. В некоторых вариантах осуществления dDMS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dDMS, включают, без ограничения,  $\beta$ -аналоги или  $\beta$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 9, 12 и 13. В некоторых вариантах осуществления dDMS или dDMS, дериватизированное с помощью линкера, является основанием, спаренным с dDMS или dDMS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dMMS или dMMS, дериватизированное с помощью линкера, является основанием, спаренным с  $\alpha$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 8, 10, 11 и 15, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dFEMS, причем dFEMS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dFEMS или dFEMS, дериватизированное с помощью линкера (например, dFEMSPA), включают, без ограничения,

d7AI-dFEMS, dM7AI-dFEMS, dImPy-dFEMS, dP7AI-dFEMS, dPPP-dFEMS, d8Q-dFEMS, dICS-dFEMS, dPICS-dFEMS, dMICS-dFEMS, d4MICS-dFEMS, d5MICS-dFEMS, dNICS-dFEMS, dONICS-dFEMS, d7OFP-dFEMS, d7OTP-dFEMS, d4OTP-dFEMS, dPYR-dFEMS, d4MP-dFEMS, d3MP-dFEMS, dPPYR-dFEMS, dMOP-dFEMS, d4MOP-dFEMS, dSICS-dFEMS, dSNICS-dFEMS, d5SICS-dFEMS, d4SICS-dFEMS, dTPT1-dFEMS, dTPT2-dFEMS, dFPT1-dFEMS, dFTPT3-dFEMS,

причем основание, комплементарное dFEMS, дериватизировано или не дериватизировано с помощью линкера (например, rFTPT3pA). dFEMS проиллюстрировано на фиг. 11 в качестве  $\alpha$ 14a-аналога. В некоторых вариантах осуществления dFEMS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые комплементарны dFEMS, включают, без ограничения,  $\beta$ -аналоги или  $\beta$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 9, 12 и 13. В некоторых вариантах осуществления dFEMS или dFEMS, дериватизированное с помощью линкера, является основанием, спаренным с dFEMS или dFEMS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dFEMS или dFEMS, дериватизированное с помощью линкера, является основанием, спаренным с  $\alpha$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 8, 10, 11 и 15, или производное данного

нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dBrMS, причем dBrMS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dBrMS или dBrMS, дериватизированное с помощью линкера (например, dBrMSPA), включают, без ограничения,

d7AI-dBrMS, dM7AI-dBrMS, dImPy-dBrMS, dP7AI-dBrMS, dPPP-dBrMS, d8Q-dBrMS, dICS-dBrMS, dPICS-dBrMS, dMICS-dBrMS, d4MICS-dBrMS, d5MICS-dBrMS, dNICS-dBrMS, dONICS-dBrMS, d7OFP-dBrMS, d7OTP-dBrMS, d4OTP-dBrMS, dPYR-dBrMS, d4MP-dBrMS, d3MP-dBrMS, dPPYR-dBrMS, dMOP-dBrMS, d4MOP-dBrMS, dSICS-dBrMS, dSNICS-dBrMS, d5SICS-dBrMS, d4SICS-dBrMS, dTPT1-dBrMS, dTPT2-dBrMS, dFPT1-dBrMS,

dFTPT3-dBrMS, причем основание, комплементарное dBrMS, дериватизировано или не дериватизировано с помощью линкера (например, pFTPT3pA). dBrMS проиллюстрировано на фиг. 11 в качестве  $\alpha$ 14a-аналога. В некоторых вариантах осуществления dBrMS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dBrMS, включают, без ограничения,  $\beta$ -аналоги или  $\beta$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 9, 12 и 13. В некоторых вариантах dBrMS или dBrMS, дериватизированное с помощью линкера, является основанием, спаренным с dBrMS или dBrMS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dBrMS или dBrMS, дериватизированное с помощью линкера, является основанием, спаренным с  $\alpha$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 8, 10, 11 и 15, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, причем dIMS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dIMS или dIMS, дериватизированное с помощью линкера (например, dIMSPA), включают, без ограничения,

d7AI-dIMS, dM7AI-dIMS, dImPy-dIMS, dP7AI-dIMS, dPPP-dIMS, d8Q-dIMS, dICS-dIMS, dPICS-dIMS, dMICS-dIMS, d4MICS-dIMS, d5MICS-dIMS, dNICS-dIMS, dONICS-dIMS, d7OFP-dIMS, d7OTP-dIMS, d4OTP-dIMS, dPYR-dIMS, d4MP-dIMS, d3MP-dIMS, dPPYR-dIMS, dMOP-dIMS, d4MOP-dIMS, dSICS-dIMS, dSNICS-dIMS, d5SICS-dIMS, d4SICS-dIMS, dTPT1-dIMS, dTPT2-dIMS, dFPT1-dIMS, dFTPT3-dIMS,

причем основание, комплементарное dIMS, дериватизировано или не дериватизировано с помощью линкера (например, pFTPT3pA). dIMS проиллюстрировано на фиг. 11 в качестве аналога  $\alpha$ 14a. В некоторых вариантах осуществления dIMS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dIMS, включают, без ограничения,  $\beta$ -аналоги или  $\beta$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 9, 12 и 13. В некоторых вариантах осуществления dIMS или dIMS, дериватизированное с помощью линкера, является основанием, спаренным с dIMS или dIMS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dIMS или dIMS, дериватизированное с помощью линкера, является основанием, спаренным с  $\alpha$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 8, 10, 11 и 15, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dICS, причем dICS в некоторых случаях является дериватизированным с помощью линкера. Не встре-



чающиеся в природе пары оснований, содержащие dICS или dICS, дериватизированное с помощью линкера (например, dICSPA), включают, без ограничения,

dICS-dFIMO, dICS-dMIMO, dICS-dFEMO, dICS-dPrMO, dICS-dEMO, dICS-dMEMO, dICS-dIMO, dICS-dDMO, dICS-dNMO, dICS-d5FM, dICS-dTMO, dICS-dFDMO, dICS-dVMO, dICS-dZMO, dICS-dCIMO, dICS-dTfMO, dICS-dCNMO, dICS-dNAM, dICS-dQMO,

причем основание, комплементарное dICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dICS проиллюстрировано на фиг. 9 в виде  $\beta$ 2-аналога. В некоторых вариантах осуществления dICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dICS, включают, без ограничения,  $\alpha$ -аналоги или  $\alpha$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления dICS или dICS, дериватизированное с помощью линкера, является основанием, спаренным с dICS или dICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dICS или dICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dPICS, причем dPICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dPICS или dPICS, дериватизированное с помощью линкера (например, dPICSPA), включают, без ограничения,

dPICS-dFIMO, dPICS-dMIMO, dPICS-dFEMO, dPICS-dPrMO, dPICS-dEMO, dPICS-dMEMO, dPICS-dIMO, dPICS-dMMO2, dPICS-dDMO, dPICS-dNMO, dPICS-d5FM, dPICS-d2OMe, dPICS-dTMO, dPICS-dFDMO, dPICS-dVMO, dPICS-dZMO, dPICS-dCIMO, dPICS-dTfMO, dPICS-dCNMO, dPICS-dNAM, dPICS-dQMO,

причем основание, комплементарное dPICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dPICS проиллюстрировано на фиг. 9 в виде  $\beta$ 2-аналога. В некоторых вариантах осуществления dPICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dPICS, включают, без ограничения,  $\alpha$ -аналоги или дериватизированные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах dPICS или dPICS, дериватизированное с помощью линкера, является основанием, спаренным с dPICS или dPICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dPICS или dPICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dMICS, в котором dMICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dMICS или dMICS, дериватизированное с помощью линкера (например, dMICSPA), включают, без ограничения,

dMICS-dFIMO, dMICS-dMIMO, dMICS-dFEMO, dMICS-dPrMO, dMICS-dEMO, dMICS-dMEMO, dMICS-dIMO, dMICS-dMMO2, dMICS-dDMO, dMICS-dNMO, dMICS-d5FM, dMICS-d2OMe, dMICS-dTMO, dMICS-dFDMO, dMICS-dVMO, dMICS-dZMO, dMICS-dCIMO, dMICS-dTfMO, dMICS-dCNMO, dMICS-dNAM, dMICS-dQMO,

причем основание, комплементарное dMICS, дериватизировано или не дериватизировано с помо-

шью линкера (например, dDMOpCO, dDMOpCC). dMICS проиллюстрировано на фиг. 9 в виде  $\beta$ 2-аналога. В некоторых вариантах осуществления dMICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dMICS, включают, без ограничения,  $\alpha$ -аналоги или дериватизированные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления dMICS или dMICS, дериватизированное с помощью линкера, является основанием, спаренным с dMICS или dMICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dMICS или dMICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим d4MICS, причем d4MICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие d4MICS или d4MICS, дериватизированное с помощью линкера (например, d4MICS-PA), включают, без ограничения,

d4MICS-dFIMO, d4MICS-dMIMO, d4MICS-dFEMO, d4MICS-dPrMO,  
d4MICS-dEMO, d4MICS-dMEMO, d4MICS-dIMO, d4MICS-dMMO2, d4MICS-dDMO,  
d4MICS-dNMO, d4MICS-d5FM, d4MICS-d2OMe, d4MICS-dTMO, d4MICS-dFDMO,  
d4MICS-dVMO, d4MICS-dZMO, d4MICS-dCIMO, d4MICS-dTfMO, d4MICS-dCNMO,  
d4MICS-dNAM, d4MICS-dQMO,

причем основание, комплементарное d4MICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). d4MICS проиллюстрировано на фиг. 9 в виде  $\beta$ 2-аналога. В некоторых вариантах осуществления d4MICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными d4MICS, включают, без ограничения,  $\alpha$ -аналоги или дериватизированные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах d4MICS или d4MICS, дериватизированное с помощью линкера, является основанием, спаренным с d4MICS или d4MICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления d4MICS или d4MICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим d5MICS, причем d5MICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие d5MICS или d5MICS, дериватизированное с помощью линкера (например, d5MICS-PA), включают, без ограничения,

d5MICS-dFIMO, d5MICS-dMIMO, d5MICS-dFEMO, d5MICS-dPrMO,  
d5MICS-dEMO, d5MICS-dMEMO, d5MICS-dIMO, d5MICS-dMMO2, d5MICS-dDMO,  
d5MICS-dNMO, d5MICS-d5FM, d5MICS-d2OMe, d5MICS-dTMO, d5MICS-dFDMO,  
d5MICS-dVMO, d5MICS-dZMO, d5MICS-dCIMO, d5MICS-dTfMO, d5MICS-dCNMO,  
d5MICS-dNAM, d5MICS-dQMO,

причем основание, комплементарное d5MICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). d5MICS проиллюстрировано на фиг. 9 в виде  $\beta$ 2-аналога. В некоторых вариантах осуществления d5MICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными d5MICS, включают, без ограничения,  $\alpha$ -аналоги или дериватизи-

рованные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах d5MICS или d5MICS, дериватизированное с помощью линкера, является основанием, спаренным с d5MICS или d5MICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления d5MICS или d5MICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dNICS, причем dNICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dNICS или dNICS, дериватизированное с помощью линкера (например, dNICS-PA), включают, без ограничения,

dNICS-dFIMO, dNICS-dMIMO, dNICS-dFEMO, dNICS-dPrMO, dNICS-dEMO,  
dNICS-dMEMO, dNICS-dIMO, dNICS-dDMO, dNICS-dNMO, dNICS-d5FM, dNICS-dTMO,  
dNICS-dFDMO, dNICS-dVMO, dNICS-dZMO, dNICS-dCIMO, dNICS-MMO2, dNICS-2OMe,  
dNICS-dTfMO, dNICS-dCNMO, dNICS-dNAM, dNICS-dQMO,

причем основание, комплементарное dNICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dNICS проиллюстрировано на фиг. 9 в виде  $\beta$ -аналога. В некоторых вариантах осуществления dNICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dNICS, включают, без ограничения,  $\alpha$ -аналоги или  $\alpha$ -аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах dNICS или dNICS, дериватизированное с помощью линкера, является основанием, спаренным с dNICS или dNICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dNICS или dNICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dONICS, причем dONICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dONICS или dONICS, дериватизированное с помощью линкера (например, dONICS-PA), включают, без ограничения,

dONICS-dFIMO, dONICS-dMIMO, dONICS-dFEMO, dONICS-dPrMO,  
dONICS-dEMO, dONICS-dMEMO, dONICS-dIMO, dONICS-dDMO, dONICS-dNMO,  
dONICS-d5FM, dONICS-dTMO, dONICS-dFDMO, dONICS-dVMO, dONICS-dZMO,  
dONICS-dCIMO, dONICS-MMO2, dONICS-2OMe, dONICS-dTfMO, dONICS-dCNMO,  
dONICS-dNAM, dONICS-dQMO,

причем основание, комплементарное dONICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dONICS проиллюстрировано на фиг. 9 в виде  $\beta$ -аналога. В некоторых вариантах осуществления dONICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые комплементарны dONICS, включают, без ограничения,  $\alpha$ -аналоги или дериватизированные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления dONICS или dONICS, дериватизированное с помощью линкера, является основанием, спаренным с dONICS или dONICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dONICS или dONICS, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена за-

щитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dSICS, причем dSICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dSICS или dSICS, дериватизированное с помощью линкера (например, dSICSPA), включают, без ограничения,

dSICS-dFIMO, dSICS-dMIMO, dSICS-dFEMO, dSICS-dPrMO, dSICS-dEMO,  
dSICS-dMEMO, dSICS-dIMO, dSICS-dDMO, dSICS-dNMO, dSICS-d5FM, dSICS-dTMO,  
dSICS-dFDMO, dSICS-dVMO, dSICS-dZMO, dSICS-dCIMO, dSICS-dTfMO, dSICS-dCNMO,  
dSICS-dNAM, dSICS-dQMO,

причем основание, комплементарное dSICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dSICS проиллюстрировано на фиг. 9 в виде β5-аналога. В некоторых вариантах осуществления dSICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dSICS, включают, без ограничения, α-аналоги или α-аналоги, дериватизированные с помощью линкера, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах dSICS или dSICS, дериватизированное с помощью линкера, является основанием, спаренным с dSICS или dSICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dSICS или dSICS, дериватизированное с помощью линкера, является основанием, спаренным с β-нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим dSNICS, причем dSNICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие dSNICS или dSNICS, дериватизированное с помощью линкера (например, dSNICSPA), включают, без ограничения,

dSNICS-dFIMO, dSNICS-dMIMO, dSNICS-dFEMO, dSNICS-dPrMO,  
dSNICS-dEMO, dSNICS-dMEMO, dSNICS-dIMO, dSNICS-dDMO, dSNICS-dNMO, dSNICS-  
d5FM, dSNICS-dTMO, dSNICS-dFDMO, dSNICS-dVMO, dSNICS-dZMO, dSNICS-dCIMO,  
dSNICS-dTfMO, dSNICS-dCNMO, dSNICS-dNAM, dSNICS-dQMO,

причем основание, комплементарное dSNICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). dSNICS проиллюстрировано на фиг. 9 в виде β5-аналога. В некоторых вариантах осуществления дериватизированное с помощью линкера dSNICS содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными dSNICS, включают, без ограничения, α-аналоги или дериватизированные с помощью линкера α-аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах dSNICS или dSNICS, дериватизированное с помощью линкера, является основанием, спаренным с dSNICS или dSNICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления dSNICS или dSNICS, дериватизированное с помощью линкера, является основанием, спаренным с β-нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим d4SICS, причем d4SICS в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие d4SICS или d4SICS, дериватизированное с помощью линкера (например, d4SICSPA), включают, без ограничения,

d4SICS-dFIMO, d4SICS-dMIMO, d4SICS-dFEMO, d4SICS-dPrMO, d4SICS-dEMO, d4SICS-dMEMO, d4SICS-dIMO, d4SICS-dDMO, d4SICS-dNMO, d4SICS-d5FM, d4SICS-dTMO, d4SICS-dFDMO, d4SICS-dVMO, d4SICS-dZMO, d4SICS-dCIMO, d4SICS-dTfMO, d4SICS-dCNMO, d4SICS-dNAM, d4SICS-dQMO,

причем основание, комплементарное d4SICS, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). d4SICS проиллюстрировано на фиг. 9 в виде β5-аналога. В некоторых вариантах осуществления d4SICS, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые являются комплементарными d4SICS, включают, без ограничения, α-аналоги или дериватизированные с помощью линкера α-аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах d4SICS или d4SICS, дериватизированное с помощью линкера, является основанием, спаренным с d4SICS или d4SICS, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления d4SICS или d4SICS, дериватизированное с помощью линкера, является основанием, спаренным с β-нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим d7OFP, причем d7OFP в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие d7OFP или d7OFP, дериватизированное с помощью линкера (например, d7OFPPA), включают, без ограничения,

d7OFP-dFIMO, d7OFP-dMIMO, d7OFP-dFEMO, d7OFP-dPrMO, d7OFP-dEMO, d7OFP-dMEMO, d7OFP-dIMO, d7OFP-dMMO2, d7OFP-dDMO, d7OFP-dNMO, d7OFP-d5FM, d7OFP-d2OMe, d7OFP-dTMO, d7OFP-dFDMO, d7OFP-dVMO, d7OFP-dZMO, d7OFP-dCIMO, d7OFP-dTfMO, d7OFP-dCNMO, d7OFP-dNAM, d7OFP-dQMO,

причем основание, комплементарное d7OFP, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). d7OFP проиллюстрировано на фиг. 9 в виде β5-аналога. В некоторых вариантах осуществления d7OFP, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые комплементарны d7OFP, включают, без ограничения, α-аналоги или дериватизированные с помощью линкера α-аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления d7OFP или d7OFP, дериватизированное с помощью линкера, является основанием, спаренным с d7OFP или d7OFP, дериватизированным с помощью линкера, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления d7OFP или d7OFP, дериватизированное с помощью линкера, является основанием, спаренным с β-нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

Настоящее изобретение относится к не встречающимся в природе парам оснований, содержащим d7OTP, причем d7OTP в некоторых случаях является дериватизированным с помощью линкера. Не встречающиеся в природе пары оснований, содержащие d7OTP или d7OTP, дериватизированное с помощью линкера (например, d7OTPPA), включают, без ограничения,

d7OTP-dFIMO, d7OTP-dMIMO, d7OTP-dFEMO, d7OTP-dPrMO, d7OTP-dEMO, d7OTP-dMEMO, d7OTP-dIMO, d7OTP-dMMO2, d7OTP-dDMO, d7OTP-dNMO, d7OTP-d5FM, d7OTP-d2OMe, d7OTP-dTMO, d7OTP-dFDMO, d7OTP-dVMO, d7OTP-dZMO, d7OTP-dCIMO, d7OTP-dTfMO, d7OTP-dCNMO, d7OTP-dNAM, d7OTP-dQMO,

причем основание, комплементарное d7OTP, дериватизировано или не дериватизировано с помощью линкера (например, dDMOpCO, dDMOpCC). d7OTP проиллюстрировано на фиг. 9 в виде β5-аналога. В некоторых вариантах осуществления d7OTP, дериватизированное с помощью линкера, содержит функциональную группу R, причем R представляет собой реакционноспособный линкер, содержа-

щий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, или R представляет собой сопряженный линкер, к которому присоединен карго. Аналоги нуклеиновых оснований, которые комплементарны d7OTP, включают, без ограничения,  $\alpha$ -аналоги или дериватизированные с помощью линкера  $\alpha$ -аналоги, проиллюстрированные на фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления d7OTP или d7OTR, дериватизированное с помощью линкера, является основанием, спаренным с d7OTP или дериватизированным с помощью линкера d7OTR, с образованием пары гомонуклеиновых оснований. В некоторых вариантах осуществления d7OTP или d7OTR, дериватизированное с помощью линкера, является основанием, спаренным с  $\beta$ -нуклеотидным основанием, включая, но не ограничиваясь ими, любое нуклеиновое основание, проиллюстрированное на фиг. 9, 12 и 13, или производное данного нуклеинового основания. В некоторых вариантах осуществления линкерная группа защищена защитной группой. В некоторых вариантах осуществления линкерная группа не защищена защитной группой, причем в некоторых случаях защитная группа удалена.

В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание TPT3 и второе нуклеиновое основание MMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание TPT3 и второе нуклеиновое основание DMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание TPT3 и второе нуклеиновое основание FEMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание TPT3 и второе нуклеиновое основание BrMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание TPT3 и второе нуклеиновое основание IMS.

В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание FTPT3 и второе нуклеиновое основание MMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание FTPT3 и второе нуклеиновое основание DMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание FTPT3 и второе нуклеиновое основание FEMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание FTPT3 и второе нуклеиновое основание BrMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание FTPT3 и второе нуклеиновое основание IMS.

В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание 5SICS и второе нуклеиновое основание MMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание 5SICS и второе нуклеиновое основание DMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание 5SICS и второе нуклеиновое основание FEMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание 5SICS и второе нуклеиновое основание BrMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание 5SICS и второе нуклеиновое основание IMS.

В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание SICS и второе нуклеиновое основание MMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание SICS и второе нуклеиновое основание DMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание SICS и второе нуклеиновое основание FEMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание SICS и второе нуклеиновое основание BrMS. В различных вариантах осуществления настоящего изобретения далее обеспечивается пара нуклеиновых оснований, содержащая первое нуклеиновое основание SICS и второе нуклеиновое основание IMS.

В различных вариантах осуществления настоящего изобретения далее обеспечивается двухцепочечный олигонуклеотидный дуплекс, в котором первая олигонуклеотидная цепь включает нуклеиновое основание, не встречающееся в природе (например, аналог нуклеинового основания), раскрытое в настоящем изобретении, и вторая комплементарная олигонуклеотидная цепь содержит комплементарное, способное к образованию пары нуклеиновое основание в сайте спаривания комплементарных оснований.

Например, для dTPT3, вторая комплементарная олигонуклеотидная цепь содержит dNaM, dDMO или dMMO2 или дериватизированный с помощью линкера их аналог в сайте спаривания комплементарных оснований. Таким образом, взаимодействие при образовании пар между первой олигонуклеотидной цепью и второй олигонуклеотидной цепью включает специфическое взаимодействие, возникающее при образовании пар нуклеиновых оснований, между остатком не встречающегося в природе нуклеинового основания, обеспеченного в настоящем изобретении, и комплементарным нуклеотидным основанием, которое может быть природным или не встречающимся в природе нуклеотидным основанием.

В некоторых вариантах осуществления настоящего изобретения обеспечивается двухцепочечный олигонуклеотидный дуплекс, в котором первая олигонуклеотидная цепь содержит соединение, имеющее формулу  $\beta 8a$  или  $\beta 8b$ , и вторая комплементарная олигонуклеотидная цепь содержит комплементарное, способное к образованию пары оснований нуклеиновое основание в сайте спаривания комплементарных оснований. В некоторых вариантах осуществления комплементарное, способное к образованию пары оснований нуклеиновое основание представляет собой соединение, имеющее формулу  $\alpha 14a$ ,  $\alpha 14b$ ,  $\alpha 14c$ ,  $\alpha 14d$ ,  $\alpha 14e$  или  $\alpha 14f$ . В некоторых вариантах осуществления комплементарное, способное к образованию пары оснований нуклеиновое основание включает, без ограничения, цитозин, гуанин, аденин, тимин, урацил, 2-аминоаденин-9-ил, 2-аминоаденин, 2-Ф-аденин, 2-тиоурацил, 2-тиотимин, 2-тиоцитозин, 2-пропил и алкильные производные аденина и гуанина, 2-аминоаденин, 2-аминопропил-аденин, 2-аминопиридин, 2-пиридон, 2'-дезоксуридин, 2-амино-2'-дезоксиаденозин 3-деазагуанин, 3-дезааденин, 4-тиоурацил, 4-тиотимин, урацил-5-ил, гипоксантин-9-ил (I), 5-метил-цитозин, 5-гидроксиметилцитозин, ксантин, гипоксантин, 5-бром и 5-трифторметил-урацилы и цитозины; 5-галоурацил, 5-галоцитозин, 5-пропинил-урацил, 5-пропинил-цитозин, 5-урацил, 5-замещенный, 5-галоген, 5-замещенные пиримидины, 5-гидроксицитозин, 5-бромцитозин, 5-бромурацил, 5-хлорцитозин, хлорированный цитозин, циклоцитозин, цитозинарабинозид, 5-фторцитозин, фторпиримидин, фторурацил, 5,6-дигидроцитозин, 5-иодоцитозин, гидроксимочевину, иодоурацил, 5-нитроцитозин, 5-бромурацил, 5-хлорурацил, 5-фторурацил и 5-иодоурацил, 6-алкильные производные аденина и гуанина, 6-азапиримидины, 6-азоурацил, 6-азоцитозин, азацитозин, 6-азотимин, 6-тиогуанин, 7-метилгуанин, 7-метиладенин, 7-деазагуанин, 7-деазагуанозин, 7-дезааденин, 7-деза-8-азагуанин, 8-азагуанин, 8-азааденин, 8-галоген, 8-амино, 8-тиол, 8-тиоалкил и 8-гидроксилзамещенные аденины и гуанины; N4-этилцитозин, N-2 замещенные пурины, N-6 замещенные пурины, O-6 замещенные пурины, те, которые повышают стабильность образования дуплекса, универсальные нуклеиновые кислоты, гидрофобные нуклеиновые кислоты, промискуитетные (promiscuous) нуклеиновые кислоты, нуклеиновые кислоты с увеличенным размером (size-expanded), фторированные нуклеиновые кислоты, трициклические пиримидины, феноксазинцитидин ([5,4-b][1,4]бензоксазин-2(3H)-он), фенотиазинцитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-он), "G-зажимы" (G-clamps), феноксазинцитидин (9-(2-аминоэтокси)-Н-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-он), карбазолцитидин (2H-пиримидо[4,5-b]индол-2-он), пиридоиндолцитидин (Н-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-он), 5-фторурацил, 5-бромурацил, 5-хлорурацил, 5-иодоурацил, гипоксантин, ксантин, 4-ацетилцитозин, 5-(карбоксигидроксиметил)урацил, 5-карбоксиметиламинометил-2-тиоуридин, 5-карбоксиметиламинометилурацил, дигидроурацил,  $\beta$ -D-галактозилквеозин, инозин, N6-изопентиладенин, 1-метилгуанин, 1-метилюридин, 2,2-диметилгуанин, 2-метиладенин, 2-метилгуанин, 3-метилцитозин, 5-метилцитозин, N6-аденин, 7-метилгуанин, 5-метиламинометилурацил, 5-метокси-аминометил-2-тиоурацил,  $\beta$ -D-маннозилквеозин, 5'-метоксикарбоксиметилурацил, 5-метоксиурацил, 2-метилтио-N6-изопентиладенин, урацил-5-оксисульфоновую кислоту, wybutoxosine, псевдоурацил, квеозин, 2-тиоцитозин, 5-метил-2-тиоурацил, 2-тиоурацил, 4-тиоурацил, 5-метилурацил, урацил-5-метиловый эфир оксисульфоновой кислоты, урацил-5-оксисульфоновую кислоту, 5-метил-2-тиоурацил, 3-(3-амино-3-N-2-карбоксипропил)урацил, (аср3)w и 2,6-диаминопурин и те соединения, в которых пуриновое или пиримидиновое основание заменено гетероциклом.

В некоторых вариантах осуществления настоящего изобретения обеспечивается двухцепочечный олигонуклеотидный дуплекс, в котором первая олигонуклеотидная цепь содержит соединение, имеющее формулу  $\alpha 14a$ ,  $\alpha 14b$ ,  $\alpha 14c$ ,  $\alpha 14d$ ,  $\alpha 14e$  или  $\alpha 14f$ , и вторая комплементарная олигонуклеотидная цепь содержит комплементарное, способное к образованию пары оснований нуклеиновое основание в сайте спаривания комплементарных оснований. В некоторых вариантах осуществления комплементарное, способное к образованию пары оснований нуклеиновое основание представляет собой соединение, имеющее формулу  $\beta 8a$  или  $\beta 8b$ . В некоторых вариантах осуществления комплементарное, способное к образованию пары оснований нуклеиновое основание включает, без ограничения, цитозин, гуанин, аденин, тимин, урацил, 2-аминоаденин-9-ил, 2-аминоаденин, 2-Ф-аденин, 2-тиоурацил, 2-тиотимин, 2-тиоцитозин, 2-пропил и алкильные производные аденина и гуанина, 2-аминоаденин, 2-аминопропил-аденин, 2-аминопиридин, 2-пиридон, 2'-дезоксуридин, 2-амино-2'-дезоксиаденозин 3-деазагуанин, 3-дезааденин, 4-тиоурацил, 4-тиотимин, урацил-5-ил, гипоксантин-9-ил (I), 5-метил-цитозин, 5-гидроксиметилцитозин, ксантин, гипоксантин, 5-бром и 5-трифторметил-урацилы и -цитозины; 5-галоурацил, 5-галоцитозин, 5-пропинил-урацил, 5-пропинил-цитозин, 5-урацил, 5-замещенный, 5-галоген, 5-замещенные пиримидины, 5-гидроксицитозин, 5-бромцитозин, 5-бромурацил, 5-хлорцитозин, хлорированный цитозин, циклоцито-

зин, цитозинарабинозид, 5-фторцитозин, фторпиримидин, фторурацил, 5,6-дигидроцитозин, 5-иодоцитозин, гидроксимочевину, иодоурацил, 5-нитроцитозин, 5-бром урацил, 5-хлорурацил, 5-фторурацил и 5-иодоурацил, 6-алкильные производные аденина и гуанина, 6-азапиримидины, 6-азоурацил, 6-азоцитозин, азацитозин, 6-азотимин, 6-тиогуанин, 7-метилгуанин, 7-метиладенин, 7-деазагуанин, 7-деазагуанозин, 7-деазаденин, 7-деаза-8-азагуанин, 8-азагуанин, 8-азааденин, 8-галоген, 8-амино, 8-тиол, 8-тиоалкил и 8-гидроксилзамещенные аденины и гуанины; N4-этилцитозин, N-2 замещенные пурины, N-6 замещенные пурины, O-6 замещенные пурины, те, которые повышают стабильность образования дуплекса, универсальные нуклеиновые кислоты, гидрофобные нуклеиновые кислоты, промискуитетные (promiscuous) нуклеиновые кислоты, нуклеиновые кислоты с увеличенным размером (size-expanded), фторированные нуклеиновые кислоты, трициклические пиримидины, феноксазинцитидин ([5,4-b][1,4]бензоксазин-2(3H)-он), фенотиазинцитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-он), "G-зажимы" (G-clamps), феноксазинцитидин (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-он), карбазолцитидин (2H-пиримидо[4,5-b]индол-2-он), пиридоиндолцитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-он), 5-фторурацил, 5-бром урацил, 5-хлорурацил, 5-иодоурацил, гипоксантин, ксантин, 4-ацетилцитозин, 5-(карбоксигидроксиметил)урацил, 5-карбоксиметиламинометил-2-тиоуридин, 5-карбоксиметиламинометилурацил, дигидроурацил,  $\beta$ -D-галактозилквеозин, инозин, N6-изопентиладенин, 1-метилгуанин, 1-метилюозин, 2,2-диметилгуанин, 2-метиладенин, 2-метилгуанин, 3-метилцитозин, 5-метилцитозин, N6-аденин, 7-метилгуанин, 5-метиламинометилурацил, 5-метоксиаминометил-2-тиоурацил,  $\beta$ -D-маннозилквеозин, 5'-метоксикарбоксиметилурацил, 5-метоксиурацил, 2-метилтио-N6-изопентиладенин, урацил-5-оксиуксусную кислоту, wybutoxosine, псевдоурацил, квеозин, 2-тиоцитозин, 5-метил-2-тиоурацил, 2-тиоурацил, 4-тиоурацил, 5-метилурацил, урацил-5-метилэфир оксиуксусной кислоты, урацил-5-оксиуксусную кислоту, 5-метил-2-тиоурацил, 3-(3-амино-3-N-2-карбоксипропил)урацил, (acr3)w и 2,6-диаминопурин и те соединения, в которых пуриновое или пиримидиновое основание заменено гетероциклом.

В некоторых вариантах осуществления по меньшей мере один  $R_2$  из нуклеинового основания в двухцепочечном олигонуклеотидном дуплексе представляет собой сопряженный линкер, связанный с карго. В некоторых вариантах осуществления карго представляет собой репортерную группу, белок или соединение, содержащее каталитическую функциональную группу.

В некоторых вариантах осуществления первую олигонуклеотидную цепь, содержащую аналог нуклеинового основания, описанного в настоящем изобретении, получают синтезом с нуклеотидным основанием, содержащим реакционноспособный линкер, с последующим соединением карго-реагента с первой олигонуклеотидной цепью, или где первую олигонуклеотидную цепь получают синтезом с нуклеотидным основанием, содержащим сопряженный линкер, связанный с карго.

В некоторых вариантах осуществления двухцепочечный олигонуклеотидный дуплекс имеет первую цепь, содержащую dTPT3 или его производное, и вторую цепь, содержащую dNaM, dDMO или dMMO2 или его производное, в сайте спаривания комплементарных оснований.

В различных вариантах осуществления настоящего изобретения далее обеспечивается способ проведения сайт-специфической функционализации двухцепочечного олигонуклеотидного дуплекса, содержащий включение не встречающегося в природе нуклеинового основания, содержащего реакционноспособный линкер, содержащий реакционноспособный центр, причем нуклеиновое основание имеет любую из следующих формул:  $\alpha$ 14a,  $\alpha$ 14b,  $\alpha$ 14c,  $\alpha$ 14d,  $\alpha$ 14e,  $\alpha$ 14f,  $\beta$ 8a,  $\beta$ 8b,  $\beta$ 9a,  $\beta$ 9b,  $\alpha$ 15a,  $\alpha$ 15b,  $\alpha$ 16a или  $\alpha$ 16b, в первую олигонуклеотидную цепь; затем синтез второй цепи, комплементарной первой цепи, причем вторая цепь содержит нуклеиновое основание, комплементарное не встречающемуся в природе нуклеиновому основанию в сайт-специфическом комплементарном положении в нем, при таких условиях, что первая цепь и вторая цепь образуют двухцепочечный олигонуклеотидный дуплекс; затем контактирование двухцепочечного олигонуклеотидного дуплекса, включающего не встречающегося в природе нуклеиновое основание, содержащего реакционноспособный линкер, с карго-реагентом, содержащим карго и группу, обладающую комплементарной реакционной способностью, в условиях, подходящих для осуществления реакции взаимодействия между реакционноспособным линкером и группой, обладающей комплементарной реакционной способностью, с получением сопряженного линкера; чтобы обеспечить функционализированный двухцепочечный олигонуклеотидный дуплекс с карго, присоединенным к нему через сопряженный линкер.

В различных вариантах осуществления настоящего изобретения далее обеспечивается способ проведения сайт-специфической функционализации двухцепочечного олигонуклеотидного дуплекса, содержащий включение не встречающегося в природе нуклеинового основания, содержащего реакционноспособный линкер, содержащий реакционноспособный центр, причем нуклеиновое основание выбрано из группы, состоящей из

d5SICSCO, d5SICSCC, dDMOCO, dDMOCC, dMMO2pCO, dMMO2pCC,

dTPT3, dTPT3A, dTPT3PA, dTPT3CO и dTPT3CC,

в первую олигонуклеотидную цепь; затем синтез второй цепи, комплементарной первой цепи, причем вторая цепь содержит нуклеиновое основание, комплементарное не встречающемуся в природе нук-



леиновому основанию в сайт-специфическом комплементарном положении в нем, при таких условиях, что первая цепь и вторая цепь образуют двухцепочечный олигонуклеотидный дуплекс; затем контактирование двухцепочечного олигонуклеотидного дуплекса, включающего не встречающееся в природе нуклеиновое основание, содержащее реакционноспособный линкер, с карго-реагентом, содержащим карго и группу, обладающую комплементарной реакционной способностью, в условиях, подходящих для осуществления реакции взаимодействия между реакционноспособным линкером и группой, обладающей комплементарной реакционной способностью, с получением сопряженного линкера; чтобы обеспечить функционализированный двухцепочечный олигонуклеотидный дуплекс с карго, присоединенным к нему через сопряженный линкер.

В одном из вариантов осуществления линкер связывается с карго после того, как соответствующий 5'-трифосфат встраивается в ДНК- или РНК-олигонуклеотид с использованием ДНК- или РНК-полимеразы (после снятия защиты, если это необходимо). В другом варианте осуществления синтезируется второй олигонуклеотид, который является комплементарным первой цепи, причем вторая цепь содержит не встречающийся в природе нуклеотид в положении, комплементарном не встречающемуся в природе нуклеотиду первой цепи; затем взаимодействие полученного двухцепочечного олигонуклеотида с несущим карго реагентом, который селективно реагирует с реакционноспособным центром реакционноспособного линкера, что обеспечивает функционализированный двухцепочечный олигонуклеотид (например, ДНК/ДНК, ДНК/РНК или РНК/РНК), несущий карго.

В различных вариантах осуществления настоящего изобретения далее обеспечиваются структуры, содержащие формулу  $N1 - Zx - N2$ , где  $N1$  представляет собой нуклеотид или его аналог или концевую фосфатную группу;  $N2$  представляет собой нуклеотид или его аналог или концевую гидроксильную группу;  $Z$  представляет собой соединение, имеющее любую из формул  $\alpha14a$ ,  $\alpha14b$ ,  $\alpha14c$ ,  $\alpha14d$ ,  $\alpha14e$ ,  $\alpha14f$ ,  $\beta8a$ ,  $\beta8b$ ,  $\beta9a$ ,  $\beta9b$ ,  $\alpha15a$ ,  $\alpha15b$ ,  $\alpha16a$  или  $\alpha16b$ ; и  $x$  представляет собой целое число от 1 до 20. В некоторых вариантах осуществления структура представляет собой олигонуклеотид. В некоторых вариантах осуществления олигонуклеотид представляет собой рибонуклеиновую кислоту или дезоксирибонуклеиновую кислоту. В некоторых вариантах осуществления олигонуклеотид представляет собой аптамер или датчик на основе нуклеиновой кислоты. В некоторых вариантах осуществления олигонуклеотид представляет собой гибридизационный зонд по типу "молекулярного маяка". В некоторых вариантах осуществления олигонуклеотид является аналогом РНК или аналогом ДНК.

В различных вариантах осуществления настоящего изобретения далее обеспечивается способ идентификации аптамера нуклеиновой кислоты, содержащего по меньшей мере одно соединение, обеспеченное в настоящем изобретении (например,  $\alpha14a$ ,  $\alpha14b$ ,  $\alpha14c$ ,  $\alpha14d$ ,  $\alpha14e$ ,  $\alpha14f$ ,  $\beta8a$ ,  $\beta8b$ ,  $\beta9a$ ,  $\beta9b$ ,  $\alpha15a$ ,  $\alpha15b$ ,  $\alpha16a$ ,  $\alpha16b$ ), как имеющего улучшенное желаемое свойство по отношению к молекуле-мишени, причем указанный способ включает: а) приготовление смеси кандидатов из аптамеров одноцепочечных нуклеиновых кислот, причем каждый аптамер нуклеиновой кислоты смеси кандидатов из аптамеров содержит по меньшей мере одно соединение, обеспеченное в настоящем изобретении (например,  $\alpha14a$ ,  $\alpha14b$ ,  $\alpha14c$ ,  $\alpha14d$ ,  $\alpha14e$ ,  $\alpha14f$ ,  $\beta8a$ ,  $\beta8b$ ,  $\beta9a$ ,  $\beta9b$ ,  $\alpha15a$ ,  $\alpha15b$ ,  $\alpha16a$ ,  $\alpha16b$ ); затем б) контактирование смеси кандидатов с молекулой-мишенью в условиях, подходящих для того, чтобы произошло связывание с молекулой-мишенью; затем в) выделение одного или более аптамера нуклеиновой кислоты, имеющего желаемое свойство по отношению к молекуле-мишени, из числа аптамеров из смеси кандидатов; а затем д) амплификация одного или более аптамера нуклеиновой кислоты с желаемым свойством, *in vitro*, с получением одного или более аптамеров нуклеиновых кислот, имеющих улучшенное желаемое свойство по отношению к молекуле-мишени. В некоторых вариантах осуществления способ дополнительно включает стадию е) повторения стадий в) и д). В некоторых вариантах осуществления аптамеры одноцепочечных нуклеиновых кислот выбраны из группы, состоящей из одноцепочечной ДНК и одноцепочечной РНК. В некоторых вариантах осуществления желаемое свойство представляет собой связывающую способность для мишени. В некоторых вариантах осуществления желаемое свойство представляет собой активность, индуцированную связыванием с мишенью. В некоторых вариантах осуществления желаемое свойство представляет собой каталитическую активность. В некоторых вариантах осуществления желаемое свойство представляет собой ингибирующую активность, активирующую активность или модификацию ингибирующей активности или активирующей активности. В некоторых вариантах осуществления желаемое свойство представляет собой активность, переключающую структуру, или модификацию активности, переключающей структуру. В некоторых вариантах осуществления желаемое свойство представляет собой совместную активность. В некоторых вариантах осуществления желаемая активность представляет собой улучшенную эффективность клеточного обновления.

В некоторых вариантах осуществления настоящего изобретения далее обеспечивается аптамер, содержащий соединение, имеющее любую из следующих формул:  $\alpha14a$ ,  $\alpha14b$ ,  $\alpha14c$ ,  $\alpha14d$ ,  $\alpha14e$ ,  $\alpha14f$ ,  $\beta8a$ ,  $\beta8b$ ,  $\beta9a$ ,  $\beta9b$ ,  $\alpha15a$ ,  $\alpha15b$ ,  $\alpha16a$ ,  $\alpha16b$ .

Фразы, такие как "в условиях, подходящих для обеспечения" или "в условиях, достаточных, чтобы получить" или т.п., в контексте способов синтеза, как используется здесь, относятся к условиям реакции, таким как время, температура, растворитель, концентрации реагентов и т.п., варьирование которых дос-

тупно квалифицированному экспериментатору с тем, чтобы обеспечить достаточное количество или выход продукта реакции. Не обязательно, чтобы желаемый продукт реакции был единственным продуктом реакции или чтобы исходные вещества были полностью израсходованы, при условии, что желаемый продукт реакции может быть выделен или иным образом впоследствии использован.

Выражение "химически возможно" означает расположение связей или соединение, где обычно понимаемые правила органической структуры не были нарушены; например, должно быть понятно, что структура в рамках определения пункта формулы изобретения, которая будет содержать в определенных ситуациях пентавалентный атом углерода, который не существуют в природе, не может быть в пункте формулы изобретения. Структуры, описанные в настоящем изобретении, во всех вариантах их осуществления предназначены для того, чтобы включать только "химически возможные" структуры, и любые перечисленные структуры, которые не являются химически возможными, например, в структуре, показанной с изменяющимися атомами или группами, не предназначены для того, чтобы быть раскрытыми или заявленными в настоящем изобретении.

"Аналог" химической структуры, как данный термин используется в настоящем описании, относится к химической структуре, которая сохраняет существенное сходство с исходной структурой, хотя она не может быть легко получена синтетически из исходной структуры. В некоторых вариантах осуществления аналог нуклеотида представляет собой нуклеотид, не встречающийся в природе. В некоторых вариантах осуществления аналог нуклеозида представляет собой нуклеозид, не встречающийся в природе. Родственная химическая структура, которая может быть легко получена синтетически из исходной химической структуры, именуется как "производное" (дериват).

Соответственно, "ДНК-аналог" или "РНК-аналог", как данные термины используются в настоящем описании, относятся к ДНК- или РНК-подобным полимерам, таким как пептидонуклеиновые кислоты (peptide nucleic acids (PNA)), замкнутые нуклеиновые кислоты (locked nucleic acids (LNA)), фосфоротиоаты и т.п., которые хорошо известны в данной области техники. ДНК- и РНК-аналоги, а также ДНК и РНК могут быть синтезированы в автоматизированных синтезаторах, например, с помощью фосфоамидитного химического способа или других химических подходов, приспособленных для применения синтезатора.

ДНК включает, но не ограничивается ими, комплементарную ДНК (кДНК) и геномную ДНК. ДНК может быть присоединена ковалентной или нековалентной связью к другой биомолекуле, включая, но не ограничиваясь ими, РНК и пептид. РНК включает кодирующую РНК, например матричную (информационную) РНК (мРНК). В некоторых вариантах осуществления РНК представляет собой rRNA, RNAi, snoRNA, microRNA, siRNA, snRNA, exRNA, piRNA, long ncRNA или их любую комбинацию или гибрид. В некоторых случаях РНК является компонентом рибозима. ДНК и РНК могут быть в любой форме, включая, но не ограничиваясь этим, линейную, круговую, суперспирализованную, одноцепочечную и двухцепочечную.

Термин "аминозащитная группа" или "аминозащищенная", используемый здесь, относится к тем группам, которые предназначены для защиты аминогруппы от нежелательных реакций в процессе синтеза и которые позднее могут быть удалены, чтобы высвободить амин. Чаще всего используемые аминозащитные группы описаны в *Protective Groups in Organic Synthesis*, Greene, T.W.; Wuts, P.G.M., John Wiley & Sons, New York, NY, (3rd Edition, 1999). Аминозащитные группы включают ацильные группы, такие как формильная, ацетильная, пропионильная, пивалоильная, трет-бутилацетилхлорид, 2-хлорацетильная, 2-бромацетильная, трифторацетильная, трихлорацетильная, о-нитрофеноксиацетильная, α-хлорбутирильная, бензоильная, 4-хлорбензоильная, 4-бромбензоилхлорид, 4 нитробензоильная и т.п.; сульфонильные группы, такие как бензолсульфонильная, п-толуолсульфонильная и т.п.; алкокси- или арилоксикарбонильные группы (которые образуют уретаны с защищенным амином), такие как бензилоксикарбонильная (Cbz), п-хлорбензилоксикарбонильная, п-метоксибензилоксикарбонильная, п-нитробензилоксикарбонильная, 2-нитробензилоксикарбонильная, п-бромбензилоксикарбонильная, 3,4-диметоксибензилоксикарбонильная, 3,5-диметоксибензилоксикарбонильная, 2,4-диметоксибензилоксикарбонильная, 4-метоксибензилоксикарбонильная, 2-нитро-4,5-диметоксибензилоксикарбонильная, 3,4,5-триметоксибензилоксикарбонильная, 1-(п-бифенил)-1-метилэтоксикарбонильная, α,α-диметил-3,5-диметоксибензилоксикарбонильная, бензгидрилоксикарбонильная, трет-бутилоксикарбонильная (Boc), диизопропилметоксикарбонильная, изопропоксикарбонильная, этоксикарбонильная, метоксикарбонильная, аллилоксикарбонильная (Alloc), 2,2,2-трихлорэтоксикарбонильная, 2-триметилсилилэтоксикарбонильная (Teos), феноксикарбонильная, 4-нитрофеноксикарбонильная, флуоренил-9-метоксикарбонильная (Fmoc), циклопентилоксикарбонильная, адамантилоксикарбонильная, циклогексилоксикарбонильная, фенилтиокарбонильная и т.п.; аралкильные группы, такие как бензильная, трифенилметильная, бензилоксиметильная и т.п.; и силильные группы, такие как триметилсилильная и т.п. Аминозащитные группы также включают циклические аминозащитные группы, такие как фталоильная и дитиосукцинимидильная, которые включают атом азота аминогруппы в гетероцикл. Как правило, аминозащитные группы включают формил, ацетил, бензоил, пивалоил, трет-бутилацетилхлорид, фенилсульфонил, Alloc, Teos, бензил, Fmoc, Boc и Cbz. Защитные группы также включают метилкарбамат, 9-

флуоренилметилкарбамат, 2,2,2-трихлорэтилкарбамат, трет-бутилкарбамат, 2-(триметилсилил)этилкарбамат, аллилкарбамат, бензилкарбамат, м-нитрофенилкарбамат, трифторацетамид, бензиламин, аллиламин и тритиламин. Защитные группы также включают формамиды, ацетамиды, трифторацетамиды, п-толуолсульфонил, трифторметансульфонил, триметилсилилэтансульфонамид и трет-бутилсульфонил. Специалист с обычной квалификацией может выбрать и использовать соответствующую аминозащитную группу для синтетической задачи.

ДНК- и РНК-аналоги включают аналоги PNA (peptide nucleic acid, пептидонуклеиновая кислота) и LNA (locked nucleic acid, замкнутая нуклеиновая кислота).

Пептидонуклеиновая кислота (PNA) является синтетическим аналогом ДНК/РНК, в котором пептидоподобный остов заменяет сахарофосфатный остов ДНК или РНК. Олигомеры PNA показывают более высокую прочность связывания и больше специфичности в связывании с комплементарными ДНК, причем несоответствие оснований PNA/ДНК является более дестабилизирующим, чем аналогичное несоответствие в дуплексе ДНК/ДНК. Такая прочность и специфичность связывания также относится к дуплексам PNA/РНК. PNA не легко узнаваемы нуклеазами или протеазами, что делает их устойчивыми к ферментной деградации. PNA также стабильны в широком диапазоне pH. См. также Nielsen P.E., Egholm M., Berg R.H., 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 S.M., Driver D.A., Berg R.H., Kim S.K., Nordén B., and Nielsen P.E. (1993), "PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen Bonding Rules". *Nature* 365 (6446): 566-8. doi:10.1038/365566a0. PMID 7692304.

Замкнутая нуклеиновая кислота (LNA) представляет собой модифицированный РНК-нуклеотид, в котором рибозный остаток LNA-нуклеотида модифицируется добавочным мостиком, соединяющим 2'-кислород и 4'-углерод. Мостик "замыкает" (блокирует) рибозу в конформации 3'-эндо (North), которая часто обнаруживается в дуплексах А-формы. LNA-нуклеотиды могут быть смешаны с нуклеотидами ДНК или РНК в олигонуклеотиде по желанию. Такие олигомеры могут быть синтезированы химически и являются коммерчески доступными. Конформация замкнутой рибозы усиливает стэкинг оснований и преорганизацию остова. См., например, публикации 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 pucker", *Tetrahedron Lett.* 38 (50): 8735-8. doi:10.1016/S0040-4039(97)10322-7.

Молекулярный маяк (molecular beacon) или зонд по типу "молекулярного маяка" представляет собой олигонуклеотидный гибридизационный зонд, который может обнаружить присутствие определенной последовательности нуклеиновой кислоты в гомогенном растворе. Молекулярные маяки представляют собой молекулы в форме шпильки с погашенным флуорофором в исходной конформации, чья флуоресценция восстанавливается, когда они связываются с целевой последовательностью нуклеиновой кислоты. См., например, публикации Tyagi S., Kramer F.R. (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 A.C. (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 нуклеиновое основание обычно является частью нуклеозида, представляющей собой гетероциклическое основание. Нуклеотидные основания могут быть естественного происхождения, могут быть модифицированными, могут не иметь никакого сходства с природными основаниями и могут быть синтезированы, например, органическим синтезом. В определенных вариантах осуществления нуклеиновое основание содержит любой атом или группу атомов, способных взаимодействовать с основанием другой нуклеиновой кислоты с использованием или без использования водородных связей. В определенных вариантах осуществления не встречающееся в природе нуклеиновое основание не происходит от природного нуклеинового основания. Следует отметить, что не встречающиеся в природе нуклеотидные основания не всегда обладают свойствами оснований, однако они называются нуклеотидными основаниями для простоты. В некоторых вариантах осуществления, когда упоминают нуклеиновое основание, "(d)" указывает на то, что нуклеиновое основание может быть присоединено к дезоксирибозе или рибозе.

В некоторых вариантах осуществления нуклеозид представляет собой соединение, содержащее ос-

таток нуклеинового основания и остаток сахара. Нуклеозиды включают, но не ограничиваются ими, нуклеозиды естественного происхождения (найденные в ДНК и РНК), абазические нуклеозиды, модифицированные нуклеозиды и нуклеозиды, имеющие миметические группы оснований и/или сахаров. Нуклеозиды включают нуклеозиды, содержащие любое множество заместителей. Нуклеозид может быть гликозидным соединением, образующимся через гликозидную сшивку между основанием нуклеиновой кислоты и восстанавливающей группой сахара.

В некоторых вариантах нуклеотид представляет собой соединение, в котором сахарный остаток нуклеозида образует сложный эфир с фосфорной кислотой, более предпочтительно моно-, ди- или трифосфатный эфир. Сахарный остаток такого нуклеозида или нуклеотида может представлять собой рибофуранозил, 2'-дезоксирiboфуранозил или 2'-замещенный рибофуранозил, имеющий заместитель в 2' положении. Аналогично, остаток фосфорной кислоты может быть тиофосфорной кислотой. А именно, остатки сахара и фосфорной кислоты могут быть в том же виде, как найдено в известных нуклеозидах, нуклеотидах или их производных. Рибонуклеотид, сахарный остаток которого представляет собой рибофуранозил, может быть использован в качестве элемента, составляющего РНК. Дезоксирибонуклеотид, сахарный остаток которого представляет собой дезоксирибофуранозил, может быть использован в качестве элемента, составляющего ДНК. Нуклеотид может представлять собой нуклеозид, дополнительно содержащий фосфат-связывающую группу. Нуклеотиды могут включать нуклеозиды, содержащие фосфатную группу.

Был разработан класс не встречающихся в природе пар оснований, примерами которых служат d5SICS-dNaM и d5SICS-dMMO2 (фиг. 1), для которых авторы изобретения показали репликацию (включая с помощью ПНР) и транскрипцию с помощью разнообразных естественных полимераз с эффективностью и точностью воспроизведения, приближающимися к таковым природной пары оснований (см. Malyshev, D.A.; Seo, Y.J.; Ordoukhanian, P.; Romesberg, F.E., PCR with an Expanded Genetic Alphabet. *J. Am. Chem. Soc.* 2009, 131 (41), 14620-14621; Seo, Y.J.; Matsuda, S.; Romesberg, F.E., Transcription of an Expanded Genetic Alphabet. *J. Am. Chem. Soc.* 2009, 131 (14), 5046-5047; Laverne T.; Degardin M.; Malyshev D.A.; Quach H.T.; Dhami K.; Ordoukhanian P.; Romesberg, F.E.; Expanding the scope of replicable unnatural DNA: Stepwise optimization of a predominantly hydrophobic base pair. *J. Am. Chem. Soc.* 2013, 135, 5408-5419; Seo, Y.J., Malyshev D.A., Laverne T., Ordoukhanian P., and Romesberg, F.E., *J. Am. Chem. Soc.* 2011, 133, 19878.). Данные не встречающиеся в природе пары оснований образуются между нуклеотидными аналогами, несущими не встречающиеся в природе, преимущественно гидрофобные нуклеотидные основания. Пары оснований показаны на фиг. 1, причем каждый не встречающийся в природе нуклеотид включен в олигонуклеотиды в комплементарных (т.е. допускающих образование пары) положениях; считается, что нуклеотидные основания соединены в положении, обозначенном волнистой линией, с 1' положением рибозильного или 2' положением дезоксирибозильного остатка, который сам по себе включен в РНК или ДНК фосфатными или фосфоротиоатными группами, связывающими 3'- и 5'-гидроксильные группы рибозильных или дезоксирибозильных групп, соответственно, как это происходит в случае с полностью естественными нуклеиновыми кислотами. Образование пар оснований, таким образом, имеет место как часть процесса образования структуры, содержащей комплементарные спаренные основания, как это хорошо известно в случае образования структур олигонуклеотидных дуплексов. В различных вариантах осуществления настоящего изобретения обеспечивается не встречающееся в природе нуклеиновое основание dTPT3 и его варианты, дериватизированные с помощью линкера, которые, как полагают, образуют пары с не встречающимися в природе нуклеотидными основаниями dNaM, dMMO2 и dDMO (или их вариантами, дериватизированными с помощью линкера) аналогичным образом.

Авторы изобретения продемонстрировали, что не встречающиеся в природе нуклеотиды dTPT3 и dTPT3PA эффективно включаются в ДНК ДНК-полимеразами в положении напротив dNaM (фиг. 2). Также ожидается, что как dTPT, так и dTPT3<sup>PA</sup> (PA = дихлорацетил-3-аминопропин-1-ил), а также другие, дериватизированные с помощью линкера варианты dTPT3 (фиг. 3), включая, когда R = 3-аминопропин-1-ил (dTPT3<sup>A</sup>), R = 4-оксагепат-1,6-диин-1-ил (dTPT3<sup>CO</sup>), или R = гепта-1,6-диин-1-ил (dTPT3<sup>CC</sup>), будут спариваться с dNaM, dDMO или dMMO2 или их аналогами, дериватизированными с помощью линкера. Скорости включения dTPT3 и его вариантов, дериватизированных с помощью линкера, в положении напротив dNaM приближаются к таковым природной пары оснований. Дополнительные не встречающиеся в природе пары оснований, для которых были показаны высокие скорости встраивания, включают

dTPT3-dFEMO, dTPT3-

dFIMO, dTPT3-dIMO, dFTPT3-dNaM, dFTPT3-dFEMO, dFTPT3-dFIMO и dFTPT3-dIMO.

В различных вариантах осуществления настоящего изобретения далее обеспечиваются не встречающиеся в природе нуклеотиды с аналогами нуклеиновых оснований, включая  $\alpha$ -аналоги (например, любой из фиг. 8, 10, 11, 15 и его производные);  $\beta$ -аналоги (например, любой из фиг. 9, 12, 13, 16 и его производные);

d5SICS<sup>CO</sup>, d5SICS<sup>CC</sup>, dDMO<sup>CO</sup>, dDMO<sup>CC</sup>,

dMMO2<sup>CO</sup>, dMMO2<sup>CC</sup>, dTPT3, dTPT3<sup>PA</sup>, dTPT3<sup>A</sup>, dTPT3<sup>CO</sup>, dTPT3<sup>CC</sup>

и их рибозильные формы и их аналоги (фиг. 2); в форме нуклеозидов, нуклеозид-5'-трифосфатов и их аналогов (например, рибозил и 2'-дезоксирибозил), нуклеотидов и их аналогов (например, рибозил и 2'-дезоксирибозил, фосфат и форсфототиоат), включая нуклеотидные реагенты, полученные из них для использования в синтезе РНК/ДНК (DMT-защищенные фосфорамидаты) и для использования в ферментативном включении в олигонуклеотиды с помощью ПЦР или опосредованной T7 РНК-полимеразой транскрипции, и включенные в нуклеиновые кислоты (олигонуклеотиды), такие как ДНК и РНК. Соединения, содержащие аналоги не встречающихся в природе нуклеиновых оснований, также могут быть включены в ДНК-аналоги или РНК-аналоги, например PNA, LNA и другие подобные полинуклеотид-аналогичные полимеры. Примеры аналогов нуклеиновых оснований, обеспеченных в настоящем изобретении, включают β8-аналоги, включающие формулы β8a и β8b, как показано на фиг. 12, где каждый X независимо представляет собой углерод или азот; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород; и где R является необязательным, и когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго. Примеры β8-аналогов включают dTPT3 и его дериватизированные с помощью линкера аналоги. Примеры аналогов нуклеиновых оснований, обеспеченных в настоящем изобретении, включают α14-аналоги, включающие формулы α14a-α14f, как показано на фиг. 10, где каждый X независимо представляет собой атом углерода или азота; где каждый R<sub>1</sub> независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый R<sub>2</sub> является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый R является необязательным, и когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород. Примеры α14-аналогов включают dMMS, dDMS, dFEMS, dBrMS, dIMS и их дериватизированные с помощью линкера аналоги.

В различных вариантах осуществления настоящего изобретения далее обеспечиваются не встречающиеся в природе пары оснований, включающие любой α-аналог или его производное, раскрытое в настоящем изобретении, и любой β-аналог или его производное, раскрытое в настоящем изобретении. Производные включают, но не ограничиваются ими, замены атомов и добавления линкерных групп. Линкерные группы могут быть прикреплены к аналогам в процессе синтеза или после включения нуклеинового основания в нуклеиновую кислоту. Примеры не встречающихся в природе пар оснований, включают, но не ограничиваются ими,

dTPT3-dNaM, dTPT3-dFEMO, dTPT3-dFIMO, dTPT3-dIMO,

dFTPT3-dNaM, dFTPT3-dFEMO, dFTPT3-dFIMO, dFTPT3-dIMO.

Не встречающиеся в природе пары оснований включают, но не ограничиваются ими,

dTPT3-MMS, dTPT3-DMS, dTPT3-FEMS, dTPT3-BrMS, dTPT3-IMS, dTPT3-dDMN, dTPT3-d4OMe, dTPT3-dIQ, dTPT3-d2MN, dTPT3-d3OMe, dTPT3-dQL, dTPT3-d2Np, dTPT3-dDM4, dTPT3-dDM, dTPT3-dBEN, dTPT3-d3FB, dTPT3-dMM1, dTPT3-dMMO1, dTPT3-dDM2, dTPT3-dDM5, dTPT3-d2Py, dTPT3-d5MPy, dTPT3-dEPy, dTPT3-d3MPy, dTPT3-d34DMPy, dTPT3-d45DMPy, dTPT3-d4MPy, dTPT3-d35DMPy, dTPT3-dBP, dTPT3-dBTp, dTPT3-dBF, dTPT3-dIN, dTPT3-dTp, dTPT3-dBTz, dTPT3-dMTp, dTPT3-dAM, dTPT3-dMAN, dTPT3-dMMAN, dTPT3-dADM, dTPT3-dMMAN, dTPT3-dTOK588, dTPT3-dTOK576, dTPT3-dTOK587, dTPT3-dTOK586, dTPT3-dTOK580, dTPT3-dPhMO, dTPT3-dPyMO1, dTPT3-PyMO2, dTPT3-dPMO1, dTPT3-dPMO2, dTPT3-dPMO3, dTPT3-dFuMO1, dTPT3-dFuMO2, dTPT3-TpMO1, dTPT3-dTpMO2, dTPT3-dFIMO, dTPT3-dIMO, dTPT3-dMIMO, dTPT3-dMEMO, dTPT3-dFEMO, dTPT3-dPrMO, dTPT3-dMMO2, dTPT3-d2OMe, dTPT3-dDMO, dTPT3-dTMO, dTPT3-dNMO, dTPT3-dNOPy, dTPT3-d5FM, dTPT3-dNAM, dTPT3-dAMO1, dTPT3-dAPy, dTPT3-dAMO2, dTPT3-dMAPy, dTPT3-dAMO3, dTPT3-dDMAPy, dTPT3-dFDMO, dTPT3-dVMO, dTPT3-dQMO, dTPT3-dZMO, dTPT3-dCIMO, dTPT3-dTfMO, dTPT3-CNMO, d7AI-dMMS, dM7AI-dMMS, dImPy-dMMS, dP7AI-dMMS, dPPP-dMMS, d8Q-dMMS, dICS-dMMS, dPICS-dMMS, dMICS-dMMS, d4MICS-dMMS, d5MICS-dMMS, dNICS-dMMS, dONICS-dMMS, d7OFP-dMMS, d7OTP-dMMS, d4OTP-dMMS, dPYR-dMMS, d4MP-dMMS, d3MP-dMMS, dPPYR-dMMS, dMOP-dMMS, d4MOP-dMMS, dSICS-dMMS, dSNICS-dMMS, d5SICS-dMMS, d4SICS-dMMS, dTPT1-dMMS, dTPT2-dMMS, dFPT1-dMMS, dFTPT3-dMMS, d7AI-dDMS, dM7AI-dDMS, dImPy-dDMS, dP7AI-dDMS, dPPP-dDMS, d8Q-dDMS, dICS-dDMS, dPICS-dDMS, dMICS-dDMS, d4MICS-dDMS, d5MICS-dDMS, dNICS-dDMS, dONICS-dDMS, d7OFP-dDMS, d7OTP-dDMS, d4OTP-dDMS, dPYR-dDMS, d4MP-dDMS, d3MP-dDMS, dPPYR-dDMS, dMOP-dDMS, d4MOP-dDMS, dSICS-dDMS, dSNICS-dDMS, d5SICS-dDMS, d4SICS-dDMS, dTPT1-dDMS, dTPT2-dDMS, dFPT1-dDMS, dFTPT3-dDMS, d7AI-dFEMS, dM7AI-dFEMS, dImPy-dFEMS, dP7AI-dFEMS, dPPP-dFEMS, d8Q-dFEMS, dICS-dFEMS, dPICS-dFEMS, dMICS-dFEMS, d4MICS-dFEMS, d5MICS-dFEMS, dNICS-dFEMS, dONICS-dFEMS, d7OFP-dFEMS, d7OTP-dFEMS, d4OTP-dFEMS, dPYR-dFEMS, d4MP-dFEMS, d3MP-dFEMS, dPPYR-dFEMS, dMOP-dFEMS, d4MOP-dFEMS, dSICS-dFEMS, dSNICS-dFEMS, d5SICS-dFEMS, d4SICS-dFEMS, dTPT1-dFEMS, dTPT2-dFEMS, dFPT1-dFEMS, dFTPT3-dFEMS, d7AI-dBrMS, dM7AI-dBrMS, dImPy-dBrMS, dP7AI-dBrMS, dPPP-dBrMS, d8Q-dBrMS, dICS-dBrMS, dPICS-dBrMS, dMICS-dBrMS, d4MICS-dBrMS, d5MICS-dBrMS, dNICS-dBrMS, dONICS-dBrMS, d7OFP-dBrMS, d7OTP-dBrMS, d4OTP-dBrMS, dPYR-dBrMS, d4MP-dBrMS, d3MP-dBrMS, dPPYR-dBrMS, dMOP-dBrMS, d4MOP-dBrMS, dSICS-dBrMS, dSNICS-dBrMS, d5SICS-dBrMS, d4SICS-dBrMS, dTPT1-dBrMS, dTPT2-dBrMS, dFPT1-dBrMS, dFTPT3-dBrMS, d7AI-dIMS, dM7AI-dIMS, dImPy-dIMS, dP7AI-dIMS, dPPP-dIMS, d8Q-dIMS, dICS-dIMS, dPICS-dIMS, dMICS-dIMS, d4MICS-dIMS, d5MICS-dIMS, dNICS-dIMS, dONICS-dIMS, d7OFP-dIMS, d7OTP-dIMS, d4OTP-dIMS, dPYR-dIMS, d4MP-dIMS, d3MP-dIMS, dPPYR-dIMS, dMOP-dIMS, d4MOP-dIMS, dSICS-dIMS, dSNICS-dIMS, d5SICS-dIMS, d4SICS-dIMS, dTPT1-dIMS, dTPT2-dIMS, dFPT1-dIMS и dFTPT3-dIMS;

где одно или два не встречающихся в природе нуклеиновых оснований в не встречающейся в природе паре оснований могут быть дериватизированы с помощью линкера. Примеры не встречающихся в

природе пар оснований данного раскрытия дополнительно включают любую пару, описанную в примере 1. Примеры  $\beta$ -аналогов включают те, которые представлены на фиг. 9, 12, и 13. Примеры  $\beta$ -аналогов нуклеиновых оснований включают  $\beta$ 8-аналоги, включающие формулы  $\beta$ 8a и  $\beta$ 8b, как показано на фиг. 12, где каждый X независимо представляет собой углерод или азот; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород; и где R является необязательным, и когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго. Примеры  $\beta$ -аналогов включают dTPT3, d5SICS, dFTPT3 и их производные или аналоги. Примеры  $\alpha$ -аналогов включают те, которые представлены на фиг. 8, 10, и 11. Примеры  $\alpha$ -аналогов включают  $\alpha$ 14-аналоги, включающие формулы  $\alpha$ 14a- $\alpha$ 14f, как показано на фиг. 10, где каждый X независимо представляет собой атом углерода или азота; где каждый  $R_1$  независимо представляет собой водород, алкильную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, или сопряженный линкер, к которому присоединен карго; где каждый  $R_2$  является необязательным и, когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый R является необязательным, и когда присутствует, независимо представляет собой водород, алкил, алкенил, алкинил, метил, метантиол, метанселен, галоген, циано, азидную группу, реакционноспособный линкер, содержащий реакционноспособный центр, приспособленный к связыванию с карго-реагентом, содержащим карго и группу, обладающую реакционной способностью, комплементарной реакционноспособному центру, сопряженный линкер, к которому присоединен карго; где каждый Y независимо представляет собой серу, кислород, селен или вторичный амин; где каждый E независимо представляет собой серу, селен или кислород. Примеры  $\alpha$ -аналогов включают dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO и их производные или аналоги. В некоторых вариантах осуществления пара оснований, не встречающаяся в природе, содержит  $\alpha$ -аналог и природное основание. В некоторых вариантах осуществления пара оснований, не встречающаяся в природе, включает  $\beta$ -аналог и природное основание. В некоторых аспектах настоящего изобретения далее обеспечиваются не встречающиеся в природе пары оснований, включающие два одинаковых аналога не встречающихся в природе нуклеозидов или их производные.

Не встречающаяся в природе пара оснований в различных аспектах содержит одно не встречающееся в природе нуклеиновое основание, раскрытое в настоящем изобретении (например,  $\alpha$ -аналог или его производное,  $\beta$ -аналог или его производное), и другое не встречающееся в природе нуклеиновое основание, включая, но не ограничиваясь ими, 2-аминоаденин-9-ил, 2-аминоаденин, 2-F-аденин, 2-тиоурацил, 2-тиотимин, 2-тиоцитозин, 2-пропил и алкильные производные аденина и гуанина, 2-аминоаденин, 2-аминопропил-аденин, 2-аминопиридин, 2-пиридон, 2'-дезоксигуанидин, 2-амино-2'-дезоксиаденозин, 3-дезагуанин, 3-дезааденин, 4-тиоурацил, 4-тиотимин, урацил-5-ил, гипоксантин-9-ил (I), 5-метил-цитозин, 5-гидроксиметил-цитозин, ксантин, гипоксантин, 5-бром- и 5-трифторметилурацилы и цитозины; 5-галоурацил, 5-галоцитозин, 5-пропинил-урацил, 5-пропинил-цитозин, 5-урацил, 5-замещенный, 5-галоген, 5-замещенные пиримидины, 5-гидроксицитозин, 5-бромцитозин, 5-бромурацил, 5-хлорцитозин, хлорированный цитозин, циклоцитозин, цитозинарабинозид, 5-фторцитозин, фторпиримидин, фторурацил, 5,6-дигидроцитозин, 5-йодоцитозин, гидроксимочевину, йодурацил, 5-нитроцитозин, 5-бромурацил, 5-хлорурацил, 5-фторурацил и 5-йодурацил, 6-алкильные производные аденина и гуанина, 6-азапиримидины, 6-азоурацил, 6-азоцитозин, азацитозин, 6-азотимин, 6-тиогуанин, 7-метилгуанин, 7-метиладенин, 7-дезагуанин, 7-дезагуанозин, 7-дезааденин, 7-деза-8-азагуанин, 8-азагуанин, 8-азааденин, 8-галоген, 8-амино, 8-тиол, 8-тиоалкил и 8-гидроксилзамещенные аденины и гуанины; N4-этилцитозин, N-2 замещенные пурины, N-6 замещенные пурины, O-6 замещенные пурины, те, которые повышают стабильность образования дуплекса, универсальные нуклеиновые кислоты, гидрофобные нуклеиновые кислоты, смешанные нуклеиновые кислоты, нуклеиновые кислоты с увеличенным размером (size-expanded), фторированные нуклеиновые кислоты, трициклические пиримидины, феноксазинцитидин ([5,4-b][1,4]бензоксазин-2(3H)-он), фенотиазинцитидин

дин (1Н-пиримидо[5,4-*b*][1,4]бензотиазин-2(3Н)-он), "G-зажимы" (G-clamps), феноксазинцитидин (9-(2-аминоэтокси)-Н-пиримидо[5,4-*b*][1,4]бензоксазин-2(3Н)-он), карбазолцитидин (2Н-пиримидо[4,5-*b*]индол-2-он), пиридоиндолцитидин (Н-пиридо [3',2':4,5]пирроло[2,3-*d*]пиримидин-2-он), 5-фторурацил, 5-бром урацил, 5-хлорурацил, 5-иодурацил, гипоксантин, ксантин, 4-ацетилцитозин, 5-(карбоксигидроксиметил)урацил, 5-карбоксиметиламинометил-2-тиоуридин, 5-карбоксиметиламинометилурацил, дигидроурацил,  $\beta$ -D-галактозилквеозин, инозин, N6-изопентиладенин, 1-метилгуанин, 1-метилюозин, 2,2-диметилгуанин, 2-метиладенин, 2-метилгуанин, 3-метилцитозин, 5-метилцитозин, N6-аденин, 7-метилгуанин, 5-метиламинометилурацил, 5-метоксиаминометил-2-тиоурацил,  $\beta$ -О-маннозилквеозин, 5'-метоксикарбоксиметилурацил, 5-метоксиурацил, 2-метилтио-N6-изопентиладенин, урацил-5-оксиуксусную кислоту, wybutoxosine, псевдоурацил, квеозин, 2-тиоцитозин, 5-метил-2-тиоурацил, 2-тиоурацил, 4-тиоурацил, 5-метилурацил, урацил-5-метилловый эфир оксиуксусной кислоты, урацил-5-оксиуксусную кислоту, 5-метил-2-тиоурацил, 3-(3-амино-3-N-2-карбоксипропил)урацил, (аср3)w и 2,6-диаминопурин и те соединения, в которых пуриновое или пиримидиновое основание заменено гетероциклом.  $\alpha$ -Аналоги не встречающейся в природе пары оснований включают, без ограничения, dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO и их производные или аналоги.  $\beta$ -Аналоги не встречающейся в природе пары оснований включают, без ограничения, dTPT3, d5SICS и dFTPT3.

В некоторых вариантах осуществления не встречающиеся в природе нуклеотидные основания и не встречающиеся в природе пары оснований, раскрытые в настоящем изобретении, эффективно включаются и удлиняются с помощью природных полимераз. В некоторых вариантах осуществления не встречающиеся в природе нуклеотидные основания и не встречающиеся в природе пары оснований, раскрытые в настоящем изобретении, эффективно включаются и удлиняются с помощью модифицированных полимераз. Воздействие не встречающегося в природе нуклеинового основания или производного не встречающегося в природе нуклеинового основания на распознавание полимеразой оценивается в следующих примерах вариантах осуществления определением стационарной эффективности (например, константы скорости второго порядка  $k_{cat}/K_M$ ), с которой полимеразы синтезируют не встречающуюся в природе пару оснований, вводя не встречающийся в природе нуклеотид в положение напротив его комплементарного основания в матрице, и удлиняет с конца получившийся не встречающийся в природе праймер путем вставки следующего правильного природного нуклеотида. Также могут быть измерены соответствующие скорости синтеза и удлинения для ошибочных пар с природными нуклеотидами, чтобы определить точность воспроизведения. В некоторых вариантах осуществления не требуется модификация полимераз, чтобы улучшить скорости включения или удлинения. Варианты осуществления и примеры, описанные в настоящем изобретении, могут быть выполнены с любой известной полимеразой. Полимеразы включают встречающиеся в природе полимеразы и любые модифицированные их варианты, включая, но не ограничиваясь ими, мутанты, рекомбинанты, слитые модификации, генетические модификации, химические модификации, синтетические полимеразы и аналоги. Встречающиеся в природе полимеразы и модифицированные их варианты не ограничиваются полимеразой, которая сохраняет способность катализировать реакцию полимеризации. В некоторых случаях встречающиеся в природе и/или модифицированные их варианты сохраняют способность катализировать реакцию полимеризации. Мутантные полимеразы включают полимеразы, в которых одна или более аминокислот заменены другими аминокислотами (природными или не встречающимися в природе), и полимеразы, имеющие одну или более аминокислотных вставки или делеции. В некоторых вариантах осуществления полимеразы относятся к слитым белкам, содержащим по меньшей мере два участка, связанных друг с другом, например, когда один участок содержит пептид, который может катализировать полимеризацию нуклеотидов в цепь нуклеиновой кислоты, связан с другим участком, который содержит вторую часть, такую как фермент-репортер или процессивность-модифицирующий домен. Одним примером варианта осуществления такой полимеразы служит T7 ДНК-полимераза, которая содержит полимеризирующий нуклеиновую кислоту домен и тиоредоксин-связывающий домен, в котором связывание тиоредоксина усиливает процессивность полимеразы. При отсутствии связывания тиоредоксина, T7 ДНК-полимераза является дистрибутивной полимеразой с процессивностью, составляющей только от одного до нескольких оснований. ДНК-полимеразы включают, но не ограничиваются ими, бактериальные ДНК-полимеразы, эукариотические ДНК-полимеразы, архейные ДНК-полимеразы, вирусные ДНК-полимеразы и фаговые ДНК-полимеразы. Бактериальные ДНК-полимеразы включают ДНК-полимеразы I, II и III, IV и V из *E. coli*, фрагмент Кленова из ДНК-полимеразы *E. coli*, ДНК-полимеразу из *Clostridium stercorarium* (Cst), ДНК-полимеразу из *Clostridium thermocellum* (Cth) и ДНК-полимеразу из *Sulfolobus solfataricus* (Sso). Эукариотические ДНК-полимеразы включают ДНК-полимеразы  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\eta$ ,  $\zeta$ ,  $\sigma$ ,  $\lambda$ ,  $\mu$ ,  $\iota$  и  $\kappa$ , а также Rev1-полимеразу (терминальную дезоксицитидил-трансферазу) и терминальную дезоксинуклеотидтрансферазу (TdT). Вирусные ДНК-полимеразы включают T4 ДНК-полимеразу, phi-29 ДНК-полимеразы, GA-1, phi-29-подобную ДНК-полимеразу, PZA ДНК-полимеразу, phi-15 ДНК-полимеразу, Cpl ДНК-полимеразу, CP7 ДНК-полимеразу, T7 ДНК-полимеразу и T4 полимеразу. Архейные ДНК-полимеразы включают термостабильные и/или термофильные ДНК-полимеразы, такие как ДНК-полимеразы, выде-



ленные из *Thermus aquaticus* (Tag) ДНК-полимеразы, ДНК-полимераза *Thermus filiformis* (Tfi), ДНК-полимераза *Thermococcus zilligi* (Tzi), ДНК-полимераза *Thermus thermophilus* (Tth), ДНК-полимераза *Thermus flavus* (Tfl), ДНК-полимераза *Pyrococcus woesei* (Pwo), ДНК-полимераза *Pyrococcus furiosus* (Pfu) и ДНК-полимераза Turbo Pfu, ДНК-полимераза *Thermococcus litoralis* (Tli), GB-D-полимераза *Pyrococcus* sp., ДНК-полимераза *Thermotoga maritima* (Tma), ДНК-полимераза *Bacillus stearothermophilus* (Bst), ДНК-полимераза *Pyrococcus kodakaraensis* (KOD), ДНК-полимераза Pfx, *Thermococcus* sp. JDF-3 (JDF-3) ДНК-полимераза, ДНК-полимераза *Thermococcus gorgonarius* ((Tgo), ДНК-полимераза *Thermococcus acidophilum*; ДНК-полимераза *Sulfolobus acidocaldarius*; *Thermococcus* sp. 9° N-7 ДНК-полимераза; ДНК-полимераза *Pyrodictium occultum*; ДНК-полимераза *Methanococcus voltae*; ДНК-полимераза *Methanococcus thermoautotrophicum*; ДНК-полимераза *Methanococcus jannaschii*; штамм *Desulfurococcus* ТОК ДНК-полимераза (D. Ток Pol); ДНК-полимераза *Pyrococcus abyssi*; ДНК-полимераза *Pyrococcus horikoshii*; ДНК-полимераза *Pyrococcus islandicum*; ДНК-полимераза *Thermococcus fumicolans*; ДНК-полимераза *Aeropyrum pernix*; и гетеродимерная ДНК-полимераза DP1/DP2. РНК-полимеразы включают, но не ограничиваются ими, вирусные РНК-полимеразы, такие как Т7 РНК-полимераза, Т3-полимераза, SP6 полимераза и K11-полимераза; эукариотические РНК-полимеразы, такие как РНК-полимераза I, РНК-полимераза II, РНК-полимераза III, РНК-полимераза IV и РНК-полимераза V; и архейная РНК-полимераза.

В некоторых вариантах осуществления полимеразы имеет специфичность для не встречающегося в природе нуклеотида, содержащего  $\alpha$ - или  $\beta$ -аналог нуклеинового основания, которая составляет по меньшей мере около 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 97, 98, 99, 99,5, 99,99% специфичности полимеразы к природному нуклеотиду. В некоторых вариантах осуществления полимеразы имеет специфичность для не встречающегося в природе нуклеотида, содержащего  $\alpha$ - или  $\beta$ -аналог нуклеинового основания и модифицированный сахар, которая составляет по меньшей мере около 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 97, 98, 99, 99,5, 99,99% специфичности полимеразы к природному нуклеотиду и/или не встречающемуся в природе нуклеотиду без модифицированного сахара. В некоторых вариантах осуществления полимеразы имеет специфичность для не встречающегося в природе нуклеотида, содержащего дериватизированный с помощью линкера  $\alpha$ - или  $\beta$ -аналог нуклеинового основания, которая составляет по меньшей мере около 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 97, 98, 99, 99,5, 99,99% специфичности полимеразы к природному нуклеотиду и/или не встречающемуся в природе нуклеотиду без линкера. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dTPT3. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dMMS. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dDMS. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dBrMS. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой IMS. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dFEMS. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой MMS<sup>PCO</sup>. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dMMS<sup>PA</sup>. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dFTPT3. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dTPT<sup>PA</sup>. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание представляет собой dTPT3<sup>CO</sup>. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14a или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14b или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14c или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14d или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14e или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\alpha$ 14f или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\beta$ 8a или его производного или аналога. В некоторых вариантах осуществления не встречающееся в природе нуклеиновое основание содержит формулу  $\beta$ 8b или его производного или аналога.

Полимеразы могут быть охарактеризованы по их точности воспроизведения при использовании с конкретным природным и/или не встречающимся в природе нуклеотидом или коллекциями природных и/или не встречающихся в природе нуклеотидов, причем не встречающийся в природе нуклеотид содержит  $\alpha$ - или  $\beta$ -аналог нуклеинового основания, раскрытого в настоящем изобретении. В различных вариантах осуществления точность воспроизведения в целом относится к тому, насколько аккуратно полимеразы включают правильные нуклеотиды в растущий олигонуклеотид при изготовлении копии олигонуклеотидной матрицы. Точность работы полимеразы может быть измерена как отношение числа правиль-

ных включений нуклеотидов к ошибочным включениям в случае природных и не встречающихся в природе нуклеотидов, когда природные и не встречающиеся в природе нуклеотиды присутствуют, например, в равных концентрациях, чтобы конкурировать за синтез цепи в том же самом месте в двойном комплексе полимеразы-цепь-матричная нуклеиновая кислота. Точность ДНК-полимеразы может быть рассчитана как отношение ( $k_{cat}/K_M$ ) для природного и не встречающегося в природе нуклеотида к ( $k_{cat}/K_M$ ) для ошибочного природного и не встречающегося в природе нуклеотида; где  $k_{cat}$  и  $K_M$  являются параметрами Михаэлиса-Ментен в стационарной кинетике ферментативных реакций. В некоторых вариантах осуществления значение точности воспроизведения для полимеразы составляет по меньшей мере около 100, 1000 и 10000, 100000 или  $1 \times 10^6$ , с корректирующей активностью или без нее. В некоторых вариантах осуществления значение точности воспроизведения для полимеразы составляет по меньшей мере около 100, 1000 и 10000, 100000 или  $1 \times 10^6$  для включения не встречающегося в природе нуклеотида. В некоторых вариантах осуществления не встречающийся в природе нуклеотид представляет собой dTPT3TP или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде представляет собой dNAM или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид представляет собой dNaMTP или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде представляет собой dTPT3 или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14a$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14b$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14c$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14d$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14e$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8a$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14f$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8b$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14a$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8b$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14b$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8b$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14c$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8b$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14d$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8e$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14e$  или его производное. В некоторых вариантах осуществления не встречающийся в природе нуклеотид содержит  $\beta 8f$  или его производное, и его соответствующее нуклеиновое основание на матричном олигонуклеотиде содержит  $\alpha 14f$  или его производное.

Не встречающиеся в природе пары оснований, приведенные в качестве примеров в данном описании, в некоторых вариантах осуществления, синтезируются/амплифицируются с эффективностью и точностью воспроизведения, как у природной пары оснований. Не встречающиеся в природе пары оснований включают, в различных вариантах осуществления, любой  $\alpha$ -аналог нуклеинового основания или его производное и/или любой  $\beta$ -аналог нуклеинового основания или его производное. Примеры  $\beta$ -аналогов включают dTPT3, d5SICS, dFTPT3 и их производные или аналоги. Примеры  $\alpha$ -аналогов включают dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO и их производные или аналоги. В некоторых вариантах осуществления пара оснований, не встречающаяся в природе, эффективно амплифицируется во множестве различных видов последовательностей, включая GC- и AT-богатые последовательности, рандомизированные последовательности, а также последовательности, включающие несколько не встречающихся в природе пар нуклеиновых оснований, с точностью воспроизведения, составляющей более чем 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,1, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,85, 99,9% или выше, в пересчете на дубликацию. Например, не встречающаяся в природе пара нуклеиновых оснований, содержащая одно или более не встречающихся в природе нуклеиновых оснований, имеет эффективность и/или точность синтеза, которая по меньшей мере на 60, 65, 70, 75, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 или 99% сходна с эффективностью и/или точностью амплификации природной пары оснований. В качестве другого примера, не встречающаяся в природе пара нук-

линовых оснований, содержащая одно или более не встречающихся в природе нуклеиновых оснований, имеет эффективность и/или точность синтеза, которая максимально на 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 или 0,5% менее эффективна и/или точна, чем у природной пары оснований. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований транскрибируется с хорошей эффективностью и селективностью в обеих цепях (например, dX должно служить матрицей для УТР-вставки, и dY должно служить матрицей для ХТР-вставки). В некоторых вариантах осуществления по отношению к скорости, с которой транскрибируется полностью природная последовательность, включение не встречающегося в природе нуклеотида не снижает скорость полноразмерной транскрипции. В некоторых вариантах осуществления по отношению к скорости, с которой транскрибируется полностью природная последовательность, включение не встречающегося в природе нуклеотида снижает скорость полноразмерной транскрипции менее чем в 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35 или 40 раз. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает dTPT3 или его производное или аналог и dNaM или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает dTPT3 или его производное или аналог и dNaM или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает dTPT3 или его производное или аналог и dNaM или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает dTPT3 или его производное или аналог и dNaM или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14a или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14b или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14c или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14d или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14e или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8a или его производное или аналог и  $\alpha$ 14f или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14a или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14b или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14c или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14d или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14e или его производное или аналог. В некоторых вариантах осуществления не встречающаяся в природе пара оснований включает  $\beta$ 8b или его производное или аналог и  $\alpha$ 14f или его производное или аналог.

В различных вариантах осуществления настоящего изобретения далее обеспечиваются не встречающиеся в природе пары оснований, включающие одно или более не встречающихся в природе нуклеиновых оснований (например,  $\alpha$ -нуклеиновое основание,  $\beta$ -нуклеиновое основание или  $\alpha$ -нуклеиновое основание и  $\beta$ -нуклеиновое основание), где одно или два нуклеиновых оснований содержат линкер. Линкер содержит реакционноспособный центр. Примеры реакционноспособных центров включают, но не ограничиваются ими, алкил, алкенил, алкинил, фенил, бензил, галоген, гидроксил, карбонил, альдегид, галоформил, сложный карбонатный эфир, карбоксилат, карбоксил, сложный эфир, метокси, гидроперокси, перокси, простой эфир, полуацеталь, полукеталь, ацеталь, кеталь, ортоэфир, метилendioкси, ортокربонатный эфир, карбоксамид, первичный амин, вторичный амин, имид, азид, азо, цианат, изоцианат, нитрат, нитрил, изонитрил, нитрозоокси, нитро, нитрозо, пиридил, сульфгидрил, сульфид, дисульфид, сульфинил, сульфо, тиоцианат, изотиоцианат, карбоностирил, фосфино, фосфоно, фосфат, бороно, боронат, бороно, боринат и их комбинации. Примером дериватизированного с помощью линкера нуклеинового основания является  $\text{TRPT3}^{\text{R}}$ , показанное на фиг. 2, где верхний индекс R обозначает линкер. В некоторых вариантах осуществления линкер модифицирован защитной группой, например  $\text{TRPT3}^{\text{PA}}$ , где линкер представляет собой защищенный пропаргильный линкер.

В некоторых вариантах осуществления аналог нуклеинового основания, обеспеченный в настоящем изобретении, содержит аминокфункциональный линкер или защищенный аминокфункциональный линкер (например,  $\text{dX}^{\text{PA}}$ ). В некоторых вариантах осуществления аминокфункциональный линкер представляет собой 3-аминопропин-1-ил. В некоторых вариантах осуществления аналог нуклеинового основания, обеспеченный в настоящем изобретении, включает алкин-азидный эфирный линкер для дериватизации с

помощью реакции клик-химии или защищенный алкин-азидный эфирный линкер для дериватизации с помощью реакции клик-химии. В некоторых вариантах осуществления алкин-азидный эфирный линкер представляет собой 4-оксагепта-1,6-диин-1-ил (4-oxahepta-1,6-diyn-1-yl). В некоторых вариантах осуществления аналог нуклеинового основания, обеспеченный в настоящем изобретении, включает алкин-азид-триметиленовый линкер для дериватизации с помощью реакции клик-химии или защищенный алкин-азид-триметиленовый линкер для дериватизации с помощью реакции клик-химии. В некоторых вариантах осуществления алкин-азид триметиленовый линкер представляет собой гепта-1,6-диин-1-ил. В некоторых вариантах осуществления X представляет собой  $\beta$ -аналог нуклеозида, имеющий любую из формул из фиг. 9, 12, 13 и 16. В некоторых вариантах осуществления X представляет собой ICS, PICS, MICS, 4MICS, 5MICS, NICS, ONICS, SICS, SNICS, 5SICS, 4SICS, 7OFP, 7OTP, TPT2, TPT3 или FTPT3. В некоторых вариантах осуществления X представляет собой  $\alpha$ -аналог нуклеозида, имеющий любую из формул из фиг. 8, 10, 11 и 15. В некоторых вариантах осуществления X представляет собой FIMO, MIMO, FEMO, PrMO, EMO, MEMO, IMO, MMO2, DMO, NMO, 5FM, 2OMe, TMO, FDMO, VMO, ZMO, CIMO, TfMO, CNMO, MMS, DMS, BrMS, IMS, FEMS, NAM или QMO.

В некоторых вариантах осуществления линкер представляет собой пропинильный линкер, такой как те, которые используются с природными нуклеотидами. Такие линкеры содержат пропаргиламины с амином, выступающим в качестве реакционноспособного сайта для присоединения других функциональных групп.

В различных вариантах осуществления дериватизированное с помощью линкера нуклеиновое основание содержит спейсер. Примером спейсера является ацетамидогексанамид. Спейсер может быть гидрофильным. Спейсер может соединять линкер с функциональной группой. Спейсеры включают, но не ограничиваются ими, спейсер C3 (3-углеродный спейсер), спейсер C6 (6-углеродный спейсер), фоторасщепляемый спейсер, гександиоловый спейсер, спейсер 9 (триэтиленгликольный спейсер), спейсер C12 (12-углеродный спейсер), спейсер 18 (18-атомный гексаэтиленгликольный спейсер) и 1',2'-дидезоксирибозный спейсер.

Не встречающаяся в природе пара нуклеиновых оснований, содержащая одно или два дериватизированных с помощью линкера нуклеиновых оснований, в некоторых случаях амплифицируется с эффективностью и точностью воспроизведения, которая аналогична таковой природной пары оснований или не встречающейся в природе пары оснований без дериватизации линкером. Например, не встречающаяся в природе пара нуклеиновых оснований, содержащая одно или два дериватизированных с помощью линкера не встречающихся в природе нуклеиновых оснований, имеет эффективность и/или точность синтеза, которая по меньшей мере на 60, 65, 70, 75, 80 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 или 99% аналогична эффективности и/или точности синтеза природной пары оснований или не дериватизированной линкером пары оснований, не встречающихся в природе. В качестве другого примера не встречающаяся в природе пара нуклеиновых оснований, содержащая одно или два дериватизированных с помощью линкера не встречающихся в природе нуклеиновых оснований, имеет эффективность и/или точность синтеза, которая максимально на 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 или 0,5% менее эффективна и/или точна, чем таковая у природной пары оснований или у не встречающейся в природе пары оснований, не дериватизированных с помощью линкера. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dTPT3<sup>PA</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dTPT3<sup>CO</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dMMS<sup>CO</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dMMS<sup>PA</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dNaM<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dMMO2<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dDMO<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает d5SICS<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dMMS<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dDMS<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dFEMS<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dBrMS<sup>R</sup>. В некоторых вариантах осуществления не встречающаяся в природе пара нуклеиновых оснований включает dIMS<sup>R</sup>.

В некоторых вариантах осуществления дериватизированное с помощью линкера нуклеиновое основание, не встречающееся в природе, имеет повышенную эффективность инсерции (вставки) в процессе синтеза олигонуклеотидов, по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит линкер. В некоторых вариантах осуществления дериватизированное с помощью линкера нуклеиновое основание, не встречающееся в природе, имеет пониженную эффективность вставки в процессе синтеза олигонуклеотидов по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит линкер. В некоторых случаях дериватизированное с помо-

шью линкера нуклеиновое основание, не встречающееся в природе, имеет приблизительно такую же эффективность вставки в процессе синтеза олигонуклеотидов, по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит линкер. В некоторых вариантах осуществления дериватизированное с помощью защищенного линкера нуклеиновое основание, не встречающееся в природе, имеет повышенную эффективность вставки в процессе синтеза олигонуклеотидов по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит защищенный линкер. В некоторых вариантах осуществления дериватизированное с помощью защищенного линкера нуклеиновое основание, не встречающееся в природе, имеет пониженную эффективность вставки в процессе синтеза олигонуклеотидов по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит защищенный линкер. В некоторых случаях дериватизированное с помощью защищенного линкера нуклеиновое основание, не встречающееся в природе, имеет приблизительно такую же эффективность в процессе синтеза олигонуклеотидов по сравнению с тем же не встречающимся в природе нуклеотидным основанием, которое не содержит защищенный линкер.

Примеры способов анализа для не встречающихся в природе пар оснований эффективности синтеза (вставка не встречающегося в природе нуклеинового основания напротив его партнера в матрице) и удлинения (продолжение удлинения праймера) обеспечиваются в данном описании. Одно или оба нуклеиновых основания в не встречающейся в природе паре оснований, в различных вариантах осуществления, может быть не встречающимся в природе нуклеотидным основанием, дериватизированным с помощью линкера. Один способ использует анализ предстационарной кинетики. Анализ основан на определении при фиксированном наборе условий количества праймера (например, 23-мер), который продлевается добавлением не встречающегося в природе нуклеотида в положении, противоположном комплементарному ему нуклеотиду в матрице (например, 45-мер) полимеразой (например, фрагмент Кленова ДНК-полимеразы I *E.coli*). В данном анализе эффективность синтеза не встречающейся в природе пары оснований характеризуется измерением процента включения (%inc) при заданной концентрации не встречающегося в природе и следующего правильного трифосфата, например, с использованием соотношения, такого как  $[24\text{-мер}+25\text{-мер}]/[23\text{-мер}+24\text{-мер}+25\text{-мер}]$ . В данном анализе эффективность удлинения характеризуется величиной процентного удлинения (%ext) при заданной концентрации следующего правильного нуклеотида и насыщающих концентрациях не встречающегося в природе нуклеотида, например, с использованием соотношения  $[25\text{-мер}]/[24\text{-мер}+25\text{-мер}]$ . Результаты из иллюстративного анализа предстационарной кинетики приведены в табл. 1, где не встречающийся в природе трифосфат представляет собой 5SICS, FPT1, TPT1, TPT2, TPT3, FTPT3, TPT3<sup>PA</sup> или 5SICS<sup>PA</sup>. В некоторых вариантах осуществления процент включения нуклеинового основания, не встречающегося в природе, составляет по меньшей мере 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 или 99%. В некоторых вариантах осуществления процент удлинения за счет следующего правильного нуклеотида после вставки нуклеинового основания, не встречающегося в природе, составляет по меньшей мере 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90% или более. В некоторых вариантах осуществления эффективность синтеза увеличивается дериватизированием нуклеинового основания, не встречающегося в природе. Например, добавлением линкера, защищенного линкера и/или линкера, конъюгированного с молекулой карго. В качестве другого примера дериватизация включает замены, добавления или удаления атомов. В некоторых вариантах осуществления процентное удлинение увеличивается дериватизированием нуклеинового основания, не встречающегося в природе. Дериватизация не встречающегося в природе нуклеинового основания в некоторых случаях увеличивает по меньшей мере на 1-2 порядка величину эффективности вставки нуклеотида, комплементарного парному основанию в матрице. Такое увеличение эффективности может быть связано с увеличением  $k_{cat}$  и снижением  $K_M$ .

Таблица 1

Предстационарная кинетика		
dXTP	Включение <sup>a</sup> %	Удлинение <sup>b</sup> %
5SICS	57,0 ± 0,2	15,1 ± 1,1
FPT1	7,2 ± 0,2	32,0 ± 1,5
TPT1	28,7 ± 0,5	8,8 ± 0,2

TPT2	65,7 ± 0,5	34,5 ± 0,5
TPT3	72,3 ± 0,5	49,8 ± 1,3
FTPT3	66,3 ± 0,5	33,8 ± 0,2
TPT3 <sup>PA</sup>	68,3 ± 0,4	31,5 ± 0,7
5SICS <sup>PA</sup>	7,0 ± 0,2	5,5 ± 0,1

<sup>a</sup> Условия анализа включения: 40 нМ не встречающегося в природе трифосфата, 2 мкМ дЦТФ 10 с.

<sup>b</sup> Условия анализа удлинения: 10 мкМ не встречающегося в природе трифосфата, 2 мкМ дЦТФ 10 с. dXTP спарены с dNaM.

Далее настоящее изобретение относится к способам оценки репликации. В одном способе дуплекс матричной нуклеиновой кислоты, содержащий не встречающуюся в природе пару оснований (например, dTPT3-dNaM или их аналоги), амплифицируется с помощью ПЦР. В одном примере набор реакций ПЦР составляет 48 циклов с полимеразой OneTaq. В другом примере набор реакций ПЦР использует 20 циклов амплификации с Taq-полимеразой, не имеющей экзонуклеазной активности. Эффективность определяется мониторингом уровня амплификации. Точность воспроизведения, обычно определяемая как удлинение после не встречающейся в природе пары оснований при дупликации, определяется из процента амплифицированной ДНК, которая сохраняет не встречающуюся в природе пару оснований. Процент амплифицированной ДНК, которая сохраняет не встречающуюся в природе пару оснований, может быть определен из относительных интенсивностей пиков хроматограммы секвенирования. В некоторых вариантах осуществления точность репликации не встречающейся в природе пары оснований составляет по меньшей мере 98, 98,1, 98,2, 98,3, 98,4, 98,5, 98,6, 98,7, 98,8, 98,9, 99, 99,1, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,8, 99,9, 99,91, 99,92, 99,93, 99,94, 99,95, 99,96, 99,97, 99,98 или 99,99%. Репликация не встречающейся в природе пары оснований может протекать с малым смещением последовательности или практически без смещения последовательности, причем малое смещение последовательности указывает на то, что неблагоприятная последовательность уменьшает точность воспроизведения менее чем на 1%. Примеры точностей воспроизведения описаны в примере 1 и приведены в табл. 4, 5 и 6.

В различных вариантах осуществления настоящего изобретения далее обеспечиваются олигонуклеотиды, включая одноцепочечную и двухцепочечную (например, дуплекс) ДНК и/или РНК, содержащие одно или более не встречающихся в природе нуклеиновых оснований, описанных в настоящем изобретении (например, любое  $\alpha$ -нуклеиновое основание или его аналог или производное и/или любое  $\beta$ -нуклеиновое основание или его аналог или производное). Нуклеиновое основание может быть любым  $\alpha$ -нуклеотидным основанием или  $\beta$ -нуклеотидным основанием, описанным в настоящем изобретении, включая приведенные на фиг. 2, 8, 9, 10, 11, 12, 13, 15 и 16. Двухцепочечный олигонуклеотид включает дуплекс ДНК-ДНК, гибридный дуплекс ДНК-РНК и дуплекс РНК-РНК. В некоторых вариантах осуществления олигонуклеотид содержит дериватизированное с помощью линкера нуклеиновое основание.

В некоторых вариантах осуществления олигонуклеотид содержит 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50 или более не встречающихся в природе нуклеиновых оснований. В некоторых вариантах осуществления процент не встречающихся в природе нуклеиновых оснований в олигонуклеотиде составляет от около 0 до около 1%, от около 0 до около 2%, от около 0 до около 3%, от около 0 до около 4%, от около 0 до около 5%, от около 0 до около 10%, от около 1 до около 10%, от около 1 до около 15%, от около 1 и около 20%, от около 5 до около 10%, от около 5 до около 20%, от около 10 до около 30%, от около 1 до около 50% или от около 1 до около 100%.

Примеры олигонуклеотидов, содержащих одно или более не встречающихся в природе нуклеиновых оснований, включают, но не ограничиваются ими, ДНК-аптамеры и РНК-аптамеры. ДНК- и РНК-аптамеры включают, но не ограничиваются ими, праймеры и молекулярные маяки. ДНК-аптамер может включать в себя штрих-код.

В некоторых вариантах осуществления олигонуклеотид содержит dTPT3 или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает d5SICS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dNaM или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dMMS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dDMS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dFEMS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dBrMS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает dIMS или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\beta$ 8a или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\beta$ 8b или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha$ 14a или

его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha 14b$  или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha 14c$  или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha 14d$  или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha 14e$  или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает  $\alpha 14f$  или его производное или аналог.

В некоторых вариантах осуществления олигонуклеотид содержит dTPT3 или его производное или аналог и dNaM или его производное или аналог. В некоторых вариантах осуществления олигонуклеотид включает пару оснований dTPT3-dNaM. В некоторых вариантах осуществления олигонуклеотид содержит одну или более пар оснований, выбранных из

dTPT3-dFEMO, dTPT3-dFIMO, dTPT3-dIMO, dFTPT3-dNaM, dFTPT3-

dFEMO, dFTPT3-dFIMO и dFTPT3-dIMO.

В некоторых вариантах осуществления олигонуклеотид включает одну или более пар оснований, выбранных из

dTPT3-MMS,

dTPT3-DMS, dTPT3-FEMS, dTPT3-BrMS, dTPT3-IMS, dTPT3-dDMN, dTPT3-d4OMe, dTPT3-dIQ, dTPT3-d2MN, dTPT3-d3OMe, dTPT3-dQL, dTPT3-d2Np, dTPT3-dDM4, dTPT3-dDM, dTPT3-dBEN, dTPT3-d3FB, dTPT3-dMM1, dTPT3-dMMO1, dTPT3-dDM2, dTPT3-dDM5, dTPT3-d2Py, dTPT3-d5MPy, dTPT3-dEPy, dTPT3-d3MPy, dTPT3-d34DMPy, dTPT3-d45DMPy, dTPT3-d4MPy, dTPT3-d35DMPy, dTPT3-dBP, dTPT3-dBTp, dTPT3-dBF, dTPT3-dIN, dTPT3-dTp, dTPT3-dBTz, dTPT3-dMTp, dTPT3-dAM, dTPT3-dMAN, dTPT3-dMMAN, dTPT3-dADM, dTPT3-dMMAN, dTPT3-dTOK588, dTPT3-dTOK576, dTPT3-dTOK587, dTPT3-dTOK586, dTPT3-dTOK580, dTPT3-dPhMO, dTPT3-dPyMO1, dTPT3-PyMO2, dTPT3-dPMO1, dTPT3-dPMO2, dTPT3-dPMO3, dTPT3-dFuMO1, dTPT3-dFuMO2, dTPT3-TpMO1, dTPT3-dTpMO2, dTPT3-dFIMO, dTPT3-dIMO, dTPT3-dMIMO, dTPT3-dMEMO, dTPT3-dFEMO, dTPT3-dPrMO, dTPT3-dMMO2, dTPT3-d2OMe, dTPT3-dDMO, dTPT3-dTMO, dTPT3-dNMO, dTPT3-dNOPy, dTPT3-d5FM, dTPT3-dNAM, dTPT3-dAMO1, dTPT3-dAPy, dTPT3-dAMO2, dTPT3-dMAPy, dTPT3-dAMO3, dTPT3-dDMAPy, dTPT3-dFDMO, dTPT3-dVMO, dTPT3-dQMO, dTPT3-dZMO, dTPT3-dCIMO, dTPT3-dTfMO, dTPT3-CNMO, d7AI-dMMS, dM7AI-dMMS, dImPy-dMMS, dP7AI-dMMS, dPPP-dMMS, d8Q-dMMS, dICS-dMMS, dPICS-dMMS, dMICS-dMMS, d4MICS-dMMS, d5MICS-dMMS, dNICS-dMMS, dONICS-dMMS, d7OFP-dMMS, d7OTP-dMMS, d4OTP-dMMS, dPYR-dMMS, d4MP-dMMS, d3MP-dMMS, dPPYR-dMMS, dMOP-dMMS, d4MOP-dMMS, dSICS-dMMS, dSNICS-dMMS, d5SICS-dMMS, d4SICS-dMMS, dTPT1-dMMS, dTPT2-dMMS, dFPT1-dMMS, dFTPT3-dMMS, d7AI-dDMS, dM7AI-dDMS, dImPy-dDMS, dP7AI-dDMS, dPPP-dDMS, d8Q-dDMS, dICS-dDMS, dPICS-dDMS, dMICS-dDMS, d4MICS-dDMS, d5MICS-dDMS, dNICS-dDMS, dONICS-dDMS, d7OFP-dDMS, d7OTP-dDMS, d4OTP-dDMS, dPYR-dDMS, d4MP-dDMS,

d3MP-dDMS, dPPYR-dDMS, dMOP-dDMS, d4MOP-dDMS, dSICS-dDMS, dSNICS-dDMS, d5SICS-dDMS, d4SICS-dDMS, dTPT1-dDMS, dTPT2-dDMS, dFPT1-dDMS, dFTPT3-dDMS, d7AI-dFEMS, dM7AI-dFEMS, dImPy-dFEMS, dP7AI-dFEMS, dPPP-dFEMS, d8Q-dFEMS, dICS-dFEMS, dPICS-dFEMS, dMICS-dFEMS, d4MICS-dFEMS, d5MICS-dFEMS, dNICS-dFEMS, dONICS-dFEMS, d7OFP-dFEMS, d7OTP-dFEMS, d4OTP-dFEMS, dPYR-dFEMS, d4MP-dFEMS, d3MP-dFEMS, dPPYR-dFEMS, dMOP-dFEMS, d4MOP-dFEMS, dSICS-dFEMS, dSNICS-dFEMS, d5SICS-dFEMS, d4SICS-dFEMS, dTPT1-dFEMS, dTPT2-dFEMS, dFPT1-dFEMS, dFTPT3-dFEMS, d7AI-dBrMS, dM7AI-dBrMS, dImPy-dBrMS, dP7AI-dBrMS, dPPP-dBrMS, d8Q-dBrMS, dICS-dBrMS, dPICS-dBrMS, dMICS-dBrMS, d4MICS-dBrMS, d5MICS-dBrMS, dNICS-dBrMS, dONICS-dBrMS, d7OFP-dBrMS, d7OTP-dBrMS, d4OTP-dBrMS, dPYR-dBrMS, d4MP-dBrMS, d3MP-dBrMS, dPPYR-dBrMS, dMOP-dBrMS, d4MOP-dBrMS, dSICS-dBrMS, dSNICS-dBrMS, d5SICS-dBrMS, d4SICS-dBrMS, dTPT1-dBrMS, dTPT2-dBrMS, dFPT1-dBrMS, dFTPT3-dBrMS, d7AI-dIMS, dM7AI-dIMS, dImPy-dIMS, dP7AI-dIMS, dPPP-dIMS, d8Q-dIMS, dICS-dIMS, dPICS-dIMS, dMICS-dIMS, d4MICS-dIMS, d5MICS-dIMS, dNICS-dIMS, dONICS-dIMS, d7OFP-dIMS, d7OTP-dIMS, d4OTP-dIMS, dPYR-dIMS, d4MP-dIMS, d3MP-dIMS, dPPYR-dIMS, dMOP-dIMS, d4MOP-dIMS, dSICS-dIMS, dSNICS-dIMS, d5SICS-dIMS, d4SICS-dIMS, dTPT1-dIMS, dTPT2-dIMS, dFPT1-dIMS, dFTPT3-dIMS;

где одно или два не встречающихся в природе нуклеиновых оснований не встречающейся в природе пары оснований могут быть дериватизированы с помощью линкера.

Олигонуклеотид, содержащий не встречающееся в природе нуклеиновое основание, раскрытое в настоящем изобретении, может дополнительно содержать одно или более дополнительных не встречающихся в природе оснований, включая, но не ограничиваясь ими, 2-аминоаденин-9-ил, 2-аминоаденин, 2-Ф-аденин, 2-тиоурацил, 2-тиотимин, 2-тиоцитозин, 2-пропил и алкильные производные аденина и гуанина, 2-аминоаденин, 2-аминопропил-аденин, 2-аминопиридин, 2-пиридон, 2'-дезоксигуанидин, 2-амино-2'-дезоксиаденозин, 3-деазагуанин, 3-деазааденин, 4-тиоурацил, 4-тиотимин, урацил-5-ил, гипоксантин-9-ил (I), 5-метил-цитозин, 5-гидроксиметилцитозин, ксантин, гипоксантин, 5-бром и 5-трифторметил-урацилы и цитозины; 5-галоурацил, 5-галоцитозин, 5-пропинил-урацил, 5-пропинил-цитозин, 5-урацил, 5-замещенный, 5-галоген, 5-замещенные пиримидины, 5-гидроксицитозин, 5-бромцитозин, 5-бромурацил, 5-хлорцитозин, хлорированный цитозин, циклоцитозин, цитозинарабинозид, 5-фторцитозин, фторпиримидин, фторурацил, 5,6-дигидроцитозин, 5-иодоцитозин, гидроксимочевину, иодоурацил, 5-нитроцитозин, 5-бромурацил, 5-хлорурацил, 5-фторурацил и 5-иодоурацил, 6-алкильные производные аденина и гуанина, 6-азапиримидины, 6-азоурацил, 6-азоцитозин, азацитозин, 6-азотимин, 6-тиогуанин, 7-метилгуанин, 7-метиладенин, 7-деазагуанин, 7-деазагуанозин, 7-деазааденин, 7-деаза-8-азагуанин, 8-азагуанин, 8-азааденин, 8-галоген, 8-амино, 8-тиол, 8-тиоалкил и 8-гидроксилзамещенные аденины и гуанины; N4-этилцитозин, N-2 замещенные пурины, N-6 замещенные пурины, O-6 замещенные пурины, те, которые повышают стабильность образования дуплекса, универсальные нуклеиновые кислоты, гидрофобные нуклеиновые кислоты, смешанные нуклеиновые кислоты, нуклеиновые кислоты с увеличенным размером (size-expanded), фторированные нуклеиновые кислоты, трициклические пиримидины, феноксазинцитидин ([5,4-b][1,4]бензоксазин-2(3H)-он), фенотиазинцитидин (1H-пиримидо[5,4-b][1,4]бензотиазин-2(3H)-он), "G-зажимы" (G-clamps), феноксазинцитидин (9-(2-аминоэтокси)-H-пиримидо[5,4-b][1,4]бензоксазин-2(3H)-он), карбазолцитидин (2H-пиримидо[4,5-b]индол-2-он), пиридоиндолцитидин (H-пиридо[3',2':4,5]пирроло[2,3-d]пиримидин-2-он), 5-фторурацил, 5-бромурацил, 5-хлорурацил, 5-иодоурацил, гипоксантин, ксантин, 4-ацетилцитозин, 5- (карбоксигидроксиметил)урацил, 5-карбоксиметиламинометил-2-тиоуридин, 5-карбоксиметиламинометилурацил, дигидроурацил, β-D-галактозилквеозин, инозин, N6-изопентиладенин, 1-метилгуанин, 1-метилюнозин, 2,2-диметилгуанин, 2-метиладенин, 2-метилгуанин, 3-метилцитозин, 5-метилцитозин, N6-аденин, 7-метилгуанин, 5-метиламинометилурацил, 5-метоксиаминометил-2-тиоурацил, β-O-маннозилквеозин, 5'-метоксикарбоксиметилурацил, 5-метоксиурацил, 2-метилтио-N6-изопентиладенин, урацил-5-оксиуксусную кислоту, wybutoxosine, псевдоурацил, квеозин, 2-тиоцитозин, 5-метил-2-тиоурацил, 2-тиоурацил, 4-тиоурацил, 5-метилюрацил, урацил-5-метилэтер оксиуксусной кислоты, урацил-5- оксиуксусную кислоту, 5-метил-2-тиоурацил, 3-(3-амино-3-N-2-карбоксипропил)урацил, (аср3)w и 2,6-диаминопурин и те соединения, в которых пуриновое или пиримидиновое основание заменено гетероциклом.

Олигонуклеотид, содержащий не встречающееся в природе нуклеиновое основание, раскрытое в



настоящем изобретении, может дополнительно содержать не встречающийся в природе сахарный остаток, включая, но не ограничиваясь ими, модификацию в 2' положении: OH; замещенный низший алкил, алкирил, аралкил, О-алкирил либо О-аралкил, SH, SCH<sub>3</sub>, OCN, Cl, Br, CN, CF<sub>3</sub>, OCF<sub>3</sub>, SOCH<sub>3</sub>, SO<sub>2</sub> CH<sub>3</sub>, ONO<sub>2</sub>, NO<sub>2</sub>, N<sub>3</sub>, NH<sub>2</sub>F; О-алкил, S-алкил, N-алкил; О-алкенил, S-алкенил, N-алкенил; О-алкинил, С-алкинил, N-алкинил; О-алкил-О-алкил, 2'-F, 2'-OCH<sub>3</sub>, 2'-O(CH<sub>2</sub>)<sub>2</sub>OCH<sub>3</sub>, где алкильная, алкенильная и алкинильная группа может быть замещена или замещена C<sub>1</sub>-C<sub>10</sub>, алкилом, C<sub>2</sub>-C<sub>10</sub>-алкенилом, C<sub>2</sub>-C<sub>10</sub>-алкинилом, -O[(CH<sub>2</sub>)<sub>n</sub>O]<sub>m</sub>CH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>OCH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>NH<sub>2</sub>, -O(CH<sub>2</sub>)<sub>n</sub> CH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>-ONH<sub>2</sub> и -O(CH<sub>2</sub>)<sub>n</sub>ON[(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>]<sub>2</sub>, где n и m составляют от 1 до около 10; и/или модификацию в 5' положении: 5'-винил, 5'-метил (R или S), модификацию в 4' положении: 4'-S, гетероциклоалкил, гетероциклоалкирил, аминокалкиламино, полиалкиламино, замещенный силл, РНК-расщепляющую группу, репортер-группу, интеркалятор, группу для улучшения фармакокинетических свойств олигонуклеотида или группу для улучшения фармакодинамических свойств олигонуклеотида и любую их комбинацию.

В некоторых вариантах осуществления олигонуклеотид, содержащий не встречающееся в природе нуклеиновое основание, раскрытое в настоящем изобретении, дополнительно содержит не встречающийся в природе остов. Не встречающийся в природе остов включает, но не ограничивается ими, фосфоротиоат, хиральный фосфоротиоат, фосфоридитиоат, фосфотриэфир, аминокалкилфосфотриэфир, C<sub>1</sub>-C<sub>10</sub>-фосфонаты, 3'-алкиленфосфонат, хиральные фосфонаты, фосфинаты, фосфорамидаты, 3'-аминофосфорамидат, аминокалкилфосфогамидаты, тионофосфорамидаты, тионоалкилфосфонаты, тионоалкилфосфотриэфиры и боранофосфаты.

Способы определения стабильности олигонуклеотидных дуплексов, содержащих не встречающиеся в природе пары оснований (с линкерами или без них) включают термодинамический анализ измерениями кругового дихроизма (KD) и эксперименты по УФ-плавлению. В некоторых вариантах осуществления используются исследования по стабильности ДНК-дуплекса для облегчения выбора подходящей не встречающейся в природе пары нуклеиновых оснований, не встречающегося в природе нуклеинового основания или производных или замен не встречающегося в природе нуклеинового основания. Подходящим образом выбранные не встречающиеся в природе пары оснований включают те, которые увеличивают точность олигонуклеотидной гибридизации в других положениях в дуплексе. Подходящим образом выбранные не встречающиеся в природе пары оснований включают те, которые увеличивают стабильность олигонуклеотидного дуплекса. Подходящим образом выбранные нуклеотидные основания могут быть использованы для оптимизации олигонуклеотидов для биотехнологических или терапевтических применений, где высокая точность гибридизации и дискриминация являются критическими. В некоторых случаях не встречающиеся в природе пары оснований по меньшей мере на 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% или более так же стабильны, как и природные пары оснований в олигонуклеотидном дуплексе. В некоторых случаях величина T<sub>m</sub> дуплекса, содержащего одну или более не встречающихся в природе пар оснований, на 10, 9, 8, 7, 6, 5, 4,5, 4, 3,5, 3, 2,9, 2,8, 2,7, 2,6, 2,5, 2,4, 2,3, 2,2, 2,1, 2, 1,9, 1,8, 1,7, 1,6, 1,5, 1,4, 1,3, 1,2, 1,1, 1, 0,9, 0,8, 0,7, 0,6, 0,5, 0,4, 0,3°C, 0,2, 0,1°C ниже величины T<sub>m</sub> того же дуплекса, в котором одно или более не встречающихся в природе нуклеиновых оснований замещены одним или более природными нуклеотидными основаниями. В некоторых вариантах осуществления присутствие не встречающейся в природе пары оснований в олигонуклеотидном дуплексе существенно не нарушает структуру дуплекса.

В некоторых вариантах осуществления олигонуклеотид, содержащий дериватизированное с помощью линкера нуклеиновое основание, позволяет выполнение сайт-специфической модификации данной ДНК или РНК во время или после ферментативного синтеза. Не встречающийся в природе нуклеотид, раскрытый в настоящем изобретении (например, нуклеотид, содержащий не встречающийся в природе α- или β-аналог нуклеинового основания), в некоторых случаях модифицируется линкером, который позволяет присоединение различных функциональных групп (например, карго) без уменьшения степени узнавания полимеразой. Сайт-специфические функциональные группы включают, но не ограничиваются ими, флуорофоры, ЯМР-идентификаторы для характеристики (например, F19), ИК-зонды (например, азидо- и цианогруппы), биотин (например, чтобы облегчить идентификацию и/или очистку), аффинные метки, липосомы и наночастицы. В одном варианте осуществления линкер обеспечивает биоконъюгацию через кросс-связывание (например, иод-группа). В одном варианте осуществления линкер обеспечивает спейсер для биоконъюгации через реакцию клик-химии (например, азидо- и алкин-заместители). В одном варианте осуществления олигонуклеотид, содержащий дериватизированное с помощью линкера нуклеиновое основание, является применимым в качестве праймера и/или молекулярного маяка.

В различных вариантах осуществления настоящего изобретения далее обеспечивается применение любых нуклеозидных аналогов, раскрытых в настоящем изобретении (α или β), или их аналогов или производных, в сайт-специфическом расщеплении или функционализации олигонуклеотида. В некоторых вариантах осуществления нуклеозидный аналог включает один или более линкеров, сконфигурированных для сайт-специфической модификации. Примеры нуклеотидных аналогов, содержащих линкерную группу, включают, но не ограничиваются ими,

d5SICSCO, d5SICSCC, dDMOCO, dDMOCC, dMMO2pCO, dMMO2pCC, dTPT3, dTPT3A, dTPT3PA, dTPT3CO, dMMSpCO, dMMSPA и dTPT3CC

или их рибозильные формы, или их аналоги. В различных вариантах осуществления настоящего изобретения обеспечиваются композиции собственно функционализированных олигонуклеотидов, способы получения функционализированных олигонуклеотидов и способы применения функционализированных олигонуклеотидов.

Различные варианты осуществления обеспечивают dTPT3, dTPT3PA, dTPT3A, dTPT3CO и dTPT3CC или другие дериватизированные с помощью линкера аналоги dTPT3, включенные в олигонуклеотиды, и дополнительное взаимодействие или дериватизацию данных не встречающихся в природе аналогов нуклеиновых оснований, включенных в олигонуклеотид, с различными реагентами для селективного взаимодействия с не встречающимися в природе аналогами нуклеиновых оснований в олигонуклеотиде, причем встречающиеся в природе нуклеотидные основания (A, T, G, C, U) не реагируют с данными реагентами в сколько-нибудь заметной степени. Базируемое на dTPT3 семейство не встречающихся в природе нуклеотидов, несущих линкер, является особенно важным, поскольку авторы обнаружили, что данные нуклеотиды более эффективно реплицируются ДНК-полимеразами, чем пары оснований, которые включают d5SICS или его дериватизированные с помощью линкера варианты, что значительно облегчает многие потенциальные применения. В некоторых вариантах осуществления процентное включение не встречающегося в природе нуклеотида, содержащего линкер, в олигонуклеотид составляет по меньшей мере 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 или 99%. В некоторых вариантах осуществления процентное удлинение следующим правильным нуклеотидом с образованием олигонуклеотида, где следующий правильный нуклеотид следует за включением не встречающегося в природе нуклеотида, содержащего линкер, составляет по меньшей мере 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90% или более. В некоторых вариантах осуществления добавление сайт-специфической функциональной группы уменьшает процентное включение не встречающегося в природе нуклеотида в олигонуклеотид более чем на около 50, 45, 40, 35, 30, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 или 1%. В некоторых вариантах осуществления точность воспроизведения дериватизированного с помощью линкера не встречающегося в природе нуклеотида составляет по меньшей мере 98, 98,1, 98,2, 98,3, 98,4, 98,5, 98,6, 98,7, 98,8, 98,9, 99, 99,1, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,8, 99,9, 99,91, 99,92, 99,93, 99,94, 99,95, 99,96, 99,97, 99,98 или 99,99%. Соответственно, в различных вариантах осуществления настоящего изобретения обеспечиваются способы использования не встречающихся в природе нуклеотидов, дериватизированных с помощью линкера, для продукции ДНК или РНК, которая является сайт-специфически модифицированной другой молекулой, представляющей интерес. В некоторых вариантах осуществления сайт-специфическое включение различных функциональных групп происходит до или после амплификации. В некоторых вариантах осуществления сайт-специфическую функционализацию используют для приложений SELEX.

Примером стратегии получения ДНК или РНК, которая сайт-специфически модифицирована другой молекулой, представляющей интерес, служит форсфоториотатная стратегия (фиг. 3), которая опирается на сайт-специфическое введение форсфоториотатной группы в ДНК или РНК через рибо- или дезоксирибо- $\alpha$ -тиотрифосфат одного из не встречающихся в природе нуклеозидов из фиг. 1 или 2. После включения в ДНК или РНК форсфоториотатная группа может быть использована для связывания реагентов, которые несут  $\gamma$ -бром- $\alpha,\beta$ -ненасыщенные карбонильные, йод- (или бром-)ацетильные или азиридилисульфонамидные группы для получения сайт-специфически функционализированной ДНК или РНК. Альтернативно, после включения в ДНК или РНК форсфоториотатная группа может быть использована для сайт-специфического расщепления ДНК или РНК с помощью йода в щелочном растворе или йодэтанола соответственно. Таким образом, форсфоториотатная стратегия обеспечивает сайт-специфическую модификацию остова нуклеиновой кислоты и обеспечивает способ сайт-специфического расщепления олигонуклеотидной цепи.

Еще одна стратегия получения ДНК или РНК, которая сайт-специфически модифицирована другой молекулой, представляющей интерес, именуемая линкерной стратегией (фиг. 4 и 5), использует дериватизацию не встречающихся в природе нуклеиновых оснований линкером (фиг. 2), который может быть использован для присоединения функциональных групп, представляющих интерес, либо перед полимеризацией (способом ПЦР или T7 РНК-полимеразы-опосредованной транскрипции с использованием соответствующего функционализированного нуклеотидтрифосфатного реагента, который включается в синтезируемую ДНК- или РНК-цепь), либо с помощью реакции взаимодействия линкера не встречающегося в природе нуклеинового основания после включения в олигонуклеотидную цепь с соответствующим реагентом функционализации, например NHS-содержащим реагентом, также содержащим целевую функциональную группу, где NHS реагирует со свободной аминогруппой аминифункционализированного, не встречающегося в природе нуклеинового основания, такого как d5SICSA, dMMO2A или dTPT3A.

На фиг. 4 также показана линкерная стратегия аминифункционализации с использованием  $d5SICS^A$  и  $dMMO2^{PA}$ , которая позволяет сайт-специфическое двойное мечение дуплексной ДНК.

Например, функционализация может быть достигнута после включения в олигонуклеотид не встречающегося в природе нуклеинового основания с линкером, несущим первичную аминогруппу (например,  $dTPT3^A$ ). Более конкретно, функционализация может осуществляться с помощью реакции первичной аминогруппы (например, пропаргиламиногруппы) и несущего карго реагента, включая сложный N-гидроксисукцинимидный (NHS) эфир (фиг. 4). Аналоги, разработанные для данного применения, включают в себя  $d5SICS^A$ ,  $d5SICS^{PA}$ ,  $dMMO2^A$ ,  $dMMO2^{PA}$ ,  $dTPT3^A$  и  $dTPT3^{PA}$  (напомним, что "A" относится к нуклеотиду с пропаргиламином и "PA" относится к тому же линкеру с защитной группой, см. фиг. 2 и подпись к ней). Применение  $dTPT3^{PA}$ , несущего защищенную первичную пропаргиламиногруппу, и  $dTPT3^A$ , несущего первичную пропаргиламиногруппу, для функционализации, специфической в отношении последовательности олигонуклеотидов, раскрыто и заявлено в данном изобретении.

Сайт-специфическая функционализация олигонуклеотида может также выполняться с использованием катализируемого медью(I) азид-алкинового циклоприсоединения (CuAAC) (т.е. линкерной стратегии с применением реакции "клик-химии"; фиг. 5), и для данных применений могут быть использованы  $d5SICS^{CO}$ ,  $d5SICS^{CC}$ ,  $dDMO^{CO}$ ,  $dDMO^{CC}$ ,  $dMMO2^{PCO}$ ,  $dMMO2^{PCC}$ ,  $dTPT3^{CO}$  и  $dTPT3^{CC}$  (фиг. 2). В каждом случае рибозилтрифосфаты не встречающихся в природе нуклеиновых оснований могут применяться для транскрипции с целью получения сайт-специфически меченой РНК, и дезоксирибозилтрифосфаты не встречающихся в природе нуклеиновых оснований могут использоваться, например, в ПЦР, для получения сайт-специфически меченой ДНК. Не встречающиеся в природе нуклеотидные основания, содержащие ацетиленовую (алкинильную) линкерную группу, подходящую для применения в CuAAC-конъюгации,  $d5SICS^{CO}$ ,  $d5SICS^{CC}$ ,  $dDMO^{CO}$ ,  $dDMO^{CC}$ ,  $dMMO2^{PCO}$ ,  $dMMO2^{PCC}$ ,  $dTPT3^{CO}$  и  $dTPT3^{CC}$ , способы их получения, а также способы их применения в получении таких сайт-специфически меченых олигонуклеотидов, раскрыты и заявлены в данном изобретении.

Демонстрация общей форсфоротиаотной стратегии (фиг. 3). Чтобы продемонстрировать возможности их системы, авторы приготовили  $\alpha$ -тиотрифосфат не встречающегося в природе нуклеотида,  $d5SICS$  ( $d5SICS-aS$ ), и встроили его в ДНК в положении напротив соответствующего ему, не встречающегося в природе нуклеотида  $dNaM$ , используя стандартную ПНР. Эффективность амплификации и точность включения  $d(d5SICS-aS)TP$  превышает 99% и практически идентичны результатам, полученным с  $d5SICS$ . Для функционализации данной не встречающейся в природе пары оснований, авторы подвергли сайт-специфически включенную форсфоротиаотную связь взаимодействию с иодоацетил-PEG<sub>2</sub>-биотином, чтобы пометить ДНК-дуплекс функциональностью биотина.<sup>8</sup> Чтобы охарактеризовать данный сайт-специфический аддукт, авторы инкубировали его в присутствии стрептавидина и затем количественно определяли степень функционализации анализом сдвига в геле. Авторы смогли преобразовать от 60 до 70% форсфоротиаотной связи в функционализированное производное, что является стандартной эффективностью (70%) для протоколов мечения, ранее описанных в литературе (см. Fidanza, J.A.; Ozaki, H.; McLaughlin, L.W., Site-specific labeling of DNA sequences containing phosphorothioate diesters. J. Am. Chem. Soc. 2002, 114 (14), 5509-5517). Данные конъюгированные производные обладают высокой стабильностью в условиях, характерных для тепловой денатурации дуплексов ДНК, т.е. при 50°C в течение ночи в диапазоне pH от 6,0 до 8,3 (<10% разложение), а также при 95°C в течение 3 мин при pH 8,3 (<5% разложение). Авторы предполагают, что форсфоротиаотную стратегию можно одинаково хорошо использовать с другими парами оснований, не встречающимися в природе, включая  $d5SICS$ - $dMMO2$  и  $d5SICS$ - $dNaM$ .

В различных вариантах осуществления настоящего изобретения обеспечивается фосфоротиаотная стратегия с использованием не встречающихся в природе пар оснований  $dTPT3$ - $dNaM$ ,  $dTPT3$ - $dMMO2$  или  $dTPT3$ - $dDMO$  и дериватизированных с помощью линкера их вариантов.

Фосфоротиаотная стратегия и стратегия на основе линкера не являются взаимоисключающими и в сочетании должны позволить провести модификацию данного сайта различными функциональными группами в количестве до трех, причем одна присоединяется к первому нуклеиновому основанию пары нуклеиновых оснований, вторая присоединяется к второму нуклеиновому основанию пары нуклеиновых оснований, а третья присоединяется к остову непосредственно в 5' положении не встречающегося в природе нуклеотида.

Демонстрация линкерной стратегии с первичным амином (фиг. 4). Для дальнейшей демонстрации осуществимости своей системы авторы синтезировали и охарактеризовали дериватизированные с помощью амино- и защищенного аминолинкера варианты  $d5SICS$  и  $dMMO2$  (фиг. 2). Авторы показали, что, будучи спаренными в ДНК с расположенными напротив их соответствующими, не встречающимися в природе партнерами, каждый из них хорошо амплифицируется с помощью ПЦР и транскрибируется в РНК. Связывание ДНК, содержащей  $dMMO2^A$  или  $d5SICS^A$ , полученными ПНР-амплификацией, с NHS-эфиром биотина, протекает с эффективностью от 55 до 70% соответственно.

Авторы показали, что рибонуклеотидтрифосфаты  $5SICS^{PA}$ ,  $5SICS^A$ ,  $MMO2^{PA}$  или  $MMO2^A$  транскрибируются в РНК Т7 РНК-полимеразой с высокой эффективностью и точностью воспроизведения (фиг. 7).

В различных вариантах осуществления настоящего изобретения обеспечивается сайт-

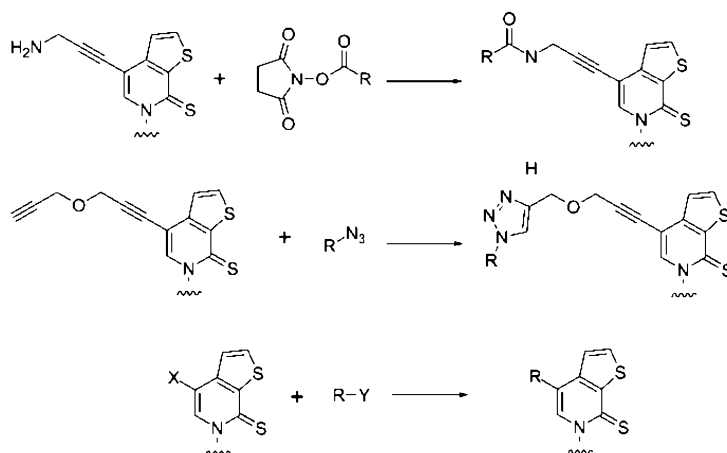
специфическая модификация ДНК или РНК с использованием dTPT3<sup>L</sup>-dNaM, dTPT3<sup>L</sup>-dMMO2 или dTPT3<sup>L</sup>-(d)DMO (где R представляет собой линкер, например, R = H для dTPT3, R = 3-аминопропин-1-ил для dTPT3<sup>A</sup>, R = дихлорацетил-3-аминопропин-1-ил для dTPT3<sup>PA</sup>, R = 4-оксагепта-1,6-диин-1-ил для dTPT3<sup>CO</sup>, R = гепта-1,6-диин-1-ил для dTPT3<sup>CC</sup>).

Демонстрация общей линкерной стратегии с алкинами (фиг. 5). Для дальнейшей демонстрации осуществимости своей системы авторы синтезировали и охарактеризовали алкинильные функционализированные варианты d5SICS, dDMO и dMMO2, включая d5SICS<sup>CO</sup>, d5SICS<sup>CC</sup>, dDMO<sup>CO</sup>, dDMO<sup>CC</sup>, dMMO2<sup>CO</sup>, dMMO2<sup>CC</sup> (фиг. 2). Каждый из данных алкин-функционализированных, не встречающихся в природе нуклеотидов должен эффективно амплифицироваться в ПНР, когда присутствует в ДНК. После амплификации ДНК, содержащая, например, пару оснований d5SICS<sup>CO</sup>-dNaM, может быть эффективно сайт-специфически модифицирована малыми молекулами или одним или более белками, обладающими азидными группами с использованием реакций клик-химии, например, катализируемых медью клик-реакций. Авторы также продемонстрировали применимость dEMO и dFEMO (фиг. 2) для включения в олигонуклеотиды и применение данных функционализированных олигонуклеотидов в реакциях клик-химии с азидами для функционализации олигонуклеотидов сайт-специфическим образом.

Демонстрация линкерной стратегии с dTPT3<sup>PA</sup> (фиг. 6). Добавление линкера к скаффолдам d5SICS и dMMO2 значительно снижает эффективность, с которой не встречающиеся в природе нуклеотиды ферментативно включаются в ДНК, что, как ожидается, ограничит их практическое применение. Тем не менее, авторы обнаружили, что скаффолд dTPT3 гораздо более толерантен к добавлению линкера (фиг. 6). Например, dTPT3<sup>PA</sup>-TP включается ДНК-полимеразами в праймер напротив dNaM в матрице с практически такой же эффективностью и точностью воспроизведения, как и природная пара оснований. Соответственно, в различных вариантах осуществления настоящего изобретения обеспечивается применение не встречающихся в природе нуклеиновых оснований на основе скаффолда dTPT3, включая dTPT3<sup>PA</sup> (защищенный амино-функциональный линкер), dTPT3<sup>A</sup> (амино-функциональный линкер) и dTPT3<sup>CO</sup> (алкин-азидный эфирный линкер для дериватизации с помощью клик-химии), и dTPT3<sup>CC</sup> (алкин-азидный триметиленовый линкер для дериватизации с помощью клик-химии) в синтезе сайт-специфически функционализированных олигонуклеотидов.

На схеме 1 показаны примеры dTPT3 с различными линкерами, которые могли бы быть использованы для сайт-специфической модификации ДНК или РНК.

Схема 1



Для ясности показаны только фрагменты структуры нуклеинового основания dTPT3, но следует понимать, что они используются, будучи частью нуклеотидов. Реакции функционализации могут осуществляться либо до введения не встречающихся в природе нуклеиновых оснований в олигонуклеотид, либо после этого. Схема 1, верхняя реакция, иллюстрирует применение dTPT3<sup>A</sup>, содержащего несущий первичный амин линкер, который ацилируют с помощью активированного сложного эфира с образованием амида, в котором R-группа содержит карго. Средняя реакция схемы 1 иллюстрирует применение dTPT3<sup>CO</sup> (также может быть использован dTPT3<sup>CC</sup>), содержащего несущий алкинильную группу линкер, который взаимодействует с азидом с получением триазола через реакцию клик-химии, где R-группа триазола, который образуется, содержит карго. Нижняя реакция иллюстрирует наиболее общий случай для производного скаффолда dTPT3, несущего линкерную группу R<sub>1</sub> с реакционноспособной группой, которая может избирательно образовывать ковалентную связь с группой R<sub>2</sub>, которая включает реакционно-способную группу, комплементарную реакционноспособной группе линкера, например тиол-малеимид, гидразин-альдегид и т.д.

В одном варианте осуществления линкер, содержащий азидную реакционноспособную группу, является применимым для присоединения карго, содержащего алкин, посредством клик-реакции. В одном варианте осуществления линкер, содержащий тиольную группу, может образовывать обратимые ди-

сульфидные связи или необратимые связи с различными, принимающими карго группами, включая, но не ограничиваясь ими, малеимид, бромид, иодид, сульфонильные производные, активные сложные эфиры и изотиоцианатные производные. В одном варианте осуществления линкер, содержащий азидную группу, вступает в реакцию с карго-молекулой, содержащей фосфиновую группу.

В одном варианте осуществления олигонуклеотид, содержащий одно или более дериватизированных с помощью линкера не встречающихся в природе нуклеиновых оснований, сконфигурирован для применения в качестве молекулярного маяка. Флуорофор молекулярного маяка представляет собой карго-молекулу, присоединенную к реакционноспособному центру дериватизированного с помощью линкера, не встречающегося в природе нуклеинового основания. Примеры карго-молекул флуорофора включают, но не ограничиваются ими, 6-FAM, флуоресцеин, Cy3™, JOE (6-карбоксихлор-2',7'-диметоксифлуоресцеин), Cy5™, TAMRA, MAX, TET™, ROX (карбоксихлор-родамин), TYE™ 563, гексахлорфлуоресцеин, TEX 615, TYE™ 665, TYE 705, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 594, Alexa Fluor® 647, Alexa Fluor® 660, Alexa Fluor® 750, IRDye® 800CW, ATTO™ 488, ATTO™ 532, ATTO™ 550, ATTO™ 565, ATTO™ Rho101, ATTO™ 590, ATTO™ 633, ATTO™ 647N, родамин Green™-X, родамин Red™-X, 5-TAMRA™, Texac Red®-X, Lightcycler® 640 и Dy 750.

В некоторых вариантах осуществления не встречающаяся в природе пара оснований, позволяет заполнить сайт-специфическое включение различных функциональных групп в ДНК для приложений SELEX (Systematic Evolution of Ligands by Exponential Enrichment, Систематическая эволюция лигандов экспоненциальным обогащением), включая генерацию ДНК- и/или РНК-аптамеров. ДНК- и РНК-аптамеры имеют разнообразные мишени, включая нуклеиновые кислоты, малые молекулы, пептиды, углеводы и клетки. SELEX включает в себя создание библиотеки молекул нуклеиновых кислот, контактирование библиотеки с молекулами-мишенями для выбора молекул нуклеиновых кислот, которые связываются с молекулами-мишенями, и амплификацию членов библиотеки, которые связываются с молекулами-мишенями. Дополнительные раунды отбора и амплификации продолжают, пока не получат достаточно аптамеров. Аптамер в одном аспекте включает в себя любое не встречающееся в природе основание, раскрытое в настоящем изобретении. В некоторых вариантах осуществления эксперимент SELEX, в котором компоненты библиотеки содержат не встречающиеся в природе нуклеотидные основания, генерирует сродство аптамеров против молекулы-мишени в результате 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или менее раундов отбора по сравнению с библиотекой, которая не содержит не встречающиеся в природе нуклеотидные основания. В некоторых вариантах осуществления аптамер, содержащий одно или более нуклеиновых оснований, не встречающихся в природе, имеет большее сродство к молекуле-мишени, чем аптамер, содержащий только природные нуклеотидные основания. Добавление одного или более не встречающихся в природе нуклеиновых оснований в библиотеку SELEX повышает химическое и структурное разнообразие получаемых ДНК- или РНК-аптамеров. В некоторых вариантах осуществления не встречающийся в природе аптамер имеет по меньшей мере наномолярное сродство в отношении его молекулы-мишени. В некоторых вариантах осуществления не встречающийся в природе аптамер имеет по меньшей мере пикомолярное сродство в отношении его молекулы-мишени. Например, не встречающийся в природе аптамер имеет сродство к его молекуле-мишени, которое составляет от 1 до 1000 пмоль. В некоторых вариантах осуществления не встречающийся в природе аптамер имеет по меньшей мере фемтомолярное сродство к его молекуле-мишени. Например, не встречающийся в природе аптамер имеет сродство к его молекуле-мишени, которое составляет от 1 до 1000 фмоль. Не встречающийся в природе аптамер, выбранный с помощью SELEX, может включать 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 или более нуклеиновых оснований, не встречающихся в природе. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит dTPT3 или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14a$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14b$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14c$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14d$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14e$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\alpha 14f$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\beta 8a$ , или его производное или аналог. В некоторых вариантах осуществления не встречающийся в природе аптамер содержит нуклеиновое основание, имеющее формулу  $\beta 8b$ , или его производное или аналог.

Различные комбинации компонентов, изложенных выше, в отношении примеров реакционных смесей и реакционных способов, могут быть предоставлены в виде набора. Такой набор может включать в себя отдельные компоненты, которые отделены друг от друга, например находиться в отдельных сосудах или упаковках. Набор может включать в себя одну или более субкомбинаций компонентов, изложенных

в данном документе, одну или более субкомбинаций, отделенных от других компонентов набора. Субкомбинации могут объединяться для создания изложенной в данном изобретении (или объединенной для проведения реакции, изложенной в настоящем изобретении) реакционной смеси. В конкретных вариантах осуществления субкомбинации компонентов, которая присутствуют в индивидуальном сосуде или упаковке, недостаточно для проведения реакции, изложенной в данном изобретении.

Однако набор, в целом, может включать в себя комплект сосудов или упаковок, содержимое которых можно объединять, чтобы осуществить реакцию, описанную в данном изобретении.

Набор может включать подходящий упаковочный материал для размещения содержимого набора. Упаковочный материал может быть сделан хорошо известными способами, предпочтительно, чтобы обеспечить стерильную, свободную от загрязнений среду. Упаковочные материалы, используемые в настоящем изобретении, могут включать, например, те, которые обычно используются в коммерческих наборах, продаваемых для применения с системами секвенирования нуклеиновой кислоты. Примеры упаковочных материалов включают, без ограничения, стекло, пластик, бумагу, фольгу и т.п., способные сохранять в пределах установленных ограничений компонент, изложенный в настоящем изобретении.

Упаковочный материал может включать в себя этикетку, которая указывает на конкретный вид применения для компонентов. Применение для набора, которое указано на этикетке, может быть одним или более из числа способов, указанных в данном описании как подходящее для конкретной комбинации компонентов, присутствующих в наборе. Например, на этикетке может быть указано, что набор применим для способа конъюгации карго-молекулы с линкерной группой не встречающегося в природе нуклеинового основания в олигонуклеотиде.

Инструкции по применению упакованных реагентов или компонентов также могут быть включены в набор. Инструкции обычно включают ясные термины, описывающие параметры реакции, такие как необходимые для смешивания относительные количества компонентов набора и образца, необходимые для смешивания реагента/образца временные периоды, температура, буферные условия и т.п.

Следует понимать, что не все компоненты, необходимые для конкретной реакции, должны присутствовать в конкретном наборе. Один или более дополнительных компонентов может быть получен из других источников. В инструкциях, предоставленных с набором, может содержаться разъяснение по поводу дополнительного(ых) компонента(ов), которые должны быть предоставлены, и где они могут быть получены.

В одном варианте осуществления набор содержит одно или более не встречающихся в природе нуклеиновых оснований или их производные и реагенты, сконфигурированные для выполнения сайт-специфической функционализации с использованием одного или более не встречающихся в природе нуклеиновых оснований или их производных.

#### Примеры

В настоящее время свободные нуклеозиды и фосфорамидаты соединений d5SICS и dNaM коммерчески доступны от Berry and Associates (Dexter, MI).

Пример 1. Скрининг на основе ПЦР для выявления не встречающихся в природе пар оснований.

Трифосфаты  $\alpha$ -группы были получены из ранее описанных нуклеозидов (Kubelka, T., Slavetinska, L., Eigner, V. and Hock, M. Synthesis of 2,6-disubstituted pyridin-3-yl C-2'-deoxyribonucleosides through chemoselective transformations of bromo-chloropyridine C-nucleosides. *Org. Biomol. Chem.*, 11, 4702-4718) в соответствии со статьей Ludwig, J. and Eckstein, F. Rapid and efficient synthesis of nucleoside 5'-0-(1-thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.*, 54, 631-635. Чистоту всех других трифосфатов подтверждали способами MALDI-TOF и UV-VIS. ДНК-полимеразы Taq и OneTaq были приобретены у New England Biolabs (Ipswich, MA). Смесь ДНТФ была приобретена у Fermentas (Glen Burnie, MD). Краситель SYBR Green I Nucleic Acid Gel Stain (10000 $\times$ ) был приобретен у компании Life Technologies (Carlsbad, CA). Синтез ДНК-матриц, D8 (Malyshev, D.A., Dhami, K., Quach, H.T., Lavergne, T., Ordoukhanian, P., Torkamani, A. and Romesberg, F.E. Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *Proc. Natl. Acad. Sci. USA*, 109, 12005-12010), используемой для раундов скрининга 1-5, и D6 (Malyshev, D.A., Seo, Y.J., Ordoukhanian, P. and Romesberg, F.E. PCR with an expanded genetic alphabet. *J. Am. Chem. Soc.*, 131, 14620-14621), используемой для всех других амплификаций, был описан ранее. Секвенирование по Сэнгеру (Sanger) проводили, как описано ранее (Malyshev, D.A., Dhami, K., Quach, H.T., Lavergne, T., Ordoukhanian, P., Torkamani, A. and Romesberg, F.E. Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *Proc. Natl. Acad. Sci. USA*, 109, 12005-12010). Необработанные "лесенки", полученные секвенированием по Сэнгеру, использовались для определения процента сохранения не встречающихся в природе пар оснований, который преобразовывали в значения точности воспроизведения в пересчете на дубликацию, как описано (Malyshev, D.A., Dhami, K., Quach, H.T., Lavergne, T., Ordoukhanian, P., Torkamani, A. and Romesberg, F.E. Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *Proc. Natl. Acad. Sci. USA*, 109, 12005-12010; Malyshev, D.A., Seo, Y.J., Ordoukhanian, P. and Romesberg, F.E. PCR with an expanded genetic alpha-



Таблица 3

## Условия ПЦР для каждого последовательного раунда ПЦР-скрининга

Компоненты реакции	Раунды	Раунд 5	Раунд 6	Финальная характеристика PCR-	
Буфер	1×OneTaq	1×OneTaq	1×OneTaq	1×OneTaq	1×OneTaq
Фермент	OneTaq	Taq	OneTaq или Taq	Taq	OneTaq
Матрица	D8 <sup>a</sup> (0,1 нг)	D8 <sup>a</sup> (0,1 нг)	D6 <sup>a</sup> (0,01 нг)	D6 <sup>a</sup> (0,01 нг)	D6 <sup>a</sup> (0,01 нг)
Термические условия					
Начальная денатурация	96 °C, 60 с	96 °C, 60 с	96 °C, 60 с	96 °C, 60 с	96 °C, 60 с
Денатурация	96 °C, 10 с	96 °C, 10 с	96 °C, 5 с	96 °C, 5 с	96 °C, 10 с
Отжиг	60 °C, 15 с	60 °C, 15 с	60 °C, 5 с	60 °C, 5 с	60 °C, 5 с
Удлинение	68 °C, 60 с	68 °C, 60 с	68 °C, 10 с	68 °C, 10 с	68 °C, 30 с
Количество циклов	16	16	24	16 +16 +20 <sup>b</sup>	16 +16 +20 <sup>b</sup>

<sup>a</sup> См табл. 2 для последовательностей матриц и праймеров.

<sup>b</sup> Первоначальное количество матрицы составляло 0,01 нг. ПЦР-смесь амплифицировали в течение 16 циклов, разбавляли в 40000 раз и амплифицировали в течение еще 16 циклов. Стадию разбавления/амплификации повторяли, в результате чего имели 52 полных циклов амплификации.

Конкретные условия анализа ПЦР. ПЦР с наиболее перспективными UBP проводили с условиями, как описано в табл. 3. Продукты ПЦР дополнительно очищали на 2% агарозном геле, а затем вырезали одну полосу и проводили последующую очистку с использованием набора Zymo Research Zymoclean Gel DNA Recovery Kit. После элюции с 20 мкл воды концентрацию ДНК измеряли с использованием связывания флуоресцентного красителя (Quant-iT dsDNA HS Assay kit, Life Technologies), и очищенные ампликоны секвенировали в трех экземплярах с обоими праймерами Primer1 и Primer2, чтобы определить степень сохранения UBP и, таким образом, точность амплификации. Амплификация ДНК, содержащей пары с участием аналогов группы α6, проводили с OneTaq-полимеразой при следующих термических циклических условиях: первоначальная денатурация при 96°C в течение 1 мин; 16 циклов 96°C в течение 10 с, 60°C в течение 15 с, 68°C в течение 1 мин. Точность воспроизведения определяли секвенированием ампликонов в направлении Primer1 в трех повторах. Амплификацию ДНК, содержащей UBP, образованные между dTPT3 и d2MN или dDM2, проводили с использованием OneTaq- или Taq-полимеразы в течение 16 циклов при следующих термических циклических условиях: 1) OneTaq: начальная денатурация при 96°C в течение 1 мин, 96°C в течение 10 с, 60°C в течение 15 с, 68°C в течение 1 мин; или 2) Taq: начальная денатурация при 96°C в течение 1 мин, 96°C в течение 5 с, 60°C в течение 5 с, 68°C в течение 10 с. Точность воспроизведения определяли секвенированием ампликонов в направлении Primer1 в трех повторах.

Результаты. Для скрининга хорошо реплицируемых UBP не встречающиеся в природе дезокси-нуклеозидтрифосфаты были сгруппированы для анализа в dMMO2/dNaM- или d5SICS/dTPT3-подобные аналоги, хотя различие не совсем ясно во всех случаях. В общей сложности 80 аналогов dMMO2/dNaM были сгруппированы в двенадцать "α-групп" (α1-α12, фиг. 8) и 31 аналог d5SICS/dTPT3 были сгруппированы в шесть "β-групп" (β1-β6; фиг. 9). Отметим, что обозначения групп, используемые в настоящем изобретении, не следует путать с обозначением аномера (все рассмотренные аналоги представляют собой β-гликозиды). Кроме того, чтобы расширить поиски взаимосвязи структуры с активностью (structure-activity relationship (SAR)) при скрининге, семь ранее описанных нуклеотидных аналогов (dTOK576-dTOK588) с замещенными пиридинными нуклеотидными основаниями (Kubelka, T., Slavetinska, L., Eigner, V. and Hockem M. Synthesis of 2,6-disubstituted pyridin-3-yl C-2'-deoxyribonucleosides through chemoselective transformations of bromo-chloropyridine C-nucleosides. Org. Biomol. Chem., 11, 4702-4718) были фосфорилированы и включены в качестве группы α6. Для скрининга 134-мерную матрицу одноцепочечной ДНК, содержащую центрально расположенный dNaM (именуемый D8), подвергли ПНП-амплификации в присутствии природных трифосфатов (200 мкМ каждого), все парные комбинации α- и β-трифосфатной группы показаны на фиг. 8 и 9 (50 мкМ каждый), и 0,02 ед/мкл ДНК-полимеразы. В ходе



первого раунда ПЦР dNaM служит матрицей при включении  $\alpha$ -аналога и затем заменяется  $\beta$ -аналогом, когда первоначальная цепь копируется во втором раунде, с получающимся UBP, амплифицируемым в последующих раундах. Продукт амплификации каждой реакции анализировали с помощью секвенирования по Сэнгеру. Присутствие не встречающегося в природе нуклеотида приводит к внезапному прекращению хроматограммы секвенирования, что позволяет проводить количественное определение уровня UBP-сохранения на основе анализа полученных данных. Процентное содержание UBP, сохранившегося в ДНК после амплификации в ходе каждого раунда скрининга, показано на фиг. 14.

В первом раунде отбора использовали 0,1 нг матрицы и 16 циклов амплификации при относительно перmissive условиях, которые включали OneTaq-полимеразу и время удлинения 1 мин. В случае данного примера OneTaq рассматривается как перmissive, потому что она представляет собой смесь Taq (полимераза семейства A) и Deep Vent (полимераза семейства B), причем последняя обладает экзонуклеазной корректирующей активностью, которая позволяет удалять ошибочно включенный трифосфат. В данных условиях только пары с участием группы  $\beta 5$  или  $\beta 6$  показали высокий уровень сохранения.

Сочетания  $\beta 5$ - или  $\beta 6$ - и  $\alpha$ -групп, которые показали самый высокий уровень сохранения, дошли до второго раунда скрининга, в котором они были разделены на более мелкие группы (помеченные a, b или c; фиг. 8 и 9). Высокое сохранение (>97%) наблюдалось с  $\beta 5a$  и  $\alpha 2c$ ,  $\alpha 9a$ ,  $\alpha 9c$ ,  $\alpha 10a$ ,  $\alpha 10c$ ,  $\alpha 12b$  или  $\alpha 12c$ ; с  $\beta 5b$  и  $\alpha 9a$ ,  $\alpha 9b$ ,  $\alpha 10c$  или  $\alpha 12b$ ; и с  $\beta 6b$  и  $\alpha 10c$ . Умеренное сохранение (от 84 до 96%) наблюдалась с  $\beta 5a$  и  $\alpha 1a$ ,  $\alpha 1b$ ,  $\alpha 6a$ ,  $\alpha 9b$ ,  $\alpha 10b$  или  $\alpha 12a$ ;  $\beta 5b$  и  $\alpha 1a$ ,  $\alpha 1b$ ,  $\alpha 2c$ ,  $\alpha 6a$ ,  $\alpha 9c$ ,  $\alpha 10a$ ,  $\alpha 12a$  или  $\alpha 12c$ ;  $\beta 6a$  и  $\alpha 1b$  или  $\alpha 10c$ ; и  $\beta 6b$  и  $\alpha 1a$ ,  $\alpha 6a$ ,  $\alpha 9a$ -с,  $\alpha 10a$ ,  $\alpha 10b$  или  $\alpha 12a$ -с.

Для третьего раунда скрининга  $\alpha$ -аналоги были проанализированы в группах из только от одного до трех соединений, и группу  $\beta 6a$  подразделили на две составные подгруппы из двух трифосфатов, dTPT1TP и dFPT1TP. Самое высокое сохранение (>90%) наблюдали с  $\beta 5a$  и  $\alpha 1a$ ,  $\alpha 2cII$ ,  $\alpha 9a$ -с,  $\alpha 10aI$ ,  $\alpha 10aII$ ,  $\alpha 10c$ ,  $\alpha 12b$  или dTfMOTP;  $\beta 5b$  и  $\alpha 9a$ ,  $\alpha 9c$  или  $\alpha 10c$ ; dFPT1TP и  $\alpha 10aI$ ; и  $\beta 6b$  и  $\alpha 1a$ ,  $\alpha 9a$ -с,  $\alpha 10aI$ ,  $\alpha 10aII$ ,  $\alpha 10c$ ,  $\alpha 12b$ , dNMOTP, dTfMOTP или dCNMOTP. Незначительно меньше сохранение (от 80 до 89%) было отмечено с  $\beta 5a$  и  $\alpha 2cI$ ,  $\alpha 12a$ , dNMOTP, dQMOTP или dTOK587TP;  $\beta 5b$  и  $\alpha 1a$ ,  $\alpha 2cII$ ,  $\alpha 10aI$ ,  $\alpha 10aII$ ,  $\alpha 12b$  или dTOK587TP; dFPT1TP и  $\alpha 10c$ ; и  $\beta 6b$  и  $\alpha 12a$ , dQMOTP, dFuMO1TP или dTOK587TP.

Для четвертого раунда скрининга все  $\alpha$ -производные из фиг. 9 были проанализированы как отдельные трифосфаты, за исключением  $\alpha 9b$  и  $\alpha 9c$ , которые остались в одной группе. Самое высокое сохранение ( $\geq 91\%$ ) наблюдали с  $\beta 5a$  и  $\alpha 9b$ ,  $\alpha 9c$ , dFIMOTP, dIMOTP,

dFEMOTP, dMMO2TP, d2OMeTP, dDMOTP, d5FMTP, dNaMTP, dVMOTP, dZMOTP, dCIMOTP, dTfMOTP, dQMOTP, d2MNTP, dDM2TP или dTOK587TP;  $\beta 5b$  и  $\alpha 9b$ ,  $\alpha 9c$ , dFIMOTP, dIMOTP, dFEMOTP, dNaMTP, dZMOTP, dCIMOTP, dQMOTP, dMM1TP, dDM2TP или dTOK587TP;  $\beta 6$ -аналогом dFPT1TP и  $\alpha$ -аналогами d2OMeTP или dNaMTP; и  $\beta 6b$  и  $\alpha 9b$ ,  $\alpha 9c$ , dFIMOTP, dIMOTP, dFEMOTP, dMMO2TP, dDMOTP, dTMOTP, dNMOTP, d5FMTP, dNaMTP, dVMOTP, dZMOTP, dCIMOTP, dTfMOTP, dQMOTP, dCNMOTP, d2MNTP, dTOK587TP или dFuMO2TP.

Для увеличения жесткости скрининга пятый раунд был проведен с Taq-полимеразой вместо OneTaq, поскольку она не имеет экзонуклеазной корректирующей активности и, таким образом, повышает чувствительность к синтезу ошибочных пар. Данный раунд также разделил все оставшиеся  $\alpha$ - и  $\beta$ -группы на отдельные трифосфаты. Самое высокое сохранение (>90%) было замечено с

dSICSTP и dNaMTP; dSNICSTP и dNaMTP; dTPT2TP и

dFDMOTP; dTPT3TP и dFIMOTP, dIMOTP или dNaMTP; и dFTPT3TP и dFIMOTP, dIMOTP, dFEMOTP, dNMOTP, dNaMTP, dCIMOTP, dTfMOTP или dCNMOTP.

Для того чтобы лучше различать UBP, авторы подвергли шестьдесят два наиболее перспективных кандидатов-UBP шестому раунду скрининга, в котором концентрация матрицы была снижена в 10 раз (до 10 пг), чтобы обеспечить больше амплификацию и, тем самым, позволить больше дискриминацию, и матрицу изменили на D6 (Malyshev, D.A., Seo, Y.J., Ordoukhanian, P. and Romesberg, F.E. PCR with an expanded genetic alphabet. J. Am. Chem. Soc., 131, 14620-14621), где три фланкирующих нуклеотида по обе стороны от не встречающегося в природе нуклеотида были рандомизированы среди природных нуклеотидов. Кроме того, стадии денатурации и отжига были уменьшены до 5 с каждая, а время удлинения было уменьшено до 10 с. В данных условиях авторы исследовали амплификацию либо с OneTaq, либо только с Taq. Результаты с OneTaq показали самое высокое сохранение (> 95%) с

dSICSTP и dNaMTP; dSNICSTP и dFEMOTP; dTPT3TP и

dFIMOTP, dIMOTP, dFEMOTP, dZMOTP или dNaMTP; и dFTPT3TP и dIMOTP или dFEMOTP.

Умеренное сохранение (от 86 до 94%) наблюдали с

dSICSTP и dFEMOTP или

dDM2TP; d5SICSTP и dNaMTP; dSNICSTP или dIMOTP; dTPT2TP и dNaMTP; dTPT3TP и dNMOTP, dCIMOTP, dQMOTP, dCNMOTP или d2MNTTP; и dFTPT3TP и dFIMOTP, dNaMTP, dZMOTP, dCIMOTP, dTfMOTP или dCNMOTP.

В то время как сохранение во время Taq-опосредованной амплификации было, в общем, уменьшено по сравнению с таковым при OneTaq-опосредованной амплификации, общие тенденции были схожими. Самое высокое сохранение (>96%) наблюдали с dTPT3TP и dFIMOTP или dIMOTP, и с dFTPT3TP и dFIMOTP. Только незначительно ниже сохранение (от 89 до 94%) наблюдали с dTPT3TP и dFEMOTP, dNaMTP или dCNMOTP; и dFTPT3TP и dIMOTP, dFEMOTP, dNaMTP, dCIMOTP, dCNMOTP или d2MNTTP.

Амплификацию с наиболее перспективными комбинациями трифосфатов, dTPT3TP или dFTPT3TP и dFIMOTP, dIMOTP, dFEMOTP или dNaMTP, затем выполняли на протяжении более 52 циклов с Taq и 10-секундным временем удлинения, чтобы исследовать более строгие условия, или с OneTaq и 30-секундным временем удлинения, чтобы изучить более практические условия (табл. 4). Обе амплифицированные цепи были секвенированы в трех повторах, чтобы определить UBP-сохранение с высокой точностью воспроизведения. dTPT3-dNaM, dTPT3-dFIMO, dFTPT3-dNaM и dFTPT3-dFIMO показали самую высокую степень сохранения с Taq-полимеразой, в то время как пары с участием dIMO и dFEMO показали несколько меньшее сохранение. С OneTaq-полимеразой dTPT3-dNaM и dFTPT3-dNaM показали высокую степень сохранения, следом за ними dFTPT3-dFIMO и dTPT3-dFIMO.

Таблица 4

dβTP	dαTP	Амплификация, ×10 <sup>12</sup>	Сохранение, %	Точность воспроизведения за дубликацию, %
<i>Taq</i> , Удлинение 10 с				
<b>TPT3</b>	<b>FIMO</b>	8,5	84 ± 3	99,60 ± 0,09
	<b>IMO</b>	6,3	81 ± 5	99,50 ± 0,15
	<b>FEMO</b>	5,0	79 ± 3	99,44 ± 0,09
	<b>NaM</b>	5,8	86,5 ± 0,5	99,66 ± 0,01
<b>FTPT3</b>	<b>FIMO</b>	4,8	84 ± 3	99,60 ± 0,09
	<b>IMO</b>	5,6	82 ± 5	99,54 ± 0,13
	<b>FEMO</b>	5,7	81 ± 4	99,51 ± 0,11
	<b>NaM</b>	3,7	91 ± 6	99,76 ± 0,15
<b>5SICS</b>	<b>NaM</b>	9,3	< 50 <sup>b</sup>	< 85 <sup>b</sup>
<i>OneTaq</i> , Удлинение 1 мин				
<b>TPT3</b>	<b>FIMO</b>	8,7	84,7 ± 1,1	99,61 ± 0,03
	<b>IMO</b>	9,4	82,9 ± 1,7	99,56 ± 0,05
	<b>FEMO</b>	10,4	82,2 ± 1,0	99,55 ± 0,03
	<b>NaM</b>	8,3	91,2 ± 1,3	99,79 ± 0,03
<b>FTPT3</b>	<b>FIMO</b>	8,2	86 ± 3	99,65 ± 0,08
	<b>IMO</b>	7,1	76,8 ± 1,6	99,38 ± 0,05
	<b>FEMO</b>	6,3	72,4 ± 1,4	99,24 ± 0,04
	<b>NaM</b>	7,0	90 ± 2	99,76 ± 0,06
<b>5SICS</b>	<b>NaM</b>	8,1	77,1 ± 0,7	99,00 ± 0,02

<sup>a</sup> Сохранение и точность воспроизведения определяли, как описано в примере 1.

<sup>b</sup> UBP-сохранение ниже 50%, и, таким образом, точность воспроизведения оценивается в <85%.

Данные скрининга позволяют предположить, что несколько пар, образованных между dTPT3 и ранее неисследованными пиридиновыми производными  $\alpha 6$ , достаточно хорошо реплицировались. Таким образом, авторы изучили в трех повторах амплификацию ДНК, содержащей данные UBP, используя OneTaq и 16 циклов амплификации с временами удлинения в 1 мин (табл. 5). Пары, образованные между dTPT3 и dTOK580, dTOK582 или dTOK586, реплицировались плохо. Тем не менее, пары, образованные между dTPT3 и dTOK588, dTOK581, dTOK576 и dTOK587, амплифицировались с сохранением в 62%, 65%, 85% и 94% соответственно.

Таблица 5

Данные по амплификации и точности воспроизведения dTPT3 по отношению к пиридиновым производным из группы  $\alpha 6$ ; DM5, MMO2, DMO, и NaM также были охарактеризованы для сравнения скаффолдов

d $\beta$ TP	d $\alpha$ TP	Амплификация	Сохранение, %	Точность воспроизведения, %
<b>TPT3</b>	<b>TOK576</b>	780	85,20 $\pm$ 1,12	98,35 $\pm$ 0,14
<b>TPT3</b>	<b>TOK580</b>	1056	< 50 <sup>b</sup>	< 85 <sup>b</sup>
<b>TPT3</b>	<b>TOK581</b>	1034	65,07 $\pm$ 0,15	95,80 $\pm$ 0,02
<b>TPT3</b>	<b>TOK582</b>	1240	< 50 <sup>b</sup>	< 85 <sup>b</sup>
<b>TPT3</b>	<b>TOK586</b>	948	< 50 <sup>b</sup>	< 85 <sup>b</sup>
<b>TPT3</b>	<b>TOK587</b>	818	93,81 $\pm$ 1,35	99,34 $\pm$ 0,15
<b>TPT3</b>	<b>TOK588</b>	666	61,98 $\pm$ 7,09	94,99 $\pm$ 1,13
<b>TPT3</b>	<b>DM5</b>	- <sup>a</sup>	- <sup>a</sup>	- <sup>a</sup>
<b>TPT3</b>	<b>MMO2</b>	1096	90,95 $\pm$ 3.63	99,06 $\pm$ 0,40
<b>TPT3</b>	<b>DMO</b>	864	84,02 $\pm$ 1,92	98,23 $\pm$ 0,23
<b>TPT3</b>	<b>NaM</b>	1004	99,23 $\pm$ 1,12	99,92 $\pm$ 0,11

<sup>a</sup> В данном образце амплификация не была обнаружена.

<sup>b</sup> Сохранение не встречающейся в природе пары оснований было ниже 50%, и точность воспроизведения, таким образом, оценивается менее чем на 85%.

Наконец, данные скрининга позволили предположить, что пары, образованные между dTPT3 и d2MN или dDM2, достаточно хорошо реплицируются, несмотря на то, что ни d2MN, ни dDM2 не обладают предположительно существенным акцептором водородной связи в орто-положении. Таким образом, данные пары были дополнительно изучены с помощью 16 циклов амплификации с OneTaq или только с Taq и с временами удлинения в 1 мин или 10 с (табл. 6). Только с одной Taq наблюдалось лишь слабое сохранение. Однако с OneTaq, сохранение было лучше для обеих пар. Сохранение пары dTPT3-dDM2 составляет 58 и 69% с временами удлинения в 1 мин и 10 с соответственно. Примечательно, dTPT3-d2MN амплифицируется с сохранениями в 96 и 94% с временами удлинения в 1 мин и 10 с соответственно.

Таблица 6

Данные по амплификации и точности воспроизведения dTPT3 по отношению к d2MN или dDM2,  $\alpha$ -аналогам без акцептора Н-связи в орто-положении

Фермент	d $\beta$ TP	d $\alpha$ TP	Время удлинения	Амплификация	Сохранение, %	Точность воспроизведения за дупликацию, %
OneTaq	<b>TPT3</b>	<b>2MN</b>	1 мин	880	95,54 $\pm$ 1,55	99,53 $\pm$ 0,17
			10 с	224	93,53 $\pm$ 1,42	99,15 $\pm$ 0,20
	<b>TPT3</b>	<b>DM2</b>	1 мин	1420	57,92 $\pm$ 6,02	94,89 $\pm$ 0,94
			10 с	376	68,46 $\pm$ 4,34	95,65 $\pm$ 0,72
Taq	<b>TPT3</b>	<b>2MN</b>	1 мин	- <sup>a</sup>	-	-
			10 с	- <sup>a</sup>	-	-
	<b>TPT3</b>	<b>DM2</b>	1 мин	334	< 50 <sup>b</sup>	< 85 <sup>b</sup>
			10 с	266	< 50 <sup>b</sup>	< 85 <sup>b</sup>

<sup>a</sup> В данном образце амплификация не была обнаружена.

<sup>b</sup> Сохранение не встречающейся в природе пары оснований было ниже 50%, и точность воспроизведения, таким образом, оценивается менее чем на 85%.

Обсуждение. В настоящем изобретении проводили скрининг на основе ПЦР для выявления наиболее перспективных UBP. Чтобы повысить SAR-содержание скрининга, были включены семь новых  $\alpha$ -производных, которые основаны на пиридинном скаффолде с различными заместителями в орто- и пара-положениях относительно гликозидной связи.

Данные по взаимосвязи структура-активность. Даже при перmissive условиях, когда присутствовала экзонуклеолитическая корректирующая активность и времена удлинения составляли 1 мин, только смешанные группы  $\alpha$ -аналогов и  $\beta$ -аналогов показали значительные уровни сохранения, демонстрируя, что эффективная репликация требует спаривания  $\alpha$ -скаффолда с  $\beta$ -скаффолдом. Однако d $\beta$ -группы, которые показали высокую степень сохранения, были только  $\beta$ 5 и  $\beta$ 6. Это свидетельствует о привилегированном статусе d5SICS/dTPT3-подобного скаффолда относительно всех остальных рассмотренных. Доминирующий вклад в высокую степень сохранения с группой  $\beta$ 5, как оказалось, обусловлен не парами с участием d5SICS, а парами с участием dSICS и, в меньшей степени, dSNICS. Например, при любых условиях, dSICS-dNaM лучше реплицировалась, чем d5SICS-dNaM. d5SICS получили в результате оптимизации dSICS для спаривания с dMMO2; по-видимому, увеличенная масса dNaM делает добавленную метильную группу вредной. Кроме того, dSNICS-dNaM реплицируется почти так же хорошо (с OneTaq) или лучше (с Taq), чем d5SICS-dNaM, что позволяет предполагать, что 6-аза-заместитель оптимизирует UBP-синтез, облегчая вставку не встречающегося в природе трифосфата напротив dNaM или повышая эффективность, с которой не встречающийся в природе нуклеотид служит матрицей для вставки dNaMTP. Наконец, dSNICS-dFEMO также лучше реплицируются, чем d5SICS-dNaM, но только в присутствии корректирующей активности, что позволяет предполагать, что в то время как вставка трифосфата может быть менее эффективной, повышение эффективности результатов удлинения приводит к общему увеличению точности воспроизведения. Доминирующий вклад в сохранение с высокой точностью воспроизведения с группой  $\beta$ 6 был обеспечен dTPT3 и dFTPT3. В общем, оба хорошо образовывали пары с dNaM, dFEMO, dFIMO или dIMO. dTPT3 образовывал пары особенно хорошо с dFIMO и dIMO, что позволяет предполагать, что иод-заместитель в пара-положении опосредует благоприятные взаимодействия, и он образовывал пары также хорошо с dFEMO и особенно с dNaM, когда присутствовала экзонуклеазная активность. dFTPT3 хорошо образовывал пары с dIMO или dFEMO в присутствии экзонуклеазной активности, а также с dFIMO и dNaM в ее отсутствии.

В то время как заместитель азота  $\alpha$ -аналогов на пиридиновой основе (группа  $\alpha$ 6), в целом, пагубен для репликации, более подробный анализ UBP, образованных с dTPT3, выявил несколько тенденций. Метил-, хлор-, или amino-заместитель в орто-положении к С-гликозидной связи приводил к плохо реп-

лизируемым парам, предположительно, из-за недостаточного удлинения после введения не встречающегося в природе трифосфата. Ортометокси-заместитель в dТОК581 приводил к улучшению репликации, предположительно, из-за его способности как вступать в гидрофобное взаимодействие с матрицей во время UBP-синтеза, так и служить акцептором водородной связи от донора водородной связи, связанного с полимеразой, во время удлинения. Полученные данные также показали, что метилсульфанильный заместитель в орто-положении в dТОК588, dТОК576 и, особенно, dТОК587 приводит к лучшей репликации. Такое улучшение, вероятно, обусловлено более оптимизированным компромиссом между возможностью гидрофобной укладки и способностью принимать водородную связь от полимеразы на конце праймера. Кроме того, заместитель в пара-положении в данной серии производных может способствовать эффективной репликации, с бром-заместителем, который является лучшим, после чего следует вторая по эффективности метилсульфанильная группа, и затем, наконец, простая метильная группа. Когда dТОК587, с его сочетанием ортометилсульфанильного и парабром-заместителей, спаривался с dTPT3, полученный UBP реплицировался OneTaq и с временами удлинения в 1 мин с точностью воспроизведения (рассчитывается из уровня сохранения) в 99,3%, что немного лучше, чем d5SICS-dMMO2 в аналогичных условиях. Несомненно, подобные ортометилсульфанильный и парабром-заместители должны быть изучены с более эффективно реплицируемыми  $\alpha$ -подобными скаффолдами, такими как dFIMO и dNaM.

Репликация пар, образованных между dTPT3 и d2MN или dDM2, также заслуживает обсуждения. ДНК, содержащая такие пары, не амплифицируется Taq в одиночку, но хорошо амплифицируется OneTaq. Данный результат был неожиданным, потому что ни d2MN, ни dDM2 не обладает акцептором ортоводородной связи, которая, согласно высказанному предположению, необходима для удлинения зарождающегося конца (природного или неприродного) праймера. В частности, когда нуклеотид расположен на растущем конце праймера, акцептор водородной связи размещается в развивающемся малой бороздке, где он принимает Н-связь от полимеразы, и данная Н-связь, как полагают, необходима для правильного выравнивания конца. При амплификации с OneTaq и времени удлинения в 1 мин, dTPT3-d2MN реплицируется с точностью воспроизведения 99,5%, которая падает до 99,1%, только когда время удлинения сокращено до 10 с. Отсутствие амплификации в отсутствие корректировочной активности, в сочетании только с небольшим снижением, наблюдаемым в присутствии корректировочной активности, когда времена удлинения были сокращены, позволяет предполагать, что удивительно высокая точность амплификации ДНК, содержащей dTPT3-d2MN, является результатом селективного удлинения UBP относительно ошибочных пар. Это говорит о том, что в случае отсутствия акцептора ортоводородной связи ошибочная пара мешает удлинению в большей степени, чем удлинение UBP.

Усилия по расширению генетического алфавита. В целом, данные подтверждают, что dTPT3-dNaM является наиболее перспективным UBP из тех, которые были протестированы. Тем не менее, пары, образованные между dTPT3 и dFEMO, dFIMO или dIMO или между dFTPT3 и dNaM, dFEMO, dFIMO или dIMO, являются также перспективными. В дополнение к наиболее перспективным UBP, отмеченным выше, следует отметить, что примечательное число дополнительных новых пар реплицируются только с умеренно сниженной точностью воспроизведения, или реплицируются с высокой точностью воспроизведения, когда амплификация осуществляется при менее строгих условиях (табл. 7). Наряду с наиболее эффективно реплицируемыми UBP, данные пары обеспечивают широкий спектр скаффолдов с различными физико-химическими свойствами для дальнейших усилий по оптимизации, где различные физико-химические свойства, как ожидается, смогут придавать задействованным нуклеотидам различные свойства, имеющие значение для фармакинетики.

Таблица 7

dβTP	dαTP	Сохранение (%)
<b>SICS</b>	<b>NaM</b>	99 <sup>a</sup>
<b>SICS</b>	<b>FEMO</b>	92 <sup>b</sup>
<b>SNICS</b>	<b>NaM</b>	90 <sup>a</sup>
<b>SNICS</b>	<b>FEMO</b>	95 <sup>b</sup>
<b>SNICS</b>	<b>IMO</b>	88 <sup>b</sup>
<b>TPT3</b>	<b>NMO</b>	89 <sup>b</sup>
<b>TPT3</b>	<b>ZMO</b>	86 <sup>b</sup>
<b>TPT3</b>	<b>CIMO</b>	90 <sup>b</sup>
<b>TPT3</b>	<b>QMO</b>	90 <sup>b</sup>
<b>TPT3</b>	<b>CNMO</b>	91 <sup>b</sup>
<b>FTPT3</b>	<b>NMO</b>	94 <sup>a</sup>
<b>FTPT3</b>	<b>ZMO</b>	88 <sup>a</sup>
<b>FTPT3</b>	<b>CIMO</b>	97 <sup>a</sup>
<b>FTPT3</b>	<b>QMO</b>	87 <sup>a</sup>
<b>FTPT3</b>	<b>CNMO</b>	94 <sup>a</sup>

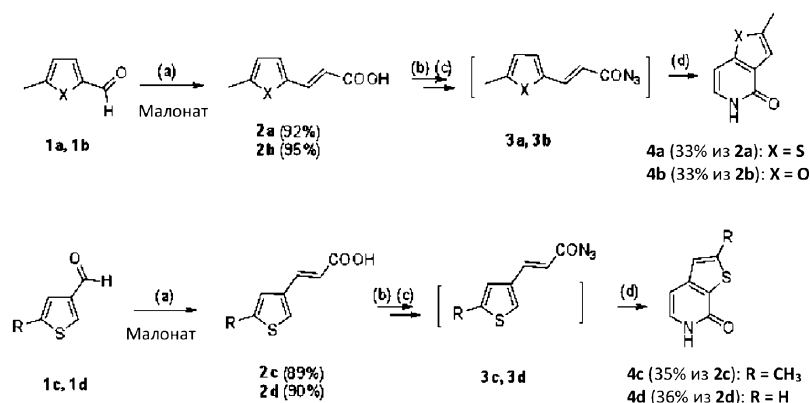
<sup>a</sup> Условия ПСР: 100 пг матрицы D8, амплифицированной в течение 16 циклов с Taq-полимеразой в условиях термоциклирования: начальная денатурация при 96°C в течение 1 мин, 96°C в течение 30 с, 60°C в течение 15 с, 68°C в течение 60 с.

<sup>b</sup> Условия ПСР: 10 пг матрицы D6, амплифицированной в течение 24 циклов с OneTaq-полимеразой в условиях термоциклирования: начальная денатурация при 96°C в течение 1 мин, 96°C в течение 5 с, 60°C в течение 5 с, 68°C в течение 10 с.

Пример 2. Общая методика синтеза трифосфата.

Протонную губку (1,3 экв.) и свободное производное нуклеозида (1,0 экв.) растворили в сухом триметилфосфате (40 экв.) и охладили до -15°C в атмосфере азота. Свежеперегнанный POCl<sub>3</sub> (1,3 экв.) добавляли по каплям и полученную смесь перемешивали при -10°C в течение 2 ч. Добавляли трибутиламин (6,0 экв.) и раствор трибутиламмонийпирофосфата (5,0 экв.) в диметилформамиде (0,5M). В течение 30 мин реакционную смесь оставляли медленно нагреваться до 0°C и затем гасили добавлением 0,5M водного Et<sub>3</sub>NH<sub>2</sub>CO<sub>3</sub> (TEAB) pH 7,5 (2 об. экв.). Смесь дважды разбавляли H<sub>2</sub>O, и продукт отделяли на колонке с DEAE-сефадексом (GE Healthcare) с градиентом элюции от 0 до 1,2M TEAB, выпаривали и перегоняли совместно с H<sub>2</sub>O (3×). Дополнительную очистку выполняли обращенно-фазовой (C18) ВЭЖХ (от 0 до 35% CH<sub>3</sub>CN в 0,1M TEAB, pH 7,5) (выход от 10 до 31%).

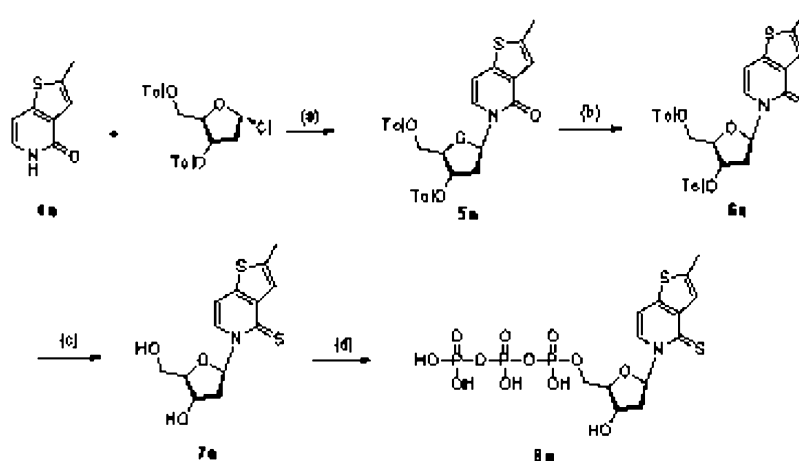
Схема S1



(a) Пиперидин, Py, 100°C, 12 ч, затем кипячение с обратным холодильником в течение 1 ч; (b) SOCl<sub>2</sub>, DMF, CHCl<sub>3</sub>, кипячение с обратным холодильником в течение 3 ч, (c) NaN<sub>3</sub>, 1,4-диоксан, H<sub>2</sub>O, 5°C, 0,5 ч; (d) дифениловый эфир, 230°C, 1 ч.

Аналоги нуклеиновых оснований 4a, 4b, 4c и 4d синтезировали на основе литературных способов 1, 2, как показано на схеме S1. Вкратце, конденсация альдегида (1a-d) с малоновой кислотой при 100°C в пиридине в качестве растворителя и пиперидине в качестве катализатора в течение 12 ч, с последующим кипячением с обратным холодильником в течение 1 ч, давала соответствующие промежуточные продукты акриловой кислоты (2a-d). Хлорирование данных кислот с тионилхлоридом в хлороформе в присутствии DMF давало ацильные хлориды, которые не были очищены, но могли бы быть использованы непосредственно в получении азидов (3a-d). Соединения 3a-d были получены в двухфазной смеси 1,4-диоксана и воды при 5°C с азидом натрия. Неочищенные смеси 3a-d в растворах  $\text{CHCl}_3$  добавляли порциями к дифенилэфру и нагревали до 230°C с получением изоцианатов, которые прошли последующую внутримолекулярную циклизацию до бициклических систем 4a-d с 6-5-членными конденсированными кольцами.

Схема S2



(a) N,O-бис(TMS)ацетамид,  $\text{SnCl}_4$  (1M в  $\text{CH}_2\text{Cl}_2$ ),  $\text{CH}_2\text{Cl}_2$ , 3 ч, 45%; (b) реагент Лавессона (Lawesson), толуол, кипячение с обратным холодильником, в течение ночи, 58%; (c) 30% NaOMe, MeOH, комн. темп., 1 ч, 85%; (d) протонная губка,  $\text{POCl}_3$ ,  $\text{Bu}_3\text{N}$ ,  $\text{Bu}_3\text{NPPi}$ ,  $(\text{MeO})_3\text{P}$ , DMF, -20°C, 31%.

Соединение 5a. К раствору 4a (54 мг, 0,33 ммоль) в  $\text{CH}_2\text{Cl}_2$  (8 мл) при комнатной температуре в атмосфере азота добавляли бис(триметилсилил)ацетамид (83 мг, 0,39 ммоль). После перемешивания в течение 40 мин добавляли 3,5-бис(толуол)-2-дезоксирибозилхлорид (196 мг, 0,39 ммоль). Реакционную смесь охлаждали до 0°C и добавляли по каплям  $\text{SnCl}_4$  (1,0M в  $\text{CH}_2\text{Cl}_2$ , 160 мкл, 0,16 ммоль). Полученный раствор перемешивали в течение 2 ч при комнатной температуре. Реакционную смесь разбавляли EtOAc, гасили насыщенным водным раствором  $\text{NaHCO}_3$ , экстрагировали этилацетатом, сушили, фильтровали и выпаривали. Неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 5a в виде белой пены (77 мг, 0,15 ммоль, 45%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7,97-6,82 (м, 11H, Ar-H), 6,44 (д,  $J = 7,5$  Гц, 1H, H-1'), 5,63 (д,  $J = 6,5$  Гц, 1H, H-3'), 4,76-4,68 (м, 2H, H-5'a, 5'b), 4,59 (д,  $J = 2,5$  Гц, H-4'), 2,89 (дд,  $J = 1,5, 0,5$  Гц, 1H, H-2'a), 2,59 (с, 3H, Ar-CH<sub>3</sub>), 2,43 (с, 3H, Ar-CH<sub>3</sub>), 2,36-2,30 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  166,6, 166,5, 158,5, 147,2, 144,8, 144,6, 140,0, 131,1, 130,3, 130,0, 129,9, 129,7, 127,1, 126,9, 125,6, 122,8, 102,7, 85,9, 83,3, 75,6, 64,8, 39,6, 22,1, 16,1.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{29}\text{H}_{28}\text{NO}_6\text{S}$  ( $\text{M} + \text{H}^+$ ) 518,1632, найдено 518,1621.

Соединение 6a. Соединение 5a (27 мг, 0,052 ммоль) высушивали 3 совместными испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1 мл). Добавляли реагент Лавессона (41,5 мг, 0,10 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 6a в виде желтой пены (16 мг, 0,03 ммоль, 58%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  8,00-7,89 (м, 4H, Ar-H), 7,71 (м, 1H, Ar-H), 7,49-7,48 (м, 1H, H-1'), 7,29-7,21 (м, 4H, Ar-H), 7,65-7,62 (м, 1H, Ar-H), 6,90 (д,  $J = 7,5$  Гц, 1H, Ar-H), 5,64-5,62 (м, 1H, H-4'), 4,85-4,74 (м, 2H, H-5'a), 4,68-4,67 (м, 1H, H-5'b), 3,38-3,34 (м, 1H, H-3'), 2,26 (с, 3H), 2,44 (с, 3H), 2,41 (с, 3H), 2,28-2,22 (м, 1H).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  174,6, 166,6, 144,9, 144,8, 142,9, 142,7, 142,3, 130,3, 130,0, 129,7, 127,9, 127,0, 126,9, 126,8, 108,3, 100,0, 91,4, 84,0, 74,9, 64,5, 39,3, 22,2, 22,1, 16,3.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{29}\text{H}_{28}\text{NO}_5\text{S}_2$  ( $\text{M} + \text{H}^+$ ) 534,1403, найдено 534,1404.

Соединение 7a. К раствору 6a (20 мг, 0,037 ммоль) в метаноле (1,0 мл) добавляли по каплям 30% NaOMe (8,66 мг, 0,16 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь затем концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле ( $\text{MeOH}/\text{CH}_2\text{Cl}_2$ ) с получением соединения 7a в виде желтого пенообразного вещества (9,2 мг, 0,031 ммоль, 85%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  8,36 (д,  $J = 4$  Гц, 1H, Ar-H), 7,58 (д,  $J = 1$  Гц, 1H, Ar-H), 7,35 (т,  $J = 4$  Гц, 1H, H-1'), 7,22 (д,  $J = 8$  Гц, 1H, Ar-H), 4,07-4,06 (м, 1H, H-4'), 4,07 (д,  $J = 4$  Гц, 1H, H-3'), 3,80 (дд,  $J = 24, 4$  Гц, 2H, H-5'a, b), 2,79-2,76 (м, 1H, H-2'a), 2,13-2,08 (м, 1H, H-2'b).

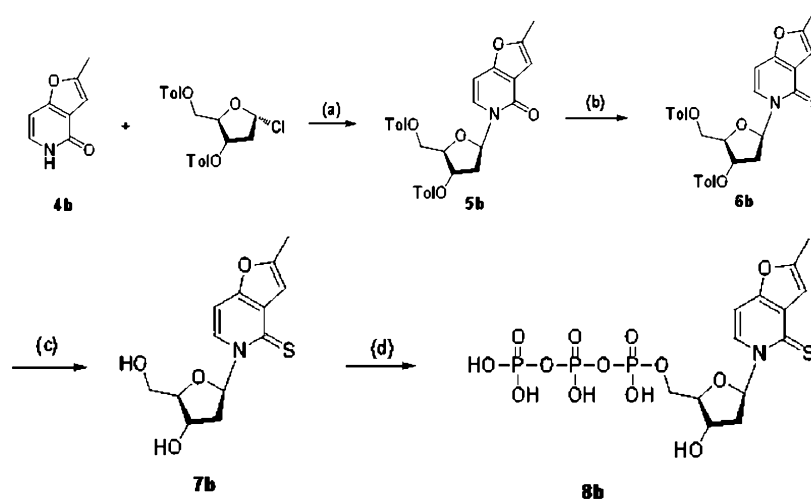
$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  173,60, 143,29, 142,26, 142,23, 129,35, 126,11, 108,01, 91,14, 88,44, 70,37, 61,35, 41,59, 14,81.

HRMS (ESI +)  $m/z$  вычислено для  $\text{C}_{13}\text{H}_{16}\text{NO}_3\text{S}_2$  ( $M + \text{H}^+$ ) 298,0566, найдено 298,0569.

Соединение 8a (dTPPTP). Соединение 8a (11,2 мг, 20,8 мкмоль, 31%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с 7a (20 мг, 67,3 мкмоль).

$^{31}\text{P}$  ЯМР (162 МГц,  $\text{D}_2\text{O}$ )  $\delta$  -10,3 (д,  $J = 19,8$  Гц, у-Р), -10,9 (д,  $J = 20,1$  Гц, а-Р), -22,8 (т,  $J = 19,4$  Гц, ф-Р). MS (MALDI-TOF-, матрица: 9-аминоакридин) ( $m/z$ ):  $[\text{M}-\text{H}]^-$  вычислено для  $\text{C}_{13}\text{H}_{17}\text{NO}_{12}\text{P}_3\text{S}_2$ , 536,3, найдено, 536,7.

Схема S3



(a) N,O-бис(ТМС)ацетамид,  $\text{SnCl}_4$  (1M в  $\text{CH}_2\text{Cl}_2$ ),  $\text{CH}_2\text{Cl}_2$ , 3 ч, 41%; (b) реагент Лавессона, толуол, кипячение с обратным холодильником, в течение ночи, 52%; (c) 30% NaOMe, MeOH, комн. темп., 1 ч, 85%; (d) протонная губка,  $\text{POCl}_3$ ,  $\text{Bu}_3\text{N}$ ,  $\text{Bu}_3\text{NPPi}$ ,  $(\text{MeO})_3\text{P}$ , DMF,  $-20^\circ\text{C}$ , 21%.

Соединение 5b. К раствору 4b (100 мг, 0,67 ммоль) в  $\text{CH}_2\text{Cl}_2$  (8 мл) при комнатной температуре в атмосфере азота добавляли бис(триметилсилил)ацетамид (165 мг, 0,81 ммоль). После перемешивания в течение 40 мин добавляли 3,5-бис(толуол)-2-дезоксирибозилхлорид (292 мг, 0,81 ммоль). Реакционную смесь охлаждали до  $0^\circ\text{C}$  и добавляли по каплям  $\text{SnCl}_4$  (1,0M в  $\text{CH}_2\text{Cl}_2$ , 200 мкл, 0,2 ммоль). Полученный раствор перемешивали в течение 2 ч при комнатной температуре. Реакционную смесь разбавляли EtOAc, гасили насыщенным водным раствором  $\text{NaHCO}_3$ , экстрагировали этилацетатом, сушили, фильтровали и выпаривали. Неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 5b в виде белой пены (137 мг, 0,27 ммоль, 41%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7,99 (д,  $J = 8,1$  Гц, 2H, Ar-H), 7,93 (д,  $J = 8,1$  Гц, 2H, Ar-H), 7,55 (д,  $J = 7,7$  Гц, 1H, Ar-H), 7,33 - 7,28 (м, 2H, Ar-H), 7,27 - 7,20 (м, 2H, Ar-H), 6,82 (дд,  $J = 8,3, 5,6$  Гц, 1H, Ar-H), 6,57 (д,  $J = 0,9$  Гц, 1H, Ar-H), 6,41 (д,  $J = 7,7$  Гц, 1H, H-1'), 5,68 - 5,61 (м, 1H, H-4'), 4,75 (дд,  $J = 12,1, 3,4$  Гц, 2H, H-5'a, b), 4,62 (q,  $J = 3,1$  Гц, 1H, H-3'), 2,94 (ддд,  $J = 14,3, 5,6, 1,7$  Гц, 1H, H-2'a), 2,48 - 2,39 (s,  $3 \times 3\text{H}$ , Ar-CH<sub>3</sub>), 2,36 - 2,26 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  166,6, 166,5, 159,2, 159,1, 154,5, 144,8, 144,6, 130,2, 130,0, 129,7, 127,5, 127,1, 126,9, 117,5, 103,3, 96,6, 86,0, 83,2, 75,5, 64,7, 39,8, 22,1, 14,1.

HRMS (ESI +)  $m/z$  вычислено для  $\text{C}_{29}\text{H}_{28}\text{NO}_7$  ( $M + \text{H}^+$ ) 502,1860, найдено 502,1885.

Соединение 6b. Соединение 5b (29 мг, 0,056 ммоль) высушивали 3 сопутствующими испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1 мл). Добавляли реагент Лавессона (41,5 мг, 0,10 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 6b в виде желтой пены (15 мг, 0,029 ммоль, 52%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  8,10-7,89 (м, 5H, Ar-H), 7,52-7,48 (м, 1H, H-1'), 7,29-7,22 (м, 4H, Ar-H), 6,8 (д,  $J = 1$  Гц, 1H, Ar-H), 6,73 (д,  $J = 7,5$  Гц, 1H, Ar-H), 5,65-5,62 (м, 1H, H-4'), 4,84-4,74 (м, 2H, H-5'a, b), 4,67-4,65 (м, 1H, H-3'), 3,36-3,32 (м, 1H, H-2'a), 2,44 (с, 3H, Ar-CH<sub>3</sub>), 2,43 (с, 3H, с, 3H, Ar-CH<sub>3</sub>), 2,41 (с, 3H, с, 3H, Ar-CH<sub>3</sub>), 2,27-2,21 (м, 1H, H-2'b).  $^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  166,6, 156,9, 153,9, 144,8, 130,3, 130,0, 129,8, 129,7, 127,9, 106,4, 96,0, 83,9, 56,6, 39,5, 22,1, 12,6.

HRMS (ESI +)  $m/z$  вычислено для  $\text{C}_{29}\text{H}_{28}\text{NO}_6\text{S}$  ( $M + \text{H}^+$ ) 518,1632, найдено 518,1638.

Соединение 7b. К раствору 6b (20 мг, 0,039 ммоль) в метаноле (1,0 мл) добавляли по каплям 30% NaOMe (8,66 мг, 0,16 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь затем концентрировали и неочищенный



продукт подвергали колоночной хроматографии на силикагеле (MeOH/CH<sub>2</sub>Cl<sub>2</sub>) с получением соединения 7b в виде желтой пены (9,3 мг, 0,033 ммоль, 85%).

<sup>1</sup>H ЯМР (500 МГц, CD<sub>3</sub>OD) δ 8,57 (д, J = 5 Гц, 1H, Ar-H), 7,42 (т, J = 4 Гц, 1H, H-1'), 7,13 (д, J = 7,5 Гц, 1H, Ar-H), 6,80 (с, 1H, Ar-H), 4,50-4,47 (м, 1H, H-4'), 4,12 (д, J = 3,5 Гц, 1H, H-3'), 3,95 (дд, J = 30, 3 Гц, 2H, H-5'a, b), 2,81-2,77 (м, 1H, H-2'a), 2,50 (с, 3H, Ar-CH<sub>3</sub>), 2,18-2,14 (м, 1H, H-2'b).

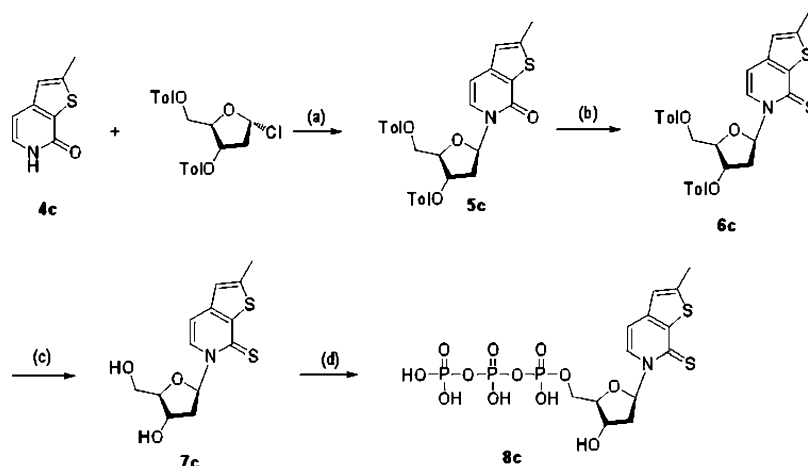
<sup>13</sup>C ЯМР (125 МГц, CD<sub>3</sub>OD) δ 172,9, 157,2, 154,5, 132,7, 131,3, 105,4, 100,8, 90,9, 88,5, 70,3, 61,3, 41,8, 12,7.

HRMS (ESI +) м/з вычислено для C<sub>13</sub>H<sub>16</sub>NO<sub>4</sub>S (M + H<sup>+</sup>) 282,0795, найдено 282,0790.

Соединение 8b (dFPT1TP). Соединение 8b (3,7 мг, 7,1 мкмоль, 10%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с 7b (20 мг, 71,2 мкмоль).

<sup>31</sup>P ЯМР (162 МГц, D<sub>2</sub>O) δ -10,4 (д, J = 20,0 Гц, у-Р), -10,9 (д, J = 19,4 Гц, а-Р), -22,8 (т, J = 20,0 Гц, ф3-Р). MS (MALDI-TOF-, матрица: 9-аминоакридин) (м/з): [M-H]<sup>-</sup> вычислено для C<sub>13</sub>H<sub>17</sub>NO<sub>13</sub>P<sub>3</sub>S, 520,3, найдено, 520,1.

Схема S4



(a) N,O-бис(ТМС)ацетамид, SnCl<sub>4</sub> (1M в CH<sub>2</sub>Cl<sub>2</sub>), CH<sub>2</sub>Cl<sub>2</sub>, 3 ч, 40%; (b) реагент Лавессона, толуол, кипячение с обратным холодильником, в течение ночи, 31%; (c) 30% NaOMe, MeOH, комн. темп., 1 ч, 81%; (d) протонная губка, POCl<sub>3</sub>, Bu<sub>3</sub>N, Bu<sub>3</sub>NPPI, (MeO)<sub>3</sub>P, DMF, -20°C, 15%.

Соединение 5c. К раствору 4с (46 мг, 0,28 ммоль) в CH<sub>2</sub>Cl<sub>2</sub> (8 мл) при комнатной температуре в атмосфере азота добавляли бис(триметилсилил)ацетамид (66 мг, 0,33 ммоль). После перемешивания в течение 40 мин добавляли 3,5-бис(толуол)-2-дезоксирибозилхлорид (120 мг, 0,33 ммоль). Реакционную смесь охлаждали до 0°C и добавляли по каплям SnCl<sub>4</sub> (1,0M в CH<sub>2</sub>Cl<sub>2</sub>, 140 мкл, 0,14 ммоль). Полученный раствор перемешивали в течение 2 ч при комнатной температуре. Реакционную смесь разбавляли EtOAc, гасили насыщенным водным раствором NaHCO<sub>3</sub>, экстрагировали этилацетатом, сушили, фильтровали и выпаривали. Неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 5с в виде белой пены (58 мг, 0,11 ммоль, 40%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 7,98-7,90 (м, 4H, Ar-H), 7,53 (д, J = 7,4 Гц, 1H, Ar-H), 7,27- 7,21 (м, 4H, Ar-H), 6,83-6,82 (м, 2H, Ar-H), 6,44 (д, J = 7,5 Гц, 1H, H-1'), 5,63 (д, J = 6,5 Гц, 1H, H-4'), 4,76 - 4,60 (м, 2H, H-5'a, b), 4,59 (д, J = 2,5 Гц, 1H, H-3'), 2,89 (дд, J = 13, 5,5 Гц, H-2'a), 2,59 (с, 3H, Ar-CH<sub>3</sub>), 2,43 (с, 3H, Ar-CH<sub>3</sub>), 2,40 (с, 3H, Ar-CH<sub>3</sub>), 2,37-2,30 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 166,5, 158,0, 149,9, 146,3, 144,8, 144,6, 130,3, 130,0, 129,7, 128,8, 127,3, 122,8, 103,7, 100,0, 85,8, 83,2, 75,5, 64,8, 39,5, 22,1, 16,7.

HRMS (ESI +) м/з вычислено для C<sub>29</sub>H<sub>28</sub>NO<sub>6</sub>S (M + H<sup>+</sup>) 518,1632, найдено 518,1631.

Соединение 6с. Соединение 5с (50 мг, 0,097 ммоль) высушивали 3-мя сопутствующими испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1,5 мл). Добавляли реагент Лавессона (83 мг, 0,20 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 6с в виде желтой пены (16 мг, 0,03 ммоль, 31%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 8,13-7,97 (м, 5H, Ar-H), 7,52-7,49 (м, 1H, H-1'), 7,37-7,29 (м, 4H, Ar-H), 6,99 (д, J = 1 Гц, 1H, Ar-H), 6,91 (д, J = 7,5 Гц, 1H, Ar-H), 5,73-5,71 (м, 1H, H-4'), 4,91-4,82 (м, 2H, H-5'a, b), 4,76-4,74 (м, 1H, H-3'), 3,41-3,37 (м, 1H, H-2'a), 2,68 (с, 3H, Ar-CH<sub>3</sub>), 2,52 (с, 3H, Ar-CH<sub>3</sub>), 2,49 (с, 3H, Ar-CH<sub>3</sub>), 2,39-2,34 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 172,1, 166,6, 154,0, 144,9, 144,7, 140,4, 130,3, 130,0, 129,7, 129,6, 127,0, 126,8, 122,7, 109,0, 91,2, 83,9, 75,0, 64,5, 39,4, 22,2, 22,1 17,0.

HRMS (ESI +) м/з вычислено для C<sub>29</sub>H<sub>28</sub>NO<sub>5</sub>S<sub>2</sub> (M + H<sup>+</sup>) 534,1403, найдено 534,1406.

Соединение 7с. К раствору 6с (20 мг, 0,037 ммоль) в метаноле (1,0 мл) добавляли по каплям 30%

NaOMe (8,66 мг, 0,16 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь затем концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле (MeOH/CH<sub>2</sub>Cl<sub>2</sub>) с получением соединения 7с в виде желтой пены (8,9 мг, 0,03 ммоль, 81%).

<sup>1</sup>H ЯМР (500 МГц, CD<sub>3</sub>OD) δ 8,48, (д, J = 7,5 Гц, 1H, Ar-H), 7,42 (т, J = 5 Гц, 1H, H-1'), 7,20 (д, J = 5 Гц, 1H, Ar-H), 7,12 (с, 1H, Ar-H), 4,51-4,48 (м, 1H, H-4'), 4,13 (д, J = 5 Гц, 1H, H-3'), 3,95 (дд, J = 30, 5 Гц, 2H, H-5'a, b), 2,81-2,78 (м, 1H, H-2'a), 2,67 (с, 3H, Ar-CH<sub>3</sub>), 2,21-2,16 (м, 1H, H-2'b).

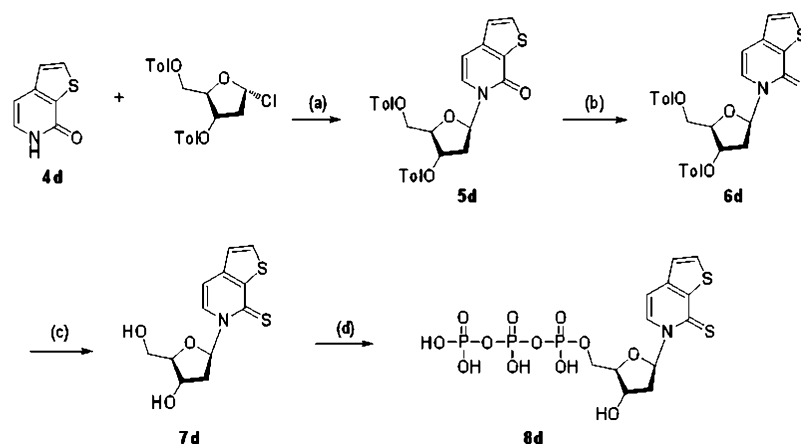
<sup>13</sup>C ЯМР (125 МГц, CD<sub>3</sub>OD) δ 171,1, 154,0, 144,1, 141,1, 131,1, 122,7, 108,8, 90,9, 88,5, 70,5, 61,4, 41,7, 15,4.

HRMS (ESI<sup>+</sup>) m/z вычислено для C<sub>13</sub>H<sub>16</sub>NO<sub>3</sub>S<sub>2</sub> (M + H<sup>+</sup>) 298,0566, найдено 298,0566.

Соединение 8с. Соединение 8с (10,8 мг, 20,2 мкмоль, 30%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с 7с (20 мг, 67,3 мкмоль).

<sup>31</sup>P ЯМР (162 МГц, D<sub>2</sub>O) δ -10,8 (д, J = 19,8 Гц, у-Р), -11,5 (д, J = 20,1 Гц, а-Р), -23,3 (т, J = 20,1 Гц, ф3-Р). MS (MALDI-TOF-, матрица: 9-аминоакридин) (m/z): [M-H]<sup>-</sup> вычислено для C<sub>13</sub>H<sub>17</sub>NO<sub>12</sub>P<sub>3</sub>S<sub>2</sub>, 536,3, найдено, 536,1.

Схема S5



(a) N,O-бис(ТМС)ацетамид, SnCl<sub>4</sub> (1M в CH<sub>2</sub>Cl<sub>2</sub>), CH<sub>2</sub>Cl<sub>2</sub>, 3 ч, 39%; (b) реагент Лавессона, толуол, кипячение с обратным холодильником, в течение ночи, 33%; (c) 30% NaOMe, MeOH, комн. темп., 1 ч, 82%; (d) протонная губка, POCl<sub>3</sub>, Bu<sub>3</sub>N, Bu<sub>3</sub>NPPi, (MeO)<sub>3</sub>P, DMF, -20°C, 30%.

Соединение 5d. К раствору 4d (200 мг, 1,32 ммоль) в CH<sub>2</sub>Cl<sub>2</sub> (8 мл) при комнатной температуре в атмосфере азота добавляли бис(триметилсилил)ацетамид (298 мг, 1,46 ммоль). После перемешивания в течение 40 мин добавляли 3,5-бис(толуол)-2-дезоксирибозилхлорид (563 мг, 1,46 ммоль). Реакционную смесь охлаждали до 0°C и добавляли по каплям SnCl<sub>4</sub> (1,0M в CH<sub>2</sub>Cl<sub>2</sub>, 660 мкл, 0,66 ммоль). Полученный раствор перемешивали в течение 2 ч при комнатной температуре. Реакционную смесь разбавляли EtOAc, гасили насыщенным водным раствором NaHCO<sub>3</sub>, экстрагировали этилацетатом, сушили, фильтровали и выпаривали. Неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 5d в виде белой пены (260 мг, 0,52 ммоль, 39%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 7,98-7,90 (м, 4H, Ar-H), 7,70 (д, J = 6 Гц, 1H, Ar-H), 7,55 (д, J = 9,5 Гц, 1H, Ar-H), 7,28-7,16 (м, 5H, Ar-H), 6,84-6,85 (м, 1H, Ar-H), 6,57 (д, J = 9,5 Гц, H-1'), 5,66-5,64 (м, 1H, H-4'), 4,75-4,72 (м, 2H, H-5'a, b), 4,61 (м, 1H, H-3'), 2,95-2,90 (м, 1H, H-2' a), 2,43 (с, 3H, Ar-CH<sub>3</sub>), 2,40 (с, 3H, Ar-CH<sub>3</sub>), 2,39-2,31 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 166,2, 166,1, 158,1, 145,1, 144,4, 144,2, 133,8, 129,9, 129,6, 129,3, 129,1, 126,9, 126,5, 124,2, 103,5, 85,5, 82,9, 75,1, 64,4, 39,2, 21,7.

HRMS (ESI<sup>+</sup>) m/z вычислено для C<sub>20</sub>H<sub>20</sub>Cl<sub>2</sub>N<sub>2</sub>O<sub>5</sub>S (M + H) 504,1475, найдено 504,1480.

Соединение 6d. Соединение 5d (50 мг, 0,1 ммоль) сушили 3 сопутствующими испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1 мл). Добавляли реагент Лавессона (48 мг, 0,12 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле (гексан/EtOAc), получая соединение 6d в виде желтой пены (17 мг, 0,033 ммоль, 33%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 8,14-7,82 (м, 7H, Ar-H), 7,51 (дд, J = 7,5, 6,0 Гц, 1H, H-1'), 7,32-7,23 (м, 5H, Ar-H), 6,99 (д, J = 7,2 Гц, 1H, Ar-H), 7,4-5,73 (м, 1H, H-4'), 4,92-4,83 (м, 2H, H-5'a, b), 4,78-4,77 (м, 1H, H-3'), 3,43-3,40 (м, 1H, H-2'a), 2,51 (с, 3H, Ar-CH<sub>3</sub>), 2,48 (с, 3H, Ar-CH<sub>3</sub>), 2,39-2,36 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 173,5, 166,6, 144,9, 144,8, 139,5, 138,0, 134,5, 130,3, 130,0, 129,7, 129,5, 126,8, 124,7, 109,5, 91,4, 84,0, 75,0, 64,5, 39,4, 22,2, 22,1.

HRMS (ESI<sup>+</sup>) m/z вычислено для C<sub>28</sub>H<sub>26</sub>NO<sub>5</sub>S<sub>2</sub> (M + H<sup>+</sup>) 520,1247, найдено 520,1241.

Соединение 7d. К раствору 6d (20 мг, 0,039 ммоль) в метаноле (1,0 мл) добавляли по каплям 30%

NaOMe (8,66 мг, 0,16 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь затем концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле (MeOH/CH<sub>2</sub>Cl<sub>2</sub>) с получением соединения 7d в виде желтой пены (9,0 мг, 0,032 ммоль, 82%).

<sup>1</sup>H ЯМР (500 МГц, CD<sub>3</sub>OD) δ 8,48 (д, J = 5 Гц, 1H, Ar-H), 8,01 (д, J = 5 Гц, 1H, Ar-H), 7,40-7,38 (м, 2H, Ar-H), 7,29 (д, J = 10 Гц, 1H, H-1'), 4,47-4,46 (м, 1H, H-4'), 4,10 (м, 1H, H-3'), 3,94-3,88 (м, 2H, H-5'a, b), 2,77-2,76 (м, 1H, H-2'a), 2,19-2,14 (м, 1H, H-2'b).

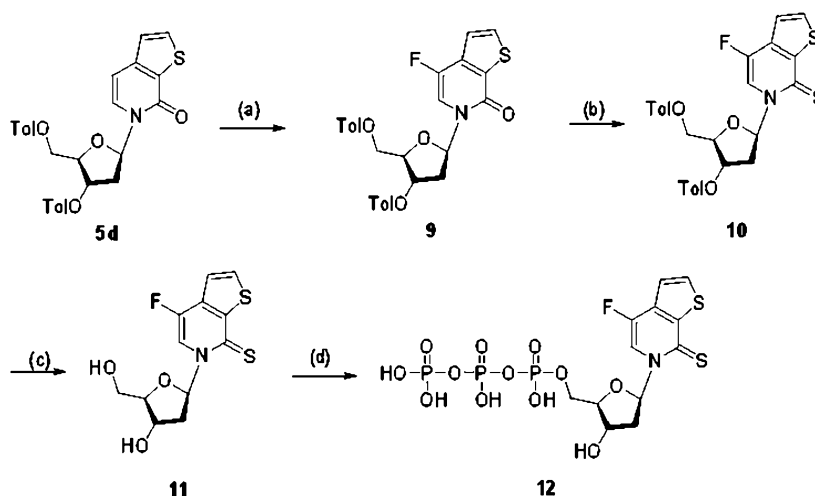
<sup>13</sup>C ЯМР (125 МГц, CD<sub>3</sub>OD) δ 171,2, 144,7, 139,6, 137,6, 130,5, 124,2, 108,8, 90,7, 88,2, 70,1, 61,0, 41,3.

HRMS (ESI +) m/z вычислено для C<sub>12</sub>H<sub>14</sub>NO<sub>3</sub>S<sub>2</sub> (M + H<sup>+</sup>) 284,041, обнаружено 284,0410.

Соединение 8d (dTPT3TP). Соединение 8d (5,7 мг, 10,9 мкмоль, 31%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с 7d (10 мг, 35,3 мкмоль).

<sup>31</sup>P ЯМР (162 МГц, D<sub>2</sub>O) δ -9,3 (д, J = 19,5 Гц, у-Р), -10,8 (д, J = 19,8 Гц, а-Р), -22,4 (т, J = 20,0 Гц, ф3-Р). MS (MALDI-TOF-, матрица: 9-аминоакридин) (m/z): [M-H]<sup>-</sup> вычислено для C<sub>12</sub>H<sub>15</sub>NO<sub>12</sub>P<sub>3</sub>S<sub>2</sub><sup>-</sup>, 521,9, найдено, 521,9.

Схема S6



(a) i) Selectfluor, MeOH/CH<sub>3</sub>CN, кипячение с обратным холодильником, 3 ч; ii) TfOH-CH<sub>2</sub>Cl<sub>2</sub> (1:1 об./об.), 1 ч, 85%; (b) реагент Лавессона, толуол, кипячение с обратным холодильником в течение ночи 32%; (c) 30% NaOMe, MeOH, комн. темп., 1 ч, 85%; (d) протонная губка, POCl<sub>3</sub>, Bu<sub>3</sub>N, Bu<sub>3</sub>NPPi, (MeO)<sub>3</sub>P, DMF, -20°C, 10%.

Соединение 9. Соединение 5d (55 мг, 0,11 ммоль) растворяли в 1,0 мл MeOH-CH CN (1:1 об./об.), добавляли selectfluor (42 мг, 0,12 ммоль) и смесь нагревали с обратным холодильником в течение 3 ч, затем растворитель выпаривали, остаток растворяли в EtOAc (20 мл), органическую фазу промывали водой три раза. Затем органический растворитель выпаривали и твердый остаток сушили в 3 сопутствующих испарениях с безводным толуолом. Остаток растворяли в 1 мл TfOH-CH<sub>2</sub>Cl<sub>2</sub> (1:1 об./об.) и смесь перемешивали при комнатной температуре в течение 1 ч, затем смесь концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле (гексан/EtOAc) с получением соединения 9 в виде белого твердого вещества (49 мг, 0,093 ммоль, 85%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 7,98-7,92 (м, 4H, Ar-H), 7,75 (д, J = 5 Гц, 1H, Ar-H), 7,52 (д, J = 7,5 Гц, 1H, Ar-H), 7,32-7,21 (м, 5H, Ar-H), 6,82-6,78 (м, 1H, H-1'), 5,64-5,61 (м, 1H, H-4'), 4,80-4,59 (м, 2H, H-5'a, b), 4,62-4,59 (м, 1H, H-3'), 2,93-2,87 (м, 1H, H-2'a), 2,43 (с, 3H, Ar-CH<sub>3</sub>), 2,39 (с, 3H, Ar-CH<sub>3</sub>), 2,34-2,27 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 166,2, 166,1, 156,4, 144,5, 144,3, 137,7, 137,5, 134,6, 129,9, 129,6, 129,3, 126,6, 126,4, 120,2, 112,1, 111,7, 85,5, 83,1, 75,0, 64,2, 39,1, 21,8, 21,7.

<sup>19</sup>F ЯМР (376 МГц, CDCl<sub>3</sub>) δ -151,5. HRMS (ESI+) m/z вычислено для C<sub>28</sub>H<sub>25</sub>FNO<sub>6</sub>S (M + H<sup>+</sup>) 522,1381, найдено 522,1380.

Соединение 10. Соединение 9 (20 мг, 0,038 ммоль) высушивали 3 сопутствующими испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1 мл). Добавляли реагент Лавессона (18,5 мг, 0,046 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/EtOAc), получая соединение 10 в виде желтой пены (6,5 мг, 0,012 ммоль, 32%).

<sup>1</sup>H ЯМР (500 МГц, CDCl<sub>3</sub>) δ 8,11-7,85 (м, 6H, Ar-H), 7,40-7,39 (м, 2H, Ar-H, H-1'), 7,28-7,21 (м, 4H, Ar-H), 5,64-5,63 (м, 1H, H-4'), 4,83 (м, 2H, H-5'a, b), 4,69 (м, 1H, H-3'), 3,34-3,29 (м, 1H, H-2'a), 2,44 (с, 1H, Ar-CH<sub>3</sub>), 2,40 (с, 3H, Ar-CH<sub>3</sub>), 2,30-2,26 (м, 1H, H-2'b).

<sup>13</sup>C ЯМР (125 МГц, CDCl<sub>3</sub>) δ 170,9, 166,6, 166,5, 144,9, 144,8, 138,6, 130,3, 130,0, 129,9, 129,7, 129,7,

126,9, 126,7, 120,5, 116,3, 116,0, 100,0, 91,6, 84,3, 74,7, 64,3, 39,2, 22,2, 22,1.

$^{19}\text{F}$  ЯМР (376 МГц,  $\text{CDCl}_3$ )  $\delta$  -142,9.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{28}\text{H}_{25}\text{FNO}_5\text{S}_2$  ( $M + H^+$ ) 538,1153, найдено 538,1155.

Соединение 11. К раствору 10 (10 мг, 0,019 ммоль) в метаноле (1,5 мл) добавляли по каплям 30% NaOMe (4,33 мг, 0,08 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь затем концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле ( $\text{MeOH}/\text{CH}_2\text{Cl}_2$ ) с получением соединения 11 в виде желтой пены (4,9 мг, 0,016 ммоль, 85%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  8,68 (д,  $J = 5$  Гц, 1H, Ar-H), 8,12 (д,  $J = 5$  Гц, 1H, Ar-H), 7,52 (д,  $J = 5$  Гц, 1H, Ar-H), 7,28 (т,  $J = 6,5$  Гц, 1H, H-1'), 4,48 (м, 1H, H-4'), 4,10 (м, 1H, H-3'), 3,94 (дд,  $J = 35,3$  Гц, 2H, H-5'a, b), 2,78-2,75 (м, 1H, H-2'a), 2,24-2,19 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  170,2, 150,2, 148,3, 139,0, 131,7, 131,6, 119,8, 117,8, 117,4, 91,5, 88,7, 70,1, 61,0, 41,5.

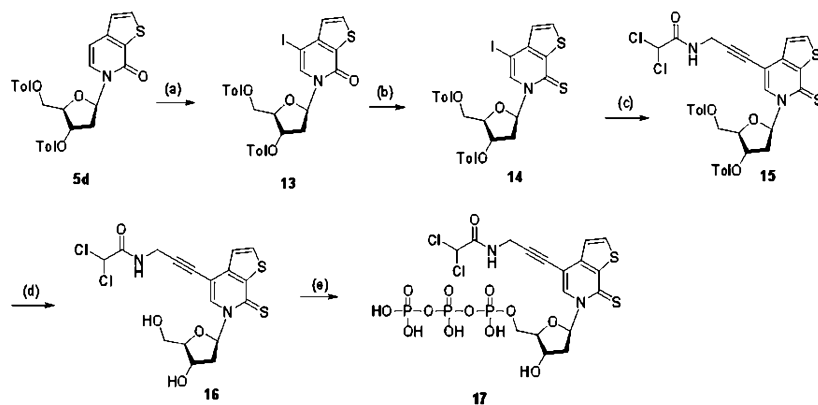
$^{19}\text{F}$  ЯМР (376 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  -145,3. HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{12}\text{H}_{13}\text{FNO}_3\text{S}_2$  ( $M + H^+$ ) 302,0315, найдено 302,0314.

Соединение 12 (dFPT3TP). Соединение 12 (2,0 мг, 3,7 мкмоль, 22%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с соединения 11 (5 мг, 16,6 мкмоль).

$^{31}\text{P}$  ЯМР (162 МГц,  $\text{D}_2\text{O}$ )  $\delta$  -10,9 (д,  $J = 20,0$  Гц, у-Р), -11,6 (д,  $J = 21,1$  Гц, АР), -23,3 (т,  $J = 23,1$  Гц, f3-Р).

$^{19}\text{F}$  ЯМР (376 МГц,  $\text{D}_2\text{O}$ )  $\delta$  -138,5 (s). MS (MALDI-TOF-, матрица: 9-аминоакридин) ( $m/z$ ):  $[M-H]^+$  вычислено для  $\text{C}_{12}\text{H}_{14}\text{FNO}_{12}\text{P}_3\text{S}_2$ -, 539,9, найдено 540,1.

Схема S7



(a)  $\text{ICl}$ ,  $\text{CH}_2\text{Cl}_2$ , от  $0^\circ\text{C}$  до комн. темп., в течение ночи, 63%; (b) реагент Лавессона, толуол, кипячение с обратным холодильником в течение ночи 27%; (c) 2,2-дихлор-N-проп-2-ин-1-илацетамид,  $(\text{PPh}_3)_4\text{Pd}$ , CuI,  $\text{Et}_3\text{N}$ , DMF, комн. темп., в течение ночи, 91%; (d) 30% NaOMe, MeOH, комн. темп., 1 ч, 74%; (e) протонная губка,  $\text{POCl}_3$ ,  $\text{Bu}_3\text{N}$ ,  $\text{Bu}_3\text{NPPi}$ ,  $(\text{MeO})_3\text{P}$ , DMF,  $-20^\circ\text{C}$ , 25%.

Соединение 13. К раствору 5d (73 мг, 0,145 ммоль) в  $\text{CH}_2\text{Cl}_2$  (1 мл) при  $0^\circ\text{C}$  в атмосфере азота по каплям добавляли монохлорид йода (1,0M в  $\text{CH}_2\text{Cl}_2$ , 0,15 мл, 0,15 ммоль). Полученную смесь перемешивали при комнатной температуре в течение ночи. Реакционную смесь гасили насыщенным водным раствором  $\text{NaHCO}_3$  и насыщенным водным раствором  $\text{Na}_2\text{S}_2\text{O}_3$ , экстрагировали  $\text{CH}_2\text{Cl}_2$ , сушили, фильтровали и выпаривали. Неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/ $\text{EtOAc}$ ), получая соединение 13 в виде белой пены (57 мг, 0,091 ммоль, 63%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7,98-7,93 (м, 4H, Ar-H), 7,83 (с, 1H, Ar-H), 7,72 (д,  $J = 5$  Гц, 1H, Ar-H), 7,28-7,17 (м, 5H, Ar-H), 6,78-6,75 (м, 1H, H-1'), 5,65-5,63 (м, 1H, H-4'), 4,76 (м, 2H, H-5'a, b), 4,63-4,62 (м, 1H, H-3'), 2,95-2,91 (м, 1H, H-2'a), 2,43 (с, 3H, Ar-CH<sub>3</sub>), 2,35 (с, 3H, Ar-CH<sub>3</sub>), 2,34-2,29 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  166,6, 166,5, 157,6, 147,4, 144,9, 144,6, 133,7, 132,7, 130,3, 130,1, 129,8, 129,7, 129,4, 128,5, 127,0, 126,8, 86,1, 83,8, 75,7, 64,7, 39,9, 22,2, 22,1.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{28}\text{H}_{25}\text{INO}_5\text{S}$  ( $M + H^+$ ) 630,0442, найдено 630,0440.

Соединение 14. Соединение 13 (30 мг, 0,048 ммоль) высушивали 3 сопутствующими испарениями с безводным толуолом. Остаток растворяли в безводном толуоле (1 мл), добавляли реагент Лавессона (23 мг, 0,057 ммоль) и смесь нагревали в течение ночи с обратным холодильником. После фильтрации на хлопке фильтрат концентрировали и неочищенный продукт подвергали хроматографии на колонке с силикагелем (гексан/ $\text{EtOAc}$ ), получая соединение 14 в виде желтой пены (8,4 мг, 0,013 ммоль, 27%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  8,31 (с, 1H, Ar-H), 7,99-7,82 (м, 5H, Ar-H), 7,39-7,36 (м, 1H, H-1'), 7,29-7,20 (м, 5H, Ar-H), 5,65-5,64 (м, 1H, H-4'), 4,83-4,81 (м, 2H, H-5'a, b), 4,71-4,70 (м, 1H, H-3'), 3,35 (дд,  $J = 15, 5,5$  Гц, 1H, H-2'a), 2,44 (с, 3H, Ar-CH<sub>3</sub>), 2,39 (с, 3H, Ar-CH<sub>3</sub>), 2,27-2,21 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  172,9, 166,6, 144,9, 144,7, 144,6, 141,8, 137,8, 135,0, 130,3, 130,2, 129,8, 129,7, 128,6, 126,9, 126,7, 91,7, 84,5, 75,3, 64,6, 39,5, 22,2, 22,1.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{28}\text{H}_{25}\text{INO}_5\text{S}_2$  ( $M + \text{H}^+$ ) 646,0213, найдено 646,0219.

Соединение 15. К раствору 14 (10 мг, 0,015 ммоль) в DMF (2 мл) в атмосфере азота добавляли  $(\text{PPh}_3)_4\text{Pd}$  (1,7 мг, 0,0015 ммоль),  $\text{CuI}$  (0,57 мг, 0,011 ммоль) и  $\text{Et}_3\text{N}$  (5 мкл, 0,030 ммоль). Реакционную смесь дегазировали и добавляли раствор  $\text{C}_{12}\text{HCHCONHCH}_2\text{CCH}$  (3,8 мг, 0,0225 ммоль) в DMF (0,5 мл). Реакционную смесь перемешивали в течение ночи при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь разбавляли  $\text{EtOAc}$ , гасили насыщенным водным раствором  $\text{NaHCO}_3$ , экстрагировали этилацетатом, сушили, фильтровали и выпаривали. Неочищенный продукт подвергали колоночной хроматографии на силикагеле ( $\text{MeOH}/\text{CH}_2\text{Cl}_2$ ) с получением соединения 15 в виде желтой пены (9,2 мг, 0,0135 ммоль, 91%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CDCl}_3$ )  $\delta$  8,26 (с, 1H, Ar-H), 7,99-7,82 (м, 5H, Ar-H), 7,40-7,37 (м, 2H, Ar-H, H-1'), 7,29-7,21 (м, 4H, Ar-H), 6,71 (br, 1H, NH), 6,95 (с, 1H,  $\text{CHCl}_2$ ), 5,65-5,64 (м, 1H, H-4'), 4,85-4,79 (м, 2H, H-5'a, b), 4,73 (м, 1H, H-3'), 4,26-4,11 (м, 2H,  $\text{NHCH}_2$ ), 3,38-3,34 (м, 1H, H-2'a), 2,44 (с, 3H, Ar-CH<sub>3</sub>), 2,40 (с, 3H, Ar-CH<sub>3</sub>), 2,31-2,25 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CDCl}_3$ )  $\delta$  173,3, 166,6, 164,1, 144,8, 139,0, 138,3, 133,4, 130,3, 130,1, 129,8, 129,7, 126,9, 124,3, 104,8, 91,8, 88,3, 84,5, 78,8, 75,2, 66,4, 64,7, 39,6, 31,3, 22,2, 22,1.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{33}\text{H}_{29}\text{Cl}_2\text{N}_2\text{O}_6\text{S}_2$  ( $M + \text{H}^+$ ) 683,0839, найдено 683,0854.

Соединение 16. К раствору 15 (9,2 мг, 0,0135 ммоль) в метаноле (1,0 мл) добавляли по каплям 30%  $\text{NaOMe}$  (2,92 мг, 0,32 ммоль). Реакционную смесь перемешивали в течение 1 ч при комнатной температуре и контролировали с помощью ТСХ. Реакционную смесь концентрировали и неочищенный продукт подвергали колоночной хроматографии на силикагеле ( $\text{MeOH}/\text{CH}_2\text{Cl}_2$ ) с получением соединения 16 в виде желтой пены (4,5 мг, 0,01 ммоль, 74%).

$^1\text{H}$  ЯМР (500 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  8,69 (с, 1H, Ar-H), 8,06 (д,  $J = 5$  Гц, 1H, Ar-H), 7,53 (д,  $J = 5$  Гц, 1H, Ar-H), 7,30 (т,  $J = 5$  Гц, 1H, H-1'), 6,33 (с, 1H,  $\text{CHCl}_2$ ), 4,47-4,46 (м, 1H, H-4'), 4,36 (с, 2H,  $\text{NHCH}_2$ ), 4,11-4,08 (м, 1H, H-3'), 3,97 (дд,  $J = 12$  Гц, 3 Гц, 2H, H-5'a, b), 2,79-2,74 (м, 1H, H-2'a), 2,21-2,16 (м, 1H, H-2'b).

$^{13}\text{C}$  ЯМР (125 МГц,  $\text{CD}_3\text{OD}$ )  $\delta$  172,9, 138,4, 134,3, 123,9, 122,8, 104,7, 100,0, 91,2, 88,6, 77,1, 70,3, 66,4, 61,1, 41,7, 30,2.

HRMS (ESI+)  $m/z$  вычислено для  $\text{C}_{17}\text{H}_{17}\text{Cl}_2\text{N}_2\text{O}_4\text{S}_2$  ( $M + \text{H}^+$ ) 447,0001, найдено 447,0020.

Соединение 17 ( $\text{dTP}^{\text{PA}}\text{TP}$ ). Соединение 17 (2,2 мг, 3,1 мкмоль, 28%) синтезировали с использованием общей процедуры синтеза трифосфатов, описанной выше, начиная с соединения 16 (5 мг, 11,2 мкмоль).

$^{31}\text{P}$  ЯМР (162 МГц,  $\text{D}_2\text{O}$ )  $\delta$  -10,85 (д,  $J = 19,9$  Гц, у-Р), -11,63 (д,  $J = 20,0$  Гц, AP), -23,07 (с), -23,26 (т,  $J = 19,7$  Гц, f3-Р). MS (MALDI-TOF-, матрица: 9-аминоакридин) ( $m/z$ ):  $[\text{M}-\text{H}]^-$  вычислено для  $\text{C}_{17}\text{H}_{18}\text{Cl}_2\text{N}_2\text{O}_{13}\text{P}_3\text{S}_2^-$ , 684,9, найдено, 685,0.

Пример 3. Общая методика анализа ПЦР-амплификации для определения точности воспроизведения.

Материалы. ДНК-полимеразы Taq и OneTaq были приобретены у New England Biolabs (Ipswich, MA). Смесь дНТФ была приобретена у Fermentas (Glen Burnie, MD). Краситель SYBR Green I Nucleic Acid Gel Stain (10000 x) был приобретен у компании Life Technologies (Carlsbad, CA).

ДНК-олигонуклеотиды. Полные олигонуклеотидные последовательности приведены в табл. 8. Полностью природные праймеры были приобретены у Integrated DNA Technologies (Coralville, Iowa). Реагенты и растворители для синтеза не встречающихся в природе праймеров 1-3 были получены из Glen Research (Sterling, VA) и/или Applied Biosystems (Foster City, CA). Олигонуклеотиды были приготовлены с использованием стандартного автоматизированного синтеза ДНК с ультра-мягкими встречающимися в природе фосфорамидами (Glen Research) и dNaM-фосфорамидатом (Berry & Associates, Inc., Dexter, MI) на стеклянных подложках с контролируемыми размерами пор (0,20 мкмоль, 1000 Å, Glen Research) и в синтезаторе ABI Expedite 8905. После автоматизированного синтеза олигонуклеотиды освобождали с подложки, снимали защиту инкубацией в концентрированном водном растворе аммиака в течение ночи при комнатной температуре, очищали ДМТ-очисткой (картридж glen-pak<sup>TM</sup>, Glen Research) и обессоливали через сефадекс G-25 (колонки NAP-25, GE Healthcare). Концентрацию одноцепочечных олигонуклеотидов определяли по УФ-поглощению при 260 нм.

Способ ПНР. ПЦР-амплификации проводили в общем объеме 25 мкл и с условиями, специфическими для каждого анализа, как описано в табл. 9. После амплификации аликвоту в 5 мкл анализировали в 6%-ном неденатурирующем полиакриламидном (ПААГ) геле вместе с набором служащих метчиками ДНК-фрагментов с шагом в 50 п.н. (50bp ladder) (Life Technologies), чтобы подтвердить размер ампликона. Оставшийся раствор очищали спин-колонкой (ДНК Clean и Concentrator-5; Zymo Research, Irvine, CA), а затем в 4%-ном агарозном геле, извлекали набором Zymoclean Gel DNA Recovery Kit (Zymo Research), определяли количественно по связыванию флуоресцентного красителя (Quant-iT dsDNA HS Assay kit, Life Technologies) и секвенировали на анализаторе 3730 ДНК Analyzer (Applied Biosystems). Точность воспроизведения определяли как среднее %-ное сохранение не встречающейся в природе пары

оснований за дубликацию, как описано ниже.

Определение точности воспроизведения. Процент сохранения не встречающейся в природе пары оснований (F) был измерен с использованием исходных данных секвенирования и нормирован по точности воспроизведения за дубликацию. Вкратце, присутствие не встречающегося в природе нуклеотида приводит к резкой терминации в профиле секвенирования, тогда как мутация с появлением природного нуклеотида приводит к "read-through" (полному прочтению). Степень "полного прочтения", таким образом, обратно пропорциональна сохранению не встречающейся в природе пары оснований. Чтобы использовать данные секвенирования для количественного измерения точности ПЦР, авторы провели эксперименты по калибровке в диапазоне от 50 до 100% сохранения не встречающейся в природе пары оснований. Следовательно, низкое сохранение (<50%) и высокий уровень "полного прочтения" делают количественное определение неточным.

Количественное определение высокой степени сохранения (>50%) выполняли путем регулировки точек запуска и остановки для программного обеспечения Sequencing Analysis (Applied Biosystems), с последующим определением средней интенсивности сигнала для каждого канала индивидуально (A, C, G и T) для пиков в пределах данных установленных точек (от 35 до 45 нуклеотидов в длину) до (раздел L) и S36 после (раздел R) не встречающегося в природе нуклеотида. Отношение R/L нормализовали с использованием калибровочных графиков секвенирования, чтобы учесть как шум в хроматограммах секвенирования, так и "полное прочтение" в контрольных образцах. Соотношение R/L после нормализации (R/L<sub>norm</sub>) соответствует проценту встречающихся в природе последовательностей в пуле. Наконец, F был рассчитан как  $1 - (R/L)_{\text{norm}}$ , и сохранение не встречающейся в природе пары оснований за дубликацию (точность воспроизведения, f) рассчитывали как  $1/(F^{\log 2A})$ , где A представляет собой амплификацию и  $\log 2A$  представляет собой количество дубликаций. Каждый образец до и после ПНР-амплификации секвенировали в трех повторах в каждом направлении, чтобы минимизировать ошибку секвенирования. Соответствующие данные приведены в табл. 10. В соответствии со стандартными условиями ПЦР, ДНК, содержащую dTPT3-dNaM, амплифицировали полимеразой OneTaq с эффективностью, которая только в 4 раза меньше, чем у ДНК, содержащей только природные пары оснований, и с точностью воспроизведения, превышающей 99,98%. Такая точность воспроизведения соответствует частоте ошибок в  $10^{-4}$  на нуклеотид, которая перекрывается с частотой ошибок, составляющей от  $10^{-4}$  до  $10^{-7}$ , для полностью природной ДНК с обычно используемыми системами ПЦР. С Taq-полимеразой эффективность только в 2,5 раза ниже, чем у природной пары оснований, и точность воспроизведения равна 99,7%. Такая точность воспроизведения соответствует частоте ошибок в  $10^{-3}$ , которая подобна той, которая наблюдается с Taq-опосредованной амплификацией природной ДНК.

Название	Последовательность (от 5' к 3')	Примечания
	Области праймеров подчеркнуты	
Fend1	<u>CACACAGGAAACAGCTATGAC</u>	Праймеры для ПЦР (матрицы D6 и 134-мер)
Fend2	<u>GAAATTAATACGACTCACTATAGG</u>	
Fend1-poly-dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT <u>CACACAG</u> <u>GAAACAGCTATGAC</u>	Праймеры для секвенирования по Сэнгеру (матрицы D6 и 134-мер)
Fend2-poly-dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT <u>GAAAT</u> <u>TAATACGACTCACTATAGG</u>	
D6	<u>CACACAGGAAACAGCTATGAC</u> CCCGGGTTATTAC ATGCGCTAGCACTTGAATTCACCAGACGNNN <b>NaM</b> NNNCGGGACCCATAGTAAATCTCCTTCTT AAAGTTAAGCTTAAC <u>CCCTATAGTGAGTCGTATT</u> <u>AATTTC</u>	N = рандомизированный природный нуклеотид
134-мер	<u>CACACAGGAAACAGCTATGAC</u> CCCGGGTTATTAC ATGCGCTAGCACTTGAATTCACAATACT <b>NaM</b> TCTTTAAGGAAACCATAGTAAATCTCCTTCTT	
	AAAGTTAAGCTTAAC <u>CCCTATAGTGAGTCGTATT</u> <u>AATTTC</u>	
Primer1	<b>NaM</b> <u>CCTGCGTCAATGTAATGTTC</u>	Праймеры для ПЦР с Temp1-3
Primer2	<u>TTCACGGT</u> <b>NaM</b> <u>AGCACGCATAGG</u>	
Primer3	<u>CCAATGTACC</u> <b>NaM</b> <u>TGCGTATGTTC</u>	
Primer-rev	<u>CCCTGCGTTTATCTGCTCTC</u>	
Temp1	<u>CCCTGCGTTTATCTGCTCTC</u> TCGGTCGTTCGGC TGCGGCGG <u>AACATTACATTGACGCAGG</u>	Нуклеотиды, выделенные жирным шрифтом, образуют ошибочную пару с dNaM в первом цикле ПЦР
Temp2	<u>CCCTGCGTTTATCTGCTCTC</u> TCGGTCGTTCGGC TGCGGCG <u>CCTATGCGTGCTTACCGTGAA</u>	
Temp3	<u>CCCTGCGTTTATCTGCTCTC</u> TCGGTCGTTCGGC TGCGGGAACATACGCATGGTACATTGG	

Таблица 9

## Условия ПЦР

	OneTaq	Taq	ПЦР для мечения биотином (неприродная пара оснований расположена в центрально)	ПЦР для мечения биотином (неприродная пара оснований расположена в положениях 1, 9, 11)
Буфер	1x OneTaq	1x Taq	1x OneTaq	1x OneTaq
Фермент, ед/мкл	OneTaq, 0,02	Taq, 0,02	OneTaq, 0,02	OneTaq, 0,02
Матрица	D6 (0,01 нг)	D6 (0,01 нг)	134-мер (0,5 нг)	60-мер (0,5 нг)
dNTPs, мкМ	200	200	200	200
dNaMTP, мкМ	100	100	100	100
dXTP, мкМ	100	100	100 от d5SICS <sup>PA</sup> TP или dTPT3 <sup>PA</sup> TP	100 от d5SICS <sup>PA</sup> TP или dTPT3 <sup>PA</sup> TP
Mg <sup>2+</sup> , мМ	3	3	3	3
Праймеры, мкМ	1	1	1	1
SYBR Green I	0,5x	0,5x	0,5x	0,5x
Термические условия				
Начальная денатурация	-	-	96 °C, 1 мин	96 °C, 1 мин
Денатурация	96 °C, 10 с	96 °C, 10 с	96 °C, 15 с	96 °C, 15 с
Отжиг	60 °C, 15 с	60 °C, 15 с	60 °C, 30 с	64 °C, 30 с
Удлинение	68 °C, 60s с	68 °C, 15 с	68 °C, 2 мин	68 °C, 2 мин
Количество циклов	16+16+16	20	12	12

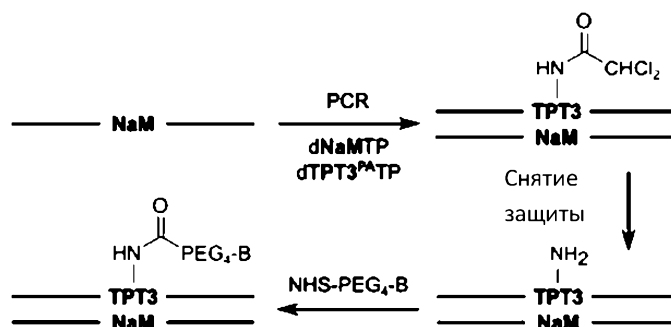
Таблица 10

	OneTaq-ПЦР (48 циклов)			Taq-ПЦР (20 циклов)		
dXTP	Амплифика ция x 10 <sup>12</sup>	Сохранение , %	Точность воспроизведения , %	Амплифика ция x 10 <sup>13</sup>	Сохранение, %	Точность воспроизведе- ния, %
5SICS	9,4	96,3 ± 1,7	99,91 ± 0,04	7,7	86,7 ± 1,0	98,90 ± 0,01
TPT3	12,9	>99	>99,98	11,7	95,6 ± 1,7	99,66 ± 0,13
TPT3 <sup>PA</sup>	4,7	98,6 ± 1,2	99,97 ± 0,03	3,5	85 ± 4	98,7 ± 0,4
5SICS <sup>P</sup> A	9,2	45 ± 2	98,16 ± 0,12	6,4	<sup>a</sup> —	<sup>a</sup> —
<sup>a</sup> Неприродная пара основания, потерянная во время амплификации						



Пример 4. Сайт-специфическое мечение TPT3: анализ по влиянию стрептавидина (SA) на электрофоретический сдвиг в геле.

Синтезировали 134-мер ДНК, содержащей центрально расположенную пару dTPT3<sup>PA</sup>-dNaM или d5SICS<sup>PA</sup>-dNaM. ДНК-матрицы амплифицировали с помощью ПНР в условиях, описанных в табл. 9. По завершении непосредственно к образцам ПНР добавляли NaOH (1М, 12,5 мкл) до конечной концентрации 0,2М и инкубировали в течение 5 ч при комнатной температуре. После добавления NaOAc (3М, pH 5,5, 7,5 мкл) и 200 мкл холодного этанола образцы смешивали, инкубировали на льду в течение ночи и осаждали ДНК центрифугированием при 10000 gfu в течение 30 мин при 4°C. Надосадочную жидкость удаляли и осадок тщательно промывали 80%-ным этанолом. Образцы ресуспендировали в 50 мкл буфера для отжига (50 мМ Na-фосфат, pH 7,5, 100 мМ NaCl, 1 мМ ЭДТА), нагревали до 95°C и охлаждали до комнатной температуры в течение 30 мин. Раствор NHS-биотин-PEG4 (Thermo Scientific) в буфере для отжига (40 мМ, 50 мкл) смешивали с ДНК-пробами и инкубировали в течение ночи при комнатной температуре. Образцы очищали спин-колонкой (ДНК Clean и Concentrator-5, Zymo Research) и элюировали в 10 мкл буфера для элюции. Половину образца (5 мкл) смешивали с 1 мкг стрептавидина (Promega) в буфере для отжига, инкубировали в течение 30 мин при 37°C, смешивали с 5-кратным неденатурирующим буфером для нанесения (Qiagen) и наносили на 6%-ный неденатурирующий ПААГ. Оставшуюся половину смешивали с 5-кратным неденатурирующим буфером для нанесения и наносили непосредственно на гель в качестве контроля. После проведения электрофореза при 110 V в течение 30 мин, гель вымачивали в 1× красителе Sybr Gold Nucleic Acid Stain (Life Technologies) в течение 30 мин и визуализировали с использованием Molecular Imager Gel Doc XR+, оснащенного фильтром 520DF30 (Bio-Rad). Схема описанной стратегии мечения показана ниже.



Пример 5. Общая методика транскрипции не встречающейся в природе пары оснований.

Для характеристики транскрипции не встречающихся в природе пар оснований, образуемых dTPT3 и dNaM, или их аналогами или производными (где производные включают линкерные группы), рибонуклеотиды и дезоксинуклеотиды синтезируются и преобразуются в соответствующие трифосфаты или дезоксифосфорамидиты, и дезоксифосфорамидиты включаются в ДНК-матрицы с использованием автоматизированного синтеза ДНК. Эксперименты по транскрипции проводятся с 100 нМ ДНК-субстрата, 1× Takara буфера (40 мМ Трис-HCl, pH 8,0, 8 мМ MgCl<sub>2</sub>, 2 мМ спермидина), DEPC-обработанной стерилизованной водой без нуклеаз (Fisher), T7 полимеразой (50 ед.), 20 мкМ каждого природного НТФ, α-<sup>32</sup>P-АТФ (2,5 мКи, MP Biomedicals), и либо 5 мкМ TPT3TP, либо 5 мкМ Nam TP. После инкубации в течение 2 ч при 37°C реакцию останавливают добавлением 10 мкл раствора для нанесения на гель (10М мочевины, 0,05% бромфенол синий) и реакционную смесь загружают в 20%-ный полиакриламидный гель с 7М мочевины, подвергают электрофорезу и анализируют с помощью формирования изображения на люминесцентном фосфорном покрытии (phosphorimaging). Эффективность транскрипции исследуют с помощью измерения (при низкой процентной конверсии) количества полноразмерного продукта, образующегося в зависимости от времени.

Пример 6. Общая методика термодинамического анализа ДНК-дуплекса, содержащего не встречающуюся в природе пару оснований.

Эксперименты по УФ-плавлению осуществляются с использованием спектрофотометра Cary 300 Bio UV-visible. Поглощение образца (3 мкл олигонуклеотида, содержащего не встречающуюся в природе пару оснований, 10 мМ PIPES-буфера, pH 7,0, 100 мМ NaCl, 10 мМ MgCl<sub>2</sub>) контролируют на длине волны 260 нм от 21 до 80°C со скоростью нагрева в 0,5°C/мин. Температуры плавления определяют с помощью производного способа с использованием программного обеспечения Cary Win UV thermal application software.

Термодинамические параметры определяются анализом Вант-Гоффа  $T_m^{-1} = R[\ln([C_T]/4)]\Delta H + \Delta S^\circ/\Delta H^\circ$ , где  $\Delta H^\circ$  и  $\Delta S^\circ$  являются стандартными изменениями энтальпии и энтропии, определяемыми из УФ-экспериментов соответственно, R представляет собой универсальную газовую постоянную и  $[C_T]$  представляет собой общую концентрацию олигонуклеотидной цепи. Изменения в количестве молекул воды, связанные с процессом плавления,  $\Delta n_w$ , получают из зависимости  $T_m$  от активности воды ( $\alpha_w$ ) в соответ-

ствии с уравнением  $\Delta n_w = (\Delta H/R)[\delta(T_m^{-1})/\delta(\ln \alpha_w)]$ . Наклон графика зависимости обратной температуры ( $K^{-1}$ ) плавления, отложенного против логарифма активности воды при различных концентрациях (0, 2, 5, 7, 10, 12 и 15% мас.%) этиленгликоля, берется в качестве величины  $\delta(T_m^{-1})/\delta(\ln \alpha_w)$ .

CD-эксперименты проводят со спектрополяриметром Aviv model 61 DS, оснащенным блоком Пельтье для термоэлектрического управления температурой (концентрация олигонуклеотидов 3 мкМ, 10 мМ PIPES-буфера, pH 7,0, 100 мМ NaCl, 10 мМ MgCl<sub>2</sub>). Данные собирают с использованием кварцевой кюветы с длиной пути в 1 см со сканированием от 360 до 220 нм, временной постоянной 3 с и размером шага длины волны в 0,5 нм при 25°C.

Пример 7. Отбор *in vitro* с не встречающимися в природе нуклеотидными основаниями.

Создают олигонуклеотидную библиотеку, содержащую не встречающиеся в природе нуклеиновые кислоты. Образец из библиотеки подвергают последовательному связыванию и элюции от молекулы-мишени, например белка. Пул связавшихся нуклеиновых кислот амплифицируют с помощью ПЦР и подвергают другому раунду отбора на связывание с молекулой-мишенью. Такой процесс отбора повторяется несколько раз. Чтобы увеличить давление отбора, в последних нескольких раундах отбора концентрацию молекулы-мишени и/или время инкубации уменьшают. Выжившие нуклеиновые кислоты секвенируют в качестве потенциальных аптамеров. Аффинности связывания потенциальных аптамеров определяют с помощью проточной цитометрии.

Пример 8. Общая процедура для клик-реакции с ДНК.

К раствору ДНК (0,2 пмоль) в 14 мкл ДМСО добавляли 1 мкл азид-PEG(3+3)-S-S-биотина (20 мМ в H<sub>2</sub>O), а затем 2 мкл лиганда (BimC<sub>4</sub>A)<sub>3</sub> (4 мМ в H<sub>2</sub>O), 1 мкл аскорбата натрия (100 мМ в H<sub>2</sub>O), и 1 мкл PBS-буфера (5×), затем смесь перемешивали и добавляли, в качестве последнего компонента 1 мкл свежеприготовленного раствора CuSO<sub>4</sub> (4 мМ в H<sub>2</sub>O). Раствор перемешивали в течение 2 ч при 37°C и затем полученный ДНК-продукт очищали (набор DNA Clean & Concentrator-5, Zymo Research Corp.). Очищенные образцы использовали непосредственно для анализов подвижности в геле (см ниже).

Пример 9. Общие методики мечения ДНК после амплификации (из Seo et al., JACS 2011, 133, 19878).

Для мечения после ферментативного синтеза, дцДНК со свободной аминокислотной группой инкубировали с 10 мМ EZ-Link сульфидо-NHS-SS-биотина или EZ-Link NHS-PEG<sub>4</sub>-биотина (Thermo Scientific) в течение 1 ч при комнатной температуре в фосфатном буфере для мечения (50 мМ фосфат натрия, pH 7,5, 150 мМ NaCl, 1 мМ ЭДТА) и затем очищали с использованием набора для очистки Qiagen PCR. С любыми защищенными дихлорацетильной группой производными аминов, такими как dTPT3<sup>PA</sup> или d5SICS<sup>PA</sup>, с амина сначала требуется снятие защиты, которое проводят инкубацией в течение ночи в концентрированном водном растворе аммиака при комнатной температуре. Аммиак удаляли с помощью концентратора SpeedVac (водоструйный насос с последующим вакуумным масляным насосом). Для расщепления содержащих дисульфидные связи линкеров (т.е. SS-биотин или SS-PEG<sub>4</sub>-биотин), дцДНК обрабатывали ДТТ (конечная концентрация 30 мМ) в течение 1 ч при 37°C. Для мечения остова дцДНК с фосфоротиоатом в остова инкубировали с 25 мМ EZ-Link иодацетил-PEG<sub>4</sub>-биотина (Thermo Scientific) в фосфатном буфере для мечения в течение ночи при 50°C и продукты очищали с набором для очистки Purification Kit Qiagen PCR. Все реакции манипулирования присоединенными группами с биотином были количественно оценены анализами влияния стрептавидина на электрофоретический сдвиг в геле.

Анализ подвижности в геле. Образцы ДНК (от 10 до 50 нг) смешивали с 1 мкг стрептавидина (Promega) в фосфатном буфере для мечения (50 мМ фосфат натрия, pH 7,5, 150 мМ NaCl, 1 мМ ЭДТА), инкубировали в течение 30 мин при 37°C, смешивали с 5× неденатурирующим буфером для нанесения (Qiagen) и наносили на 6%-ный неденатурирующий ПААГ. Электрофорез проводили при 150 V в течение 25-40 мин, затем гель окрашивали 1× Sybr Gold Nucleic Acid Stain (Life Technologies) в TBE в течение 30 мин и визуализировали с использованием Molecular Imager Гель Doc XR+, оснащенного фильтром 520DF30 (Bio-Rad). Были видны интенсивно окрашенные полосы, соответствующие дцДНК (при ~150 п.н.) и комплексу 1:1 между дцДНК и стрептавидином (при ~400 п.н.). Были также видны бледные полосы, соответствующие более высокого порядка (медленнее мигрирующим) комплексам ДНК и стрептавидина или неботионилированной, одноцепочечной ДНК в результате неполного отжига после ПНР, в некоторых случаях.

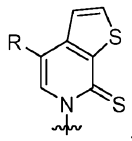
Все патенты и публикации, которые упоминаются в настоящем изобретении, включены в настоящее описание посредством ссылки в той же степени, как если бы каждая отдельная публикация была конкретно и индивидуально указана как включенная посредством ссылки во всей своей полноте.

Термины и выражения, которые были использованы, используются в качестве терминов описания, а не ограничения, и нет намерения при использовании таких терминов и выражений исключить любые эквиваленты показанных и описанных признаков или их частей, но следует признать, что возможны различные модификации в пределах объема заявленных вариантов осуществления. Таким образом, следует понимать, что, хотя в представленных вариантах осуществления специально раскрыты предпочтительные варианты осуществления и дополнительные признаки, специалисты в данной области техники могут прибегнуть к модификации и вариации понятий, раскрытых в настоящем изобретении, и что такие мо-

дификации и варианты считаются находящимися в пределах объема настоящего изобретения, как определено прилагаемой формулой изобретения.

# ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Соединение, содержащее аналог нуклеинового основания формулы



где R представляет собой водород, C<sub>2</sub>-C<sub>3</sub> алкенил, C<sub>2</sub>-C<sub>3</sub> алкинил, метокси, метантиол, метанселено, циано или азида; или R представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу; и волнистая линия показывает точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку, причем рибозильный, дезоксирибозильный или дидезоксирибозильный остаток находится в свободной форме, присоединен к монофосфатной, дифосфатной, трифосфатной, α-тиотрифосфатной, β-тиотрифосфатной или γ-тиотрифосфатной группе или включен в РНК или ДНК.

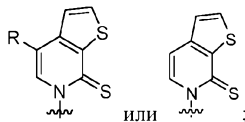
2. Соединение по п.1, в котором R представляет собой H.

3. Соединение по п.1, в котором рибозильный или дезоксирибозильный остаток содержит трифосфатную или α-тиотрифосфатную группу, связанную с его 5'-гидроксильной группой.

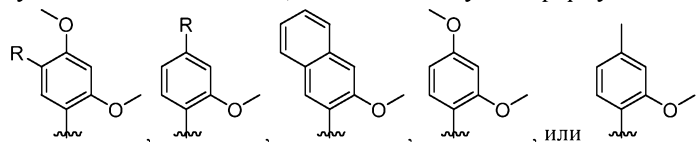
4. Соединение по п.1, в котором рибозильный или дезоксирибозильный остаток включен в олигонуклеотидную цепь РНК или ДНК соответственно.

5. Соединение по п.1, в котором R содержит группу -C≡C-CH<sub>2</sub>NHR<sub>3</sub>, где R<sub>3</sub> представляет собой водород или представляет собой аминозащитную группу; ацетиленовую группу, подходящую для применения в клик-реакции с азидом; тиольную группу, подходящую для применения в реакции с группой, способной реагировать с тиольной группой; альдегидную группу, подходящую для применения в реакции с группой, способной реагировать с альдегидной группой; азидную группу, подходящую для применения в реакции с алкином; группу -C≡C-(CH<sub>2</sub>)<sub>n</sub>-C≡CH, где n равно 1, 2, 3, 4, 5 или 6; или группу -C≡C-(CH<sub>2</sub>)<sub>n1</sub>-O-(CH<sub>2</sub>)<sub>n2</sub>-C≡CH, где n<sub>1</sub> и n<sub>2</sub>, каждый независимо, равны 1, 2 или 3.

6. Композиция для образования не встречающейся в природе пары оснований, содержащая первый аналог нуклеинового основания, имеющий формулу



и второй аналог нуклеинового основания, имеющий любую из формул

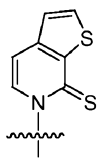


где R независимо представляет собой водород, C<sub>2</sub>-C<sub>3</sub> алкенил, C<sub>2</sub>-C<sub>3</sub> алкинил, метокси, метантиол, метанселено, циано или азида; или R представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу;

и волнистая линия показывает точку присоединения к рибозильному, дезоксирибозильному или дидезоксирибозильному остатку, причем рибозильный, дезоксирибозильный или дидезоксирибозильный остаток находится в свободной форме, присоединен к монофосфатной, дифосфатной, трифосфатной, α-тиотрифосфатной, β-тиотрифосфатной или γ-тиотрифосфатной группе или включен в РНК или ДНК.

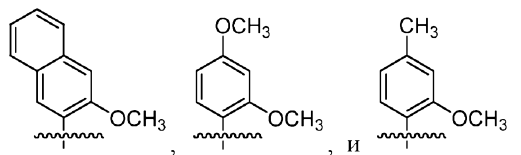
7. Двухцепочечный олигонуклеотидный дуплекс, в котором первая олигонуклеотидная цепь содержит соединение по п.1 и вторая комплементарная олигонуклеотидная цепь содержит комплементарное, способное к образованию пары нуклеиновое основание в ее сайте спаривания комплементарных оснований.

8. Двухцепочечный олигонуклеотидный дуплекс по п.7, в котором первая цепь содержит нуклеиновое основание, имеющее формулу



и вторая цепь содержит комплементарное, способное к образованию пары нуклеиновое основание,

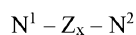
выбранное из



в ее сайте спаривания комплементарных оснований.

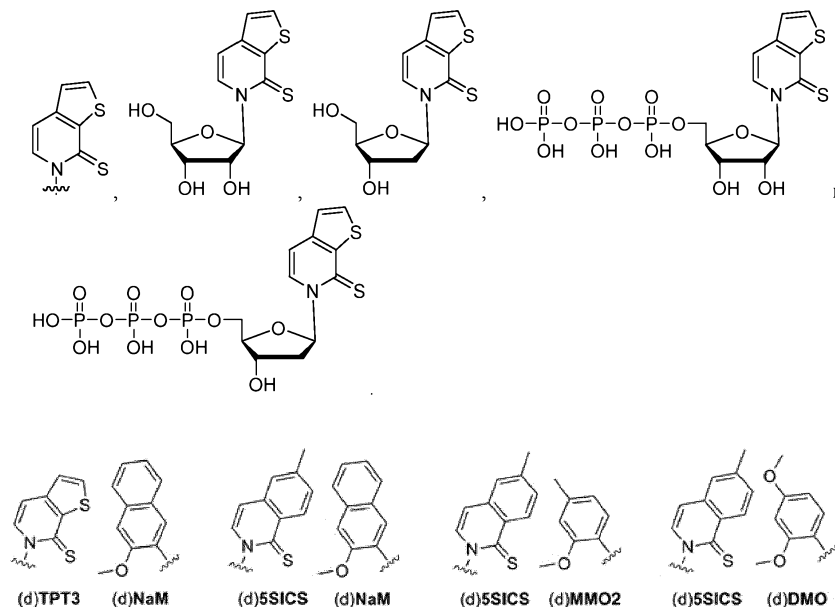
9. Способ проведения сайт-специфической функционализации двухцепочечного олигонуклеотидного дуплекса, включающий включение не встречающегося в природе нуклеинового основания, выбранного из соединения по п.1, в первую олигонуклеотидную цепь; затем синтез второй олигонуклеотидной цепи, которая комплементарна первой олигонуклеотидной цепи, причем вторая олигонуклеотидная цепь содержит нуклеиновое основание, комплементарное не встречающемуся в природе нуклеиновому основанию в сайт-специфической комплементарной позиции в нем, в таких условиях, что первая олигонуклеотидная цепь и вторая олигонуклеотидная цепь образуют двухцепочечный олигонуклеотидный дуплекс; затем контактирование двухцепочечного олигонуклеотидного дуплекса, включающего не встречающееся в природе нуклеиновое основание, с группой, обладающей реакционной способностью, комплементарной заместителю R не встречающегося в природе нуклеинового основания, в условиях, подходящих для осуществления реакции взаимодействия между заместителем R нуклеинового основания и группой, обладающей комплементарной реакционной способностью, с получением сопряженного линкера, чтобы обеспечить функционализированный двухцепочечный олигонуклеотидный дуплекс, где заместитель R не встречающегося в природе нуклеинового основания представляет собой  $C_2-C_3$  алкенил,  $C_2-C_3$  алкинил, метокси, метантиол, метанселено, циано или азидо; или заместитель R не встречающегося в природе нуклеинового основания представляет собой реакционноспособный линкер, содержащий аминогруппу, ацетиленовую группу, тиольную группу, альдегидную группу или азидную группу.

10. Олигонуклеотид, имеющий формулу

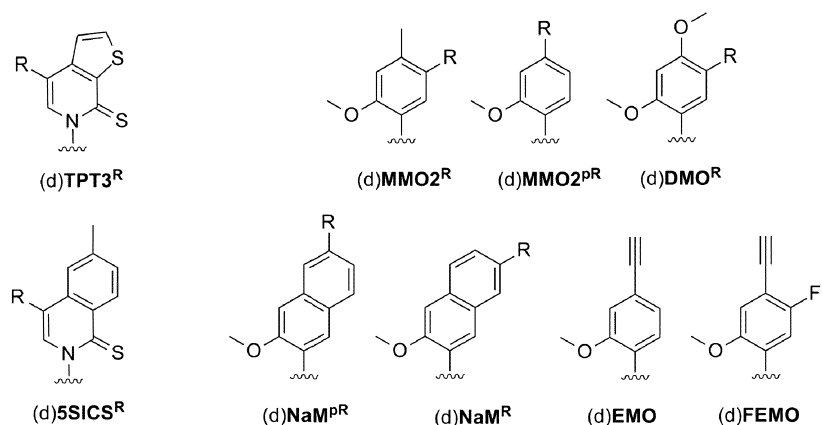


где  $N^1$  представляет собой нуклеотид или концевую фосфатную группу;  $N^2$  представляет собой нуклеотид или концевую гидроксильную группу; Z представляет собой соединение по п.1 и x означает целое число от 1 до 20.

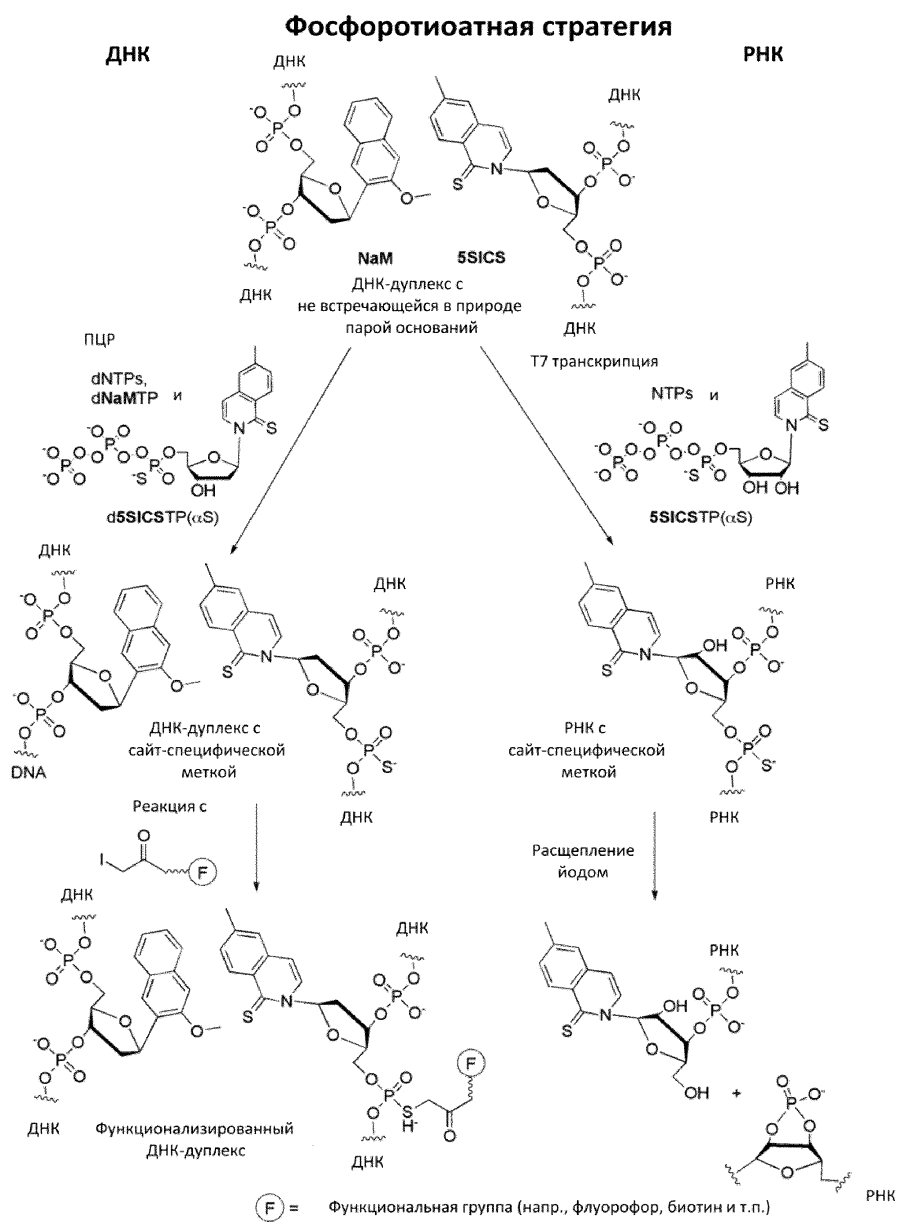
11. Соединение по п.1, имеющее формулу, выбранную из



Фиг. 1

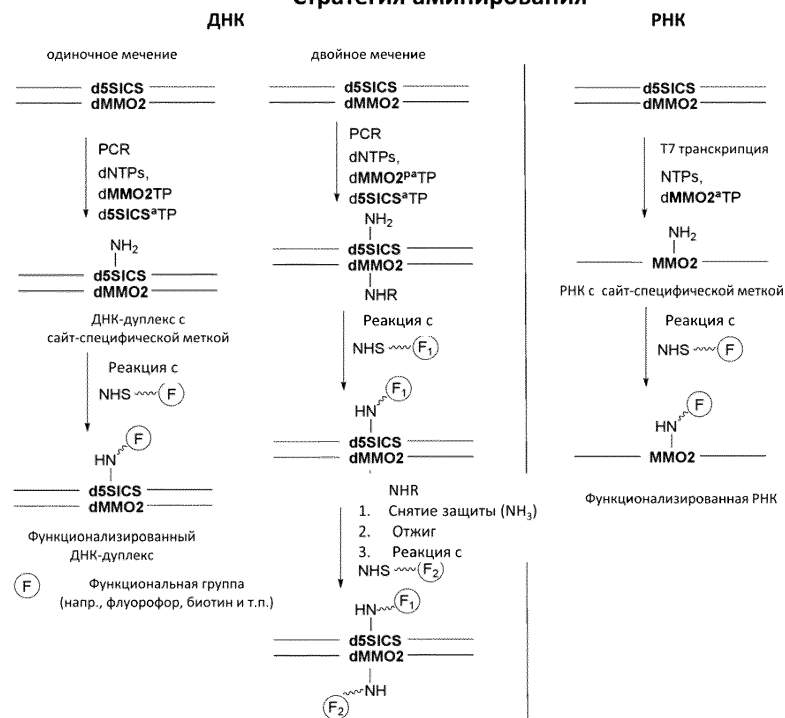


Фиг. 2

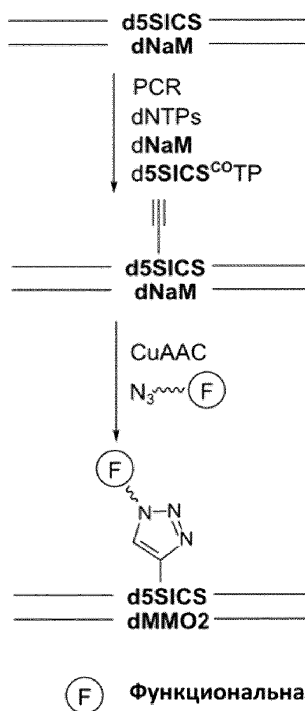


Фиг. 3

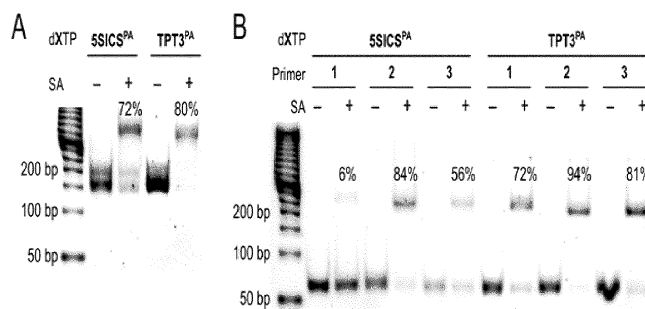
## Стратегия аминирования



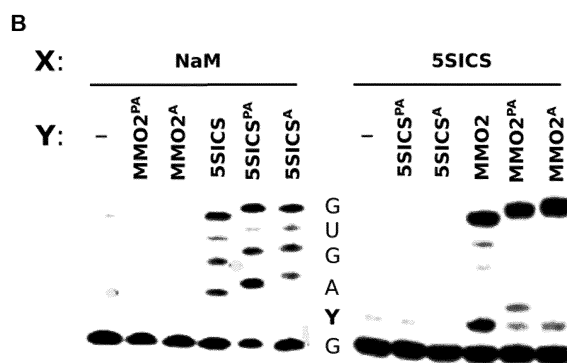
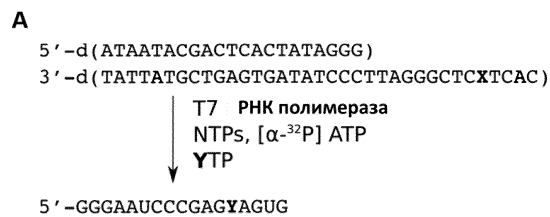
Фиг. 4



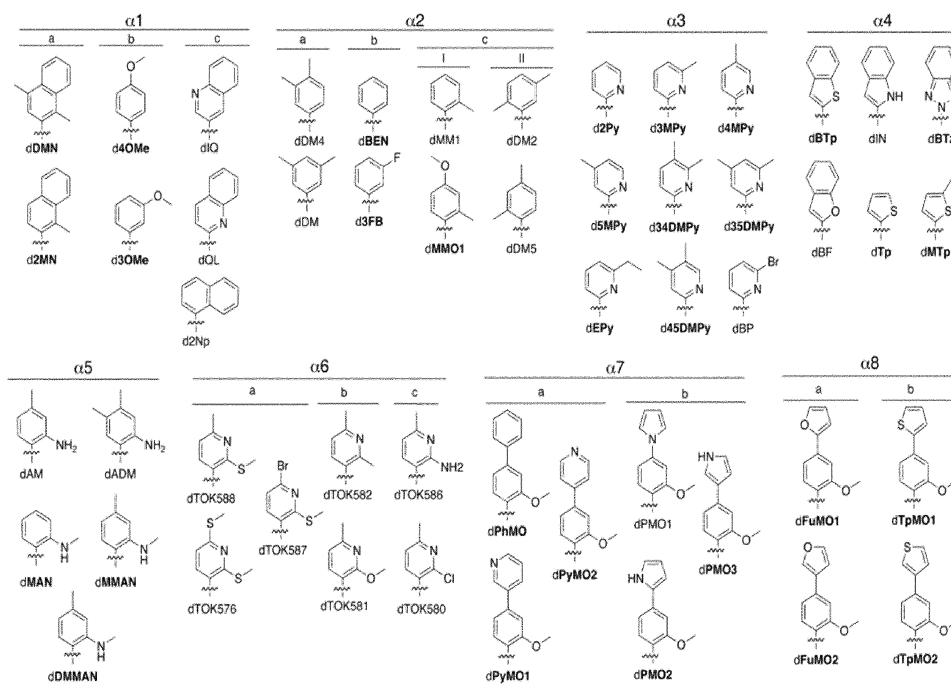
Фиг. 5

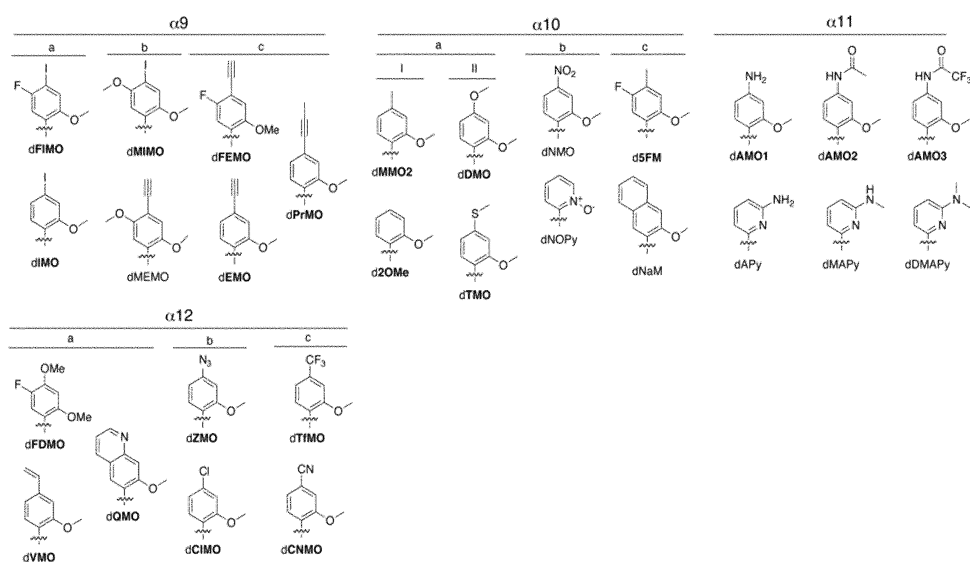


Фиг. 6

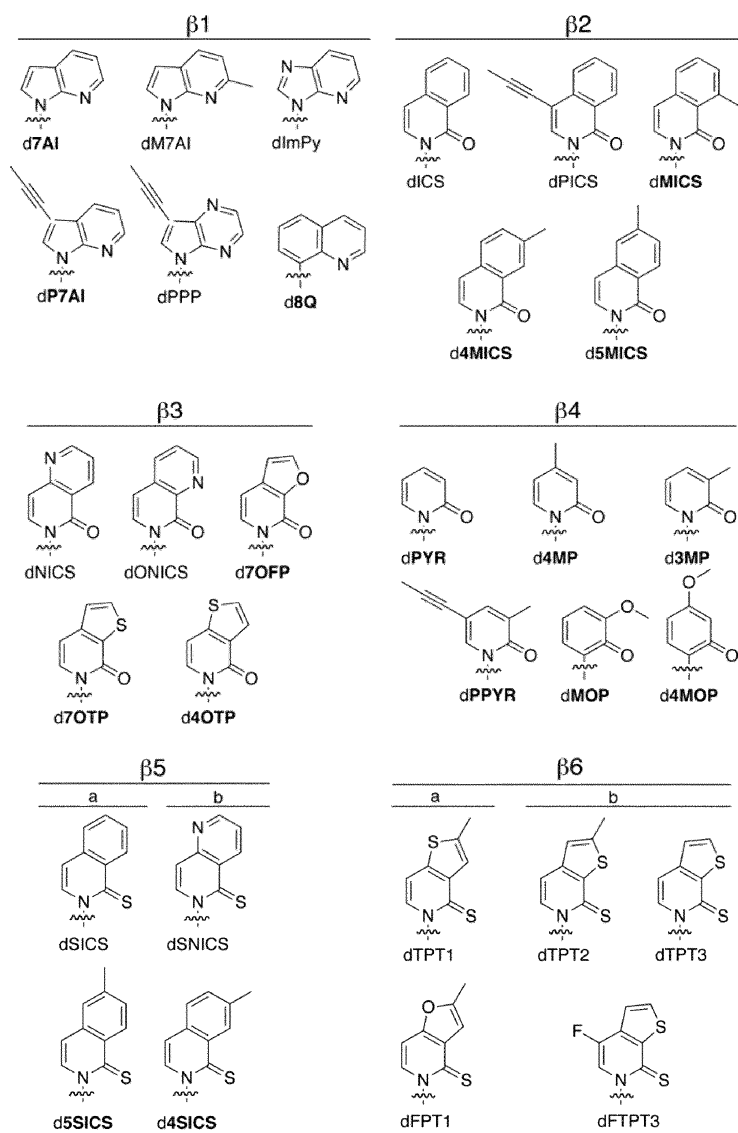


Фиг. 7



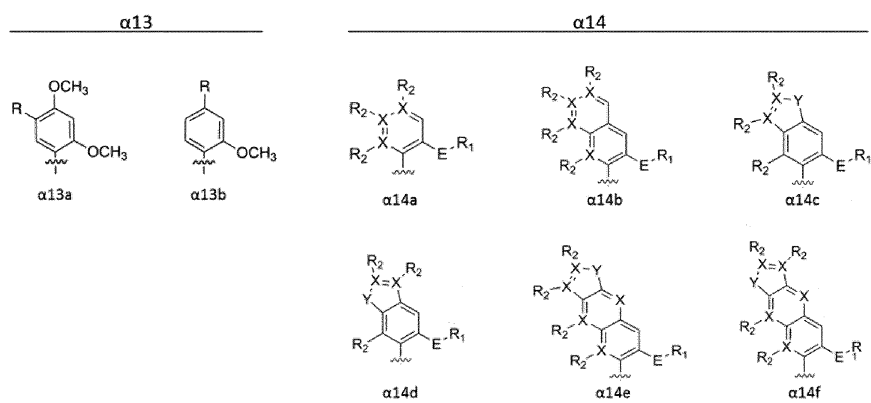


Фиг. 8

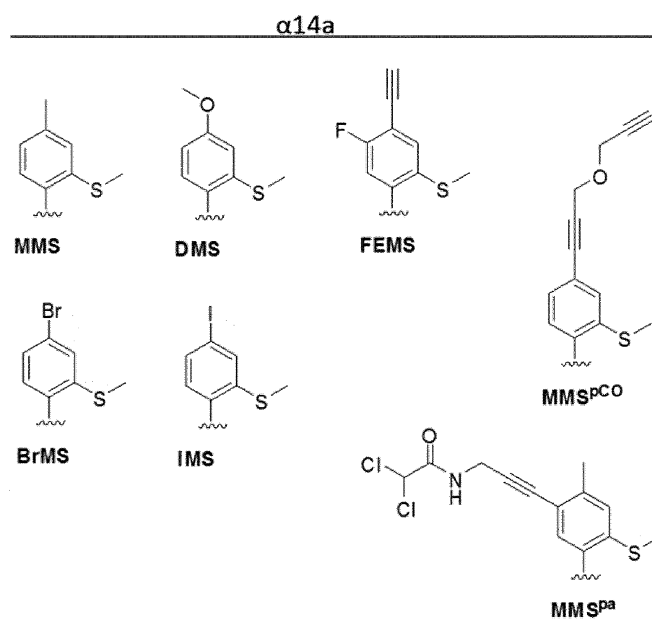


Фиг. 9

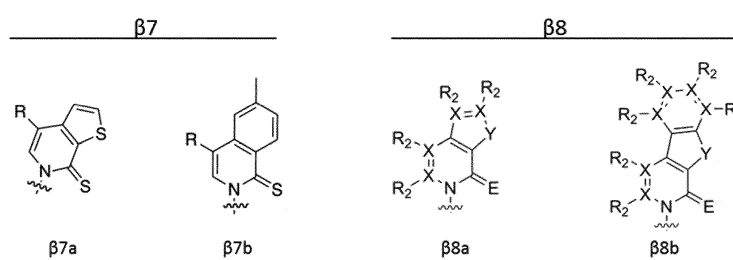




Фиг. 10

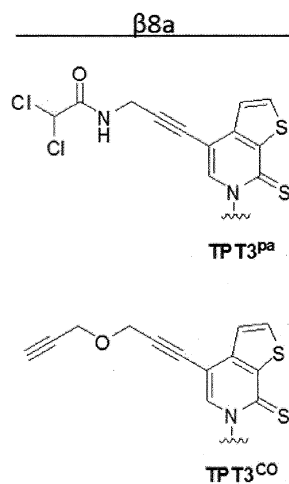


Фиг. 11



Фиг. 12

038576



Фиг. 13

Раунд 1																		Раунд 2								
$\alpha 1$	$\alpha 2$	$\alpha 3$	$\alpha 4$	$\alpha 5$	$\alpha 6$	$\alpha 7$	$\alpha 8$	$\alpha 9$	$\alpha 10$	$\alpha 11$	$\alpha 12$	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 4$	$\beta 5$	$\beta 6$					$\beta 5a$	$\beta 5b$	$\beta 6a$	$\beta 6b$	
																	94	89	$\alpha 1$	a	88	92		86		
																	98	57	$\alpha 2$	b	92	85	86	65		
																			$\alpha 3$	c						
																			$\alpha 4$	a						
																			$\alpha 5$	b						
																	92	77	$\alpha 6$	c	97	89	68			
																	58		$\alpha 7$	a	87	91	74	84		
																	57		$\alpha 8$	b						
																	92	88	$\alpha 9$	c						
																	99	90	$\alpha 10$	a					79	
																	99	87	$\alpha 11$	b					58	
																	99	87	$\alpha 12$	a					79	
																			$\beta 1$	b					69	
																			$\beta 2$	a	99	98		92		
																			$\beta 3$	b	94	100	76	94		
																			$\beta 4$	c	100	86	64	90		
																			$\beta 5$	a	100	91	73	91		
																			$\beta 6$	b	80	69		91		
																				c	100	98	89	98		
																			$\alpha 12$	a	96	94		91		
																				b	99	97	71	90		
																				c	97	87		90		

## Раунд 3

## Раунд 4

Раунд 5

		SICS	3SICS	4SICS	5SICS	SMICS	TP1	TP2	TP3	FTP3	
α1	2MN				71			83	77		
	α2	MM1									
		DM2									
α6	TOK587	77			81			66	76		
	TOK576				68						
α7	PyMO2										
α8	FuMO1										
	FuMO2										
	TrpMO1										
α9	FIMO	82	64		81		85	94	98		
	IMO	71			78		68	93	97		
	FEMO		58				71	85	97		
	PrMO										
	EMO										
	MIMO							62	71		
	MEMO										
	MMO2								76		
	2OMe										
	DMO							67			
α10	TMO										
	NOPy										
	NMO							86	94		
	5FM										
	NaM	99	70		90		70	97	99		
	α12	FDMO						95	85	98	
		VMO								82	
		ZMO	55			62		74	82	88	
CIMO		61		69		61	84	97			
TfMO							82	92			
QMO				70			71	87			
CNMO							84	94			

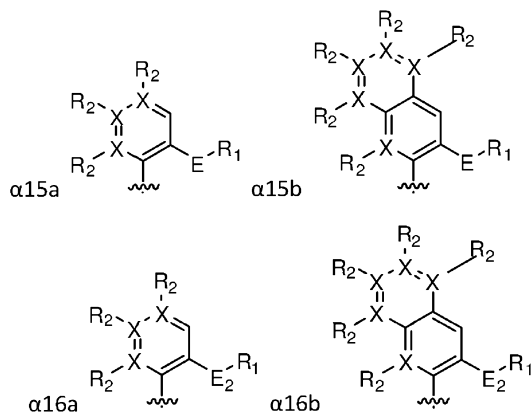
Раунд 6

		SICS	3SICS	4SICS	5SICS	SMICS	TP1	TP2	TP3	FTP3
α1	2MN	60				61		88	70	
	α2	MM1				61	58			
		DM2	88			71	62			
α6	TOK587	78			83		72	81		
	α9	FIMO	80	81				97	89	
IMO		81	75		88		95	97		
FEMO		92	83		95		95	97		
α10	NMO						89	82		
	NaM	95	86		81	93	97	88		
	α12	ZMO						96	87	
CIMO							90	94		
TfMO							76	89		
QMO							90	84		
CNMO							91	92		

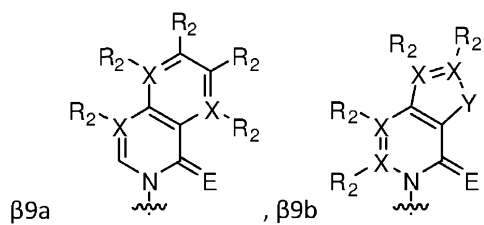
**Taq**

		SICS	3SICS	4SICS	5SICS	SMICS	TP1	TP2	TP3	FTP3
α1	2MN					71		64	91	
	α2	MM1								
DM2										
α6	TOK587	77				72		74	80	
	α9	FIMO					71		99	99
IMO						65		96	89	
FEMO								93	92	
α10	NMO						75	82		
	NaM	79			73		89	92		
	α12	ZMO						77	73	
CIMO							77	89		
TfMO							69	72		
QMO							73			
CNMO							91	94		

Фиг. 14



ФИГ. 15



ФИГ. 16