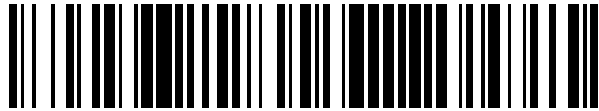


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 773 547**

51 Int. Cl.:

C07H 19/00 (2006.01)

C12N 15/09 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.08.2014 PCT/US2014/050423**

87 Fecha y número de publicación internacional: **12.02.2015 WO15021432**

96 Fecha de presentación y número de la solicitud europea: **08.08.2014 E 14834827 (9)**

97 Fecha y número de publicación de la concesión europea: **04.12.2019 EP 3041854**

54 Título: **Un procedimiento de marcaje enzimático específico de sitio de ácidos nucleicos in vitro mediante la incorporación de nucleótidos no naturales**

30 Prioridad:

08.08.2013 US 201361863649 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

13.07.2020

73 Titular/es:

**THE SCRIPPS RESEARCH INSTITUTE (100.0%)
10550 North Torrey Pines Road
La Jolla, CA 92037, US**

72 Inventor/es:

**ROMESBERG, FLOYD, E.;
MALYSHEV, DENIS, A.;
LI, LINGJUN;
LAVERGNE, THOMAS y
LI, ZHENGTAO**

74 Agente/Representante:

GARCÍA GONZÁLEZ, Sergio

ES 2 773 547 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Un procedimiento de marcaje enzimático específico de sitio de ácidos nucleicos *in vitro* mediante la incorporación de nucleótidos no naturales

5

Declaración de apoyo del gobierno de los Estados Unidos

La presente invención se realizó con apoyo del gobierno bajo la Subvención núm. GM060005, que otorgó el Instituto Nacional de Salud. El Gobierno de los Estados Unidos tiene ciertos derechos en la invención.

10

Referencia cruzada con solicitudes relacionadas

La presente solicitud reivindica el beneficio de la solicitud provisional de Estados Unidos núm. 61/863,649 presentada el 8 de agosto de 2013.

15

Antecedentes

Los oligonucleótidos y sus aplicaciones han revolucionado la biotecnología. Sin embargo, los oligonucleótidos que incluyen tanto el ADN como el ARN incluyen cada uno solo los cuatro nucleótidos naturales de adenosina (A), guanósina (G), citosina (C), timina (T) para el ADN y los cuatro nucleótidos naturales de adenosina (A), guanósina (G), citosina (C), y uridina (U) para el ARN, lo que restringe significativamente las funciones y aplicaciones potenciales de los oligonucleótidos.

20

La capacidad de sintetizar/amplificar con especificidad de secuencia a los oligonucleótidos (ADN o ARN) con polimerasas, por ejemplo por PCR o sistemas de amplificación isotérmica (por ejemplo, la transcripción con la ARN polimerasa T7), ha revolucionado la biotecnología. Además de todas las aplicaciones potenciales en nanotecnología, esto ha permitido una amplia gama de nuevas tecnologías como la evolución *in vitro* vía SELEX (Evolución sistemática de ligandos por enriquecimiento exponencial) de aptámeros y enzimas de ARN y ADN. Véase, por ejemplo, Oliphant AR, Brandl CJ y 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 y Gold L (1990), Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, 249:505-510; Ellington AD y Szostak JW (1990), *In vitro* selection of RNA molecules that bind specific ligands, *Nature*, 346:818-822.

25

30

Desafortunadamente, estas aplicaciones se restringen por la limitada diversidad químico/física presente en el alfabeto genético natural (los cuatro nucleótidos naturales A, C, G y T en el ADN, y los cuatro nucleótidos naturales A, C, G y U en el ARN). En consecuencia, hay mucho interés en técnicas que permitan la síntesis/amplificación enzimática de oligonucleótidos marcados específicamente en el sitio con grupos funcionales que no están presentes entre los nucleótidos del alfabeto genético natural. Actualmente, las opciones disponibles para la derivatización de ácidos nucleicos específica del sitio incluyen síntesis química basada en soporte sólido, síntesis química/enzimática combinada y procedimientos de marcaje final. Los procedimientos de marcaje final se limitan a los oligonucleótidos terminales, y la síntesis química se limita a oligonucleótidos cortos (<200 nucleótidos para el ADN y <70 nucleótidos para el ARN). La funcionalización enzimática depende del reconocimiento enzimático de la modificación de interés y más problemáticamente no es específica del sitio.

40

45

Malyshev DA y otros divulgaron que la replicación eficiente e independiente de la secuencia del ADN que contiene un tercer par de bases establece un alfabeto genético funcional de seis letras (*Proc. Natl. Acad. Sci. USA*, 109 (30), 2012, 12005-12010). Seo Young Jun y otros divulgaron la transcripción de un alfabeto genético expandido. (*Am. Chem. Soc.* 2009, 131 (14), 2009, 5046-5047). Yamashige Rie y otros divulgaron sistemas de pares de bases no naturales altamente específicos como un tercer par de bases para la amplificación por PCR (*Nucleic Acids Research*, 40 (6), 2012). Li Lingjun y otros divulgaron la replicación natural de un par de bases no natural para la expansión del alfabeto genético y las aplicaciones biotecnológicas del mismo (*Journal of the American Chemical Society*, 136 (3), 2014, 826-829). Fourrey y otros divulgaron el fotorreordenamiento del acceso de derivados de fenil seleniuro a nucleósidos C sustituidos con selenio (*Tetrahedron Letters*, 1980, vol. 21, 455-458). Montero y otros divulgaron sobre una síntesis selectiva de varias ribofuranosil-1-purinas (Nucleosidos de *synthese XVI*, páginas 929-935, septiembre de 1978). El documento WO 2002/062816 A1 (Davis y otros) divulga nucleótidos y dinucleótidos no naturales. Henry y otros divulgaron los determinantes de estabilidad de nucleobase no natural y reconocimiento de la polimerasa (*Journal of the American Chemical Society*, 2003, 9638-9646). Leconte y otros divulgaron el descubrimiento, la caracterización y la optimización de un par de bases no natural para la expansión del alfabeto genético (*Journal of the American Chemical Society*, 2008, 2336-2343).

50

55

60

Sumario

La presente invención se define por la materia objeto de las reivindicaciones.

65

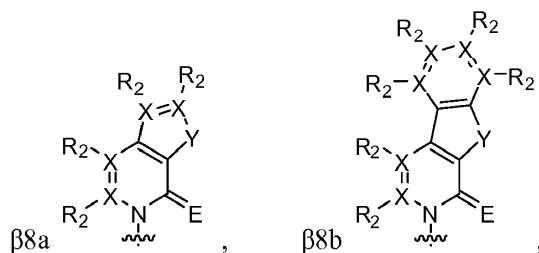
Las composiciones y procedimientos descritos en la presente memoria se basan en la expansión del alfabeto genético *in vitro*, y proporcionan la incorporación específica del sitio de nucleótidos no naturales como se describe en la presente memoria, que portan enlazadores reactivos adaptados para reaccionar con reactivos de carga que comprenden grupos de reactividad complementaria, o enlazadores que portan cargas unidos a estos, en cualquier posición de cualquier secuencia de ADN o ARN, por ejemplo, mediante el uso de PCR estándar o metodologías de transcripción isotérmica.

En diversas realizaciones, los enlazadores se unen a una carga en la etapa de nucleótido trifosfato, lo que permite la producción directa del oligonucleótido marcado específicamente para el sitio deseado, por ejemplo, mediante máquinas automatizadas de síntesis de polinucleótidos tales como máquinas de síntesis de polinucleótidos de fosfoamidita.

Los enlazadores, en otras realizaciones, incluyen un centro reactivo (por ejemplo, amina primaria, alquino, tiol, aldehído o azida), que proporciona un enlazador reactivo, que permite la modificación específica del sitio del oligonucleótido de ADN o ARN después de su síntesis. Esto se puede lograr mediante el uso de un reactivo de carga que comprende una carga (por ejemplo, una molécula, un liposoma, una nanopartícula, etc.) y un grupo de reactividad complementaria al centro reactivo de la fracción enlazadora reactiva. En algunas realizaciones, el centro reactivo de la fracción enlazadora se protege con un grupo protector estándar. La reacción de una nucleobase divulgada en la presente memoria que porta un enlazador reactivo (después de la desprotección, si es necesario) y un reactivo de carga que incorpora una carga y un grupo de reactividad complementaria al enlazador reactivo, sirve para proporcionar una nucleobase unida a una carga a través de una fracción enlazadora acoplada.

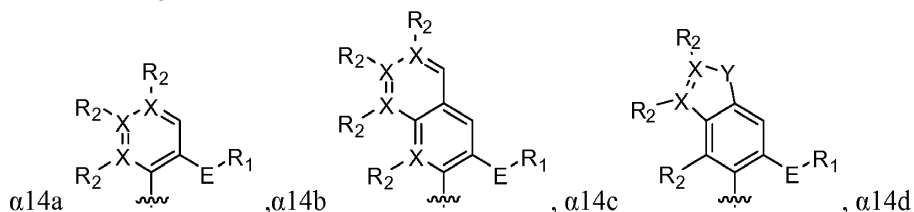
Las composiciones de esta divulgación, en diversas realizaciones, permiten la expansión del repertorio limitado de funcionalidad de los cuatro nucleótidos de ADN naturales y de los cuatro nucleótidos de ARN naturales para incluir prácticamente cualquier funcionalidad de interés, incorporada específicamente en un oligonucleótido de ADN o ARN o en un análogo de ADN o un análogo de ARN, tal como un PNA o un LNA o similar. La carga opcionalmente incluye funcionalidad que permite el reconocimiento molecular alterado (es decir, para el desarrollo de aptámeros), la reactividad alterada (incluso para catálisis) y/o sondas para la visualización y/o caracterización (es decir, para el desarrollo de diagnósticos).

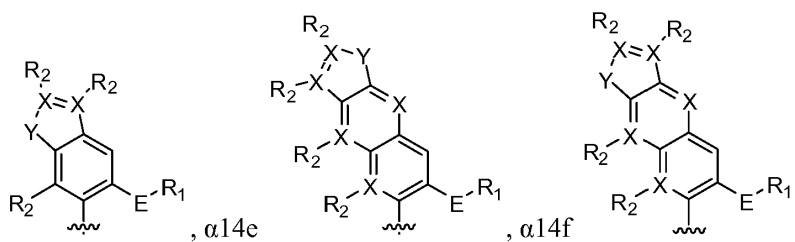
En la presente memoria se divulga, en diversas realizaciones, un compuesto que comprende un análogo de nucleobase de cualquiera de las fórmulas siguientes $\beta 8a$ o $\beta 8b$:



en las que cada X es independientemente carbono o nitrógeno; en las que cada R_2 es opcional y cuando está presente es independientemente hidrógeno, alquilo, alqueno, alquino, metoxi, metanotiol, metanoseleno, halógeno, ciano, grupo azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio, o una amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno; y en las que el análogo de la nucleobase no es 4TFP o 7TFP o una derivatización del enlazador del mismo.

En la presente memoria se divulga, en diversas realizaciones, un compuesto que comprende un análogo de nucleobase de cualquiera de las fórmulas siguientes:





5

10

15

20

en las que cada X es independientemente carbono o nitrógeno; en las que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, alquilo, alqueno, alquino, metoxi, metanotiol, metanoseleno, halógeno, ciano, grupo azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno; y en las que el análogo de la nucleobase no es FIMO, MIMO, FEMO, PrMO, EMO, MEMO, IMO, MMO2, DMO, NMO, 5FM, 20Me, TMO, FDMO, VMO, ZMO, CIMO, TfMO, CNMO, NaM o QMO.

25

Aquí y a lo largo de la memoria, la línea ondulada indica un punto de unión a una fracción ribosilo, desoxirribosilo o didesoxirribosilo; o a un análogo de una fracción ribosilo, desoxirribosilo o didesoxirribosilo, tal como un análogo de ribosa bloqueado, un grupo peptídico o similar. En algunas realizaciones, la fracción ribosilo, desoxirribosilo o didesoxirribosilo o análogo de la misma está en forma libre, conectada a un grupo monofosfato, difosfato o trifosfato; que comprende opcionalmente un grupo α -tiotrifosfato, β -tiotrifosfato, o γ -tiotrifosfato; o se incluye en un ARN o un ADN o en un análogo de ARN o un análogo de ADN.

30

En algunas realizaciones, cuando se refiere a una fracción ribosilo o una fracción desoxirribosilo de una nucleobase no natural divulgado en la presente memoria, se usa X, dX o (d)X, por ejemplo, dTPT3 o (d) TPT3 se refiere a la nucleobase TPT3 unida a cualquiera de las opciones en la posición indicada por la línea ondulada. Por lo tanto, la denominación general de dX se refiere a compuestos que tienen análogos de ribosa o desoxirribosa unidos al mismo como se indica por la línea ondulada. Cuando se refiere específicamente a un nucleósido de ribosilo, el prefijo "d" se elimina, es decir, TPT3 se refiere a una forma de ribosilo. Cuando se incorpora a un sustrato de polimerasa trifosfato (es decir, TPT3TP, dTPT3TP), se considera que el nucleótido o las variantes derivadas del enlazador se incorporan a un oligonucleótido de ARN o ADN mediante el uso de una ARN o ADN polimerasa, respectivamente.

35

40

En algunas realizaciones, un análogo de ribosilo, desoxirribosilo o didesoxirribosilo de un análogo de nucleósido divulgado en la presente memoria (por ejemplo, $\beta 8a$, $\beta 8b$, $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$) comprende un grupo funcional 2'. Los ejemplos de grupos funcionales incluyen, sin limitación, metoxi, halógeno, -O-alilo, -O-metoxietilo, amina primaria, -O-dinitrofenol, -O-dinitrofenil éter, alquilo, -O-alquilo, tiol, aminoetoximetilo, aminopropoximetilo, grupos aminoetilo, cianoetilo y guanidinoetilo. En algunas realizaciones, el análogo ribosilo, desoxirribosilo o didesoxirribosilo comprende una sustitución 4'-tio (por ejemplo, el oxígeno de la fracción de azúcar se reemplaza con un azufre).

45

En algunas realizaciones, un grupo alquilo de un análogo de nucleobase incluye, sin limitación, un grupo metilo, etilo, propilo e isopropilo. En algunas realizaciones, un grupo halógeno de un análogo de nucleobase incluye, sin limitación, flúor, cloro, bromo, y yodo.

50

En algunas realizaciones, un enlazador reactivo de un análogo de nucleobase comprende un grupo funcional que incluye, pero no se limita a, alquilo, alqueno, alquino, fenilo, bencilo, halo, hidroxilo, carbonilo, aldehído, haloformilo, éster de carbonato, carboxilato, carboxilo, éster, metoxi, hidroperoxi, peroxi, éter, hemiacetal, hemiketal, acetal, cetal, ortoéster, metilendioxi, éster de ortocarbonato, carboxamida, amina primaria, amina secundaria, imida, azida, azo, cianato, isocianato, nitrato, nitrilo, isonitrilo, oxinitroso, nitro, nitroso, piridilo, sulfhidrilo, sulfuro, disulfuro, sulfinilo, sulfuro, tiocianato, isotiocianato, carbonotioilo, fosfino, fosfono, fosfato, borono, boronato, borino, borinato y una combinación de los mismos. En algunas realizaciones, un enlazador reactivo de un análogo de nucleobase comprende un grupo amino, un grupo acetilénico, un grupo tiol, un grupo aldehído o un grupo azida.

55

60

65

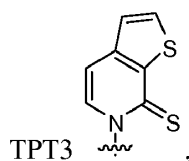
En algunas realizaciones, una fracción ribosilo o desoxirribosilo comprende un grupo trifosfato o α -tiotrifosfato unido a un 5'-hidroxilo del mismo. En algunas realizaciones, se incorpora una fracción ribosilo o desoxirribosilo en una cadena de oligonucleótidos de ARN o ADN, respectivamente, o la fracción ribosilo o desoxirribosilo o análogo del mismo se incorpora en un análogo de ARN o ADN. El análogo de ARN o análogo de ADN, en ciertas realizaciones, es un ácido nucleico peptídico (PNA) o un ácido nucleico bloqueado (LNA). El análogo de ARN o análogo de ADN, en ciertas realizaciones, es un derivado bicíclico. Los derivados bicíclicos incluyen, sin limitación, ácido nucleico con puente de 2'-O,4'-C-etileno (ENA), ácido nucleico bloqueado carbocíclico (CLNA), ácido nucleico de ciclohexeno (CENA) y ácido nucleico bloqueado con 2'-desoxi-2'-N,4'-C-etileno (AENA). En ciertas realizaciones, el análogo de ARN o el análogo de ADN es un derivado acíclico. En ciertas realizaciones, el análogo de ARN o el análogo de ADN es un ácido nucleico

desbloqueado (UNA). En ciertas realizaciones, el análogo de ARN o el análogo de ADN comprende un anillo de piranososa en lugar de una ribosa. En ciertas realizaciones, el análogo de ARN o el análogo de ADN es un ácido nucleico arabino (ANA) o un ácido nucleico de hexitol (HNA).

- 5 En algunas realizaciones, una fracción ribosilo o desoxirribosilo o análogo del mismo se sustituye con grupos protectores y activadores adecuados para uso en una máquina automatizada de síntesis de oligonucleótidos químicos. Un ejemplo de una máquina automatizada de síntesis de oligonucleótidos químicos es una máquina de síntesis de fosforoamidita.
- 10 En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente un grupo -C=C-CH₂NHR₃, en la que R₃ es hidrógeno o es un grupo protector de amino. Un ejemplo de un grupo protector de amino es un grupo dicloroacetilo. En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente un grupo acetilénico adecuado para usar en una reacción de clic con un reactivo de carga que comprende una carga y un grupo reactivo acetileno. En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente un grupo tiol adecuado para usar en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo tiol. En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente un grupo aldehído adecuado para usar en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo aldehído. En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente un grupo azida adecuado para usar en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo azida. En algunas realizaciones, al menos un R2 de un análogo de nucleobase comprende independientemente -C=C-(CH₂)_n-C≡CH, en la que n es 1, 2, 3, 4, 5 o 6; o R2 es -C≡C-(CH₂)_{n1}-O-(CH₂)_{n2}-C≡CH, en la que n1 y n2 son cada uno independientemente 1, 2 o 3. En algunas realizaciones, al menos un R2 es independientemente un enlazador acoplado unido a una carga por reacción de un grupo amino y un grupo amino-reactivo. Un ejemplo de un grupo amino-reactivo es un grupo acilante o un grupo alquilante, o es un éster de N-hidroxisuccinimida. En algunas realizaciones, al menos un R2 es independientemente un enlazador acoplado unido a una carga por reacción de un grupo acetileno y un grupo reactivo acetileno. Un ejemplo de un grupo reactivo acetileno es un grupo azida. En ciertas realizaciones, el grupo acetileno y el grupo azida se acoplan con una reacción de clic catalizada por cobre. En algunas realizaciones, al menos un R2 es independientemente un enlazador acoplado unido a una carga por reacción de un grupo tiol y un grupo reactivo tiol. En algunas realizaciones, al menos un R2 es independientemente un enlazador acoplado unido a una carga por reacción de un aldehído y un grupo reactivo aldehído. En algunas realizaciones, al menos un R2 es independientemente un enlazador acoplado unido a una carga por reacción de una azida y un grupo reactivo azida. Un ejemplo de un grupo reactivo azida es un alquino terminal o un ciclooctino tensado. En algunas realizaciones, al menos un R2 es independientemente hidrógeno, el compuesto que comprende un grupo α-tiotrifosfato, en el que un reactivo de carga que comprende un grupo carbonilo γ-bromo-α,β-insaturado, yodo, bromoacetilo o aziridinilsulfonamida se acopla al mismo.
- 20
- 25
- 30
- 35

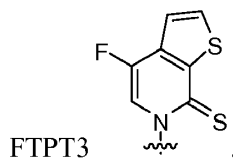
En algunas realizaciones, una carga de un análogo de nucleobase incluye, sin limitación, proteínas, péptidos, aminoácidos, oligonucleótidos, fármacos de molécula pequeña, grupos alifáticos, compuestos que comprenden grupos fotorreactivos, compuestos que comprenden grupos químicamente reactivos, compuestos que comprenden grupos catalíticos, compuestos que comprenden clorometilcetonas, lípidos, biotina, compuestos fluorescentes, compuestos de inactivación de la fluorescencia, liposomas y nanopartículas.

Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase



y derivados y análogos del mismo.

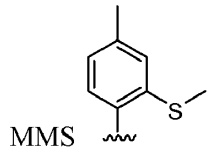
Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase



y derivados y análogos del mismo.

Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase

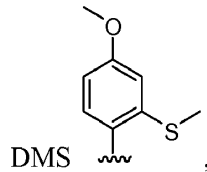
5



y derivados y análogos del mismo.

10 Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase

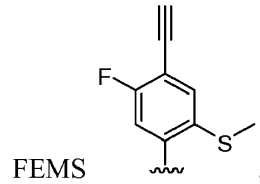
15



y derivados y análogos del mismo.

20 Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase

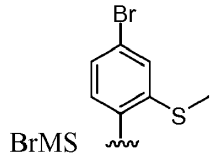
25



y derivados y análogos del mismo.

Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase

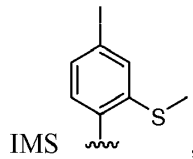
35



y derivados y análogos del mismo.

Se divulga además en la presente memoria, en diversas realizaciones, el análogo de nucleobase

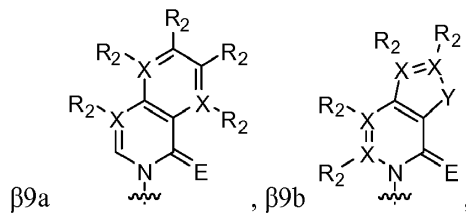
45



y derivados y análogos del mismo.

En la presente memoria se divulgan, en algunas realizaciones, pares de nucleobase que comprenden un primer análogo de nucleobase que tiene cualquiera de las fórmulas β9a o β9b; y un segundo análogo de nucleobase que tiene cualquiera de las fórmulas α15a o α15b:

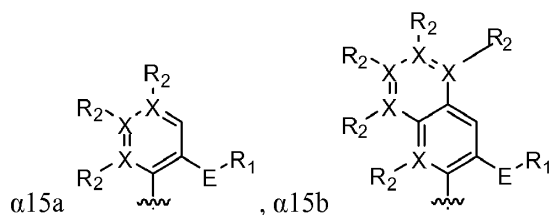
55



60

65

5



10

15

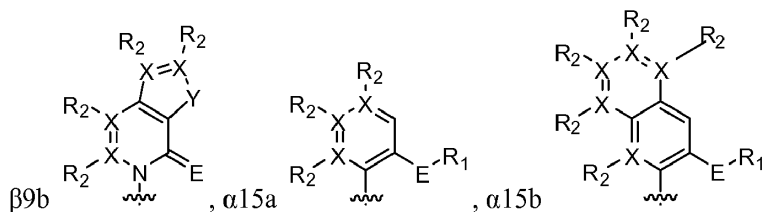
20

en las que cada X es independientemente carbono o nitrógeno; en las que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente oxígeno, azufre o selenio; y en el que el par de nucleobases no es 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-dlmo, d5SICS-dMMO2, d5SICS-dDMO, d5SICS-dNMO, d5SICS-d5FM, d5SICS-d2OMe, d5SICS-dTMO, d5SICS-dFDMO, d5SICS-dVMO, d5SICS-dZMO, d5SICS-dCIMO, d5SICS-dTfMO y d5SICS-dCNMO.

25

En la presente memoria, en algunas realizaciones, se divulga un par de nucleobase que comprende un primer análogo de nucleobase que tiene la fórmula $\beta9b$, y un segundo análogo de nucleobase que tiene cualquiera de las fórmulas $\alpha15a$ o $\alpha15b$:

30



35

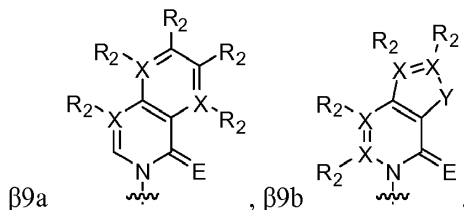
40

en las que cada X es independientemente carbono o nitrógeno; en las que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio, o amina secundaria; y en las que cada E es independientemente oxígeno, azufre o selenio.

45

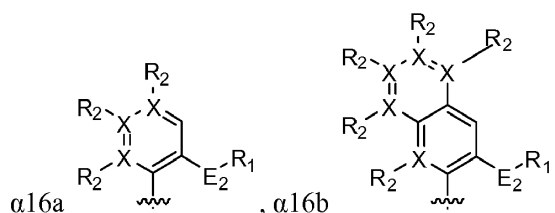
En la presente memoria se divulga, en algunas realizaciones, un par de nucleobase que comprende un primer análogo de nucleobase que tiene cualquiera de las fórmulas $\beta9a$ o $\beta9b$, y un segundo análogo de nucleobase que tiene cualquiera de las fórmulas $\alpha16a$ o $\alpha16b$:

50



55

60



65

en las que cada X es independientemente carbono o nitrógeno; en las que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que

comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio, o amina secundaria; en las que cada E es independientemente oxígeno, azufre o selenio; y en las que cada E₂ es independientemente azufre o selenio.

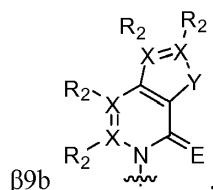
La línea ondulada indica un punto de unión a una fracción ribosilo, desoxirribosilo o didesoxirribosilo; o a un análogo de una fracción ribosilo, desoxirribosilo o didesoxirribosilo, tal como un análogo de ribosa bloqueado, un grupo peptídico o similar. En algunas realizaciones, la fracción ribosilo, desoxirribosilo o didesoxirribosilo o análogo de la misma está en forma libre, conectada a un grupo monofosfato, difosfato o trifosfato; que comprende opcionalmente un grupo α-tiotrifosfato, β-tiotrifosfato, o γ-tiotrifosfato; o se incluye en un ARN o un ADN o en un análogo de ARN o un análogo de ADN.

En algunas realizaciones, un grupo alquinilo es un grupo etinilo o propinilo. En algunas realizaciones, un grupo alquilo es un grupo metilo, etilo, propilo o isopropilo. En algunas realizaciones, un halógeno es flúor, cloro, bromo o yodo.

En algunas realizaciones, un análogo de ribosilo, desoxirribosilo, o didesoxirribosilo del par nucleobase comprende un grupo funcional 2'. Los grupos funcionales ejemplares incluyen, sin limitación, metoxi, halógeno, -O-alilo, -O-metoxietilo, amina primaria, alquilo, -O-alquilo, tiol, -O-dinitrofenol, -O-dinitrofenil éter, aminoetoximetilo, aminopropoximetilo, aminoetilo, grupos cianoetilo, y guanidinoetilo. En algunas realizaciones, un análogo de ribosilo, desoxirribosilo o didesoxirribosilo del par de nucleobase comprende una sustitución 4'-tio.

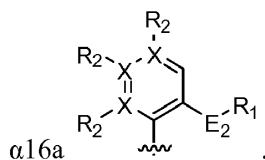
En algunas realizaciones, un par de nucleobase comprende una nucleobase que tiene la fórmula α15b. En algunas realizaciones, un par de nucleobase comprende una nucleobase que tiene la fórmula α15a. En algunas realizaciones, un par de nucleobase comprende una nucleobase que tiene la fórmula β9b. En algunas realizaciones, un par de nucleobase comprende una nucleobase que tiene la fórmula β9b, en la que cada X es carbono, Y es azufre, cada R₂ es hidrógeno y E es azufre. En algunas realizaciones, un par de nucleobase comprende una nucleobase que tiene la fórmula α15b, en la que cada X es carbono, cada R₂ es hidrógeno, R₁ es un grupo metilo y E es oxígeno. En algunas realizaciones, un par de nucleobase comprende un primer análogo de nucleobase que tiene la fórmula β9a y un segundo análogo de nucleobase que tiene la fórmula α16a. En algunas realizaciones, un par de nucleobase comprende un primer análogo de nucleobase que tiene la fórmula β9a y un segundo análogo de nucleobase que tiene la fórmula α16b. En algunas realizaciones, un par de nucleobase comprende un primer análogo de nucleobase que tiene la fórmula β9b y un segundo análogo de nucleobase que tiene la fórmula α16a. En algunas realizaciones, un par de nucleobase comprende un primer análogo de nucleobase que tiene la fórmula β9b y un segundo análogo de nucleobase que tiene la fórmula α16b.

En la presente memoria se divulga, en ciertas realizaciones, un par de nucleobase que comprende un primer análogo de nucleobase que tiene la fórmula β9b y un segundo análogo de nucleobase que tiene la fórmula β9b:



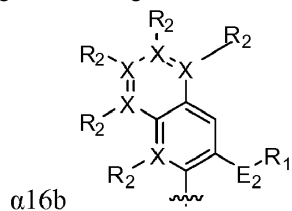
en la que cada X es independientemente carbono o nitrógeno; en la que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en la que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; y en la que cada E es independientemente oxígeno, azufre o selenio. En algunas realizaciones, el par de nucleobase es un par de homo-nucleobase.

En la presente memoria se divulga, en ciertas realizaciones, un par de nucleobase que comprende un primer análogo de nucleobase que tiene la fórmula α16a y un segundo análogo de nucleobase que tiene la fórmula α16a:



en la que cada X es independientemente carbono o nitrógeno; en la que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en la que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alqueno, alquino, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; y en la que cada E₂ es independientemente azufre o selenio. En algunas realizaciones, el par de nucleobase es un par de homo-nucleobase.

En la presente memoria se divulga, en ciertas realizaciones, un par de nucleobase que comprende un primer análogo de nucleobase que tiene la fórmula α16b y un segundo análogo de nucleobase que tiene la fórmula α16b:



en la que cada X es independientemente carbono o nitrógeno; en la que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en la que cada R₂ es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alqueno, alquino, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; y en la que cada E₂ es independientemente azufre o selenio. En algunas realizaciones, el par de nucleobase es un par de homo-nucleobase.

En la presente memoria se divulga, en ciertas realizaciones, un par de nucleobase que comprende un primer análogo de nucleobase que tiene cualquiera de las fórmulas β9a, β9b, α15a, α15b, α16a o α16b; y una segunda nucleobase seleccionada del grupo que consiste en citosina, guanina, adenina, timina, uracilo, 2-aminoadenina-9-ilo, 2-aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, 2-propilo y derivados de alquilo de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-deazaguanina, 3-deazaadenina, 4-tio-uracilo, 4-tio-timina, uracil-5-ilo, hipoxantina-9-ilo (I), 5-metilcitosina, 5-hidroximetil citosina, xantina, hipoxantina, 5-bromo, y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5-sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxiourea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo, y 5-yodouracilo, 6-alquil derivados de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azo citosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-deazaguanina, 7-deazaguanosina, 7-deaza-adenina, 7-deaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y adeninas y guaninas 8-hidroxilo sustituidas; N4-etilcitosina, purinas N-2 sustituidas, purinas N-6 sustituidas, purinas O-6 sustituidas, aquellas que aumentan la estabilidad de la formación dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos expandidos por tamaño, fluorados ácidos nucleicos, pirimidinas tricíclicas, fenoxazina citidina([5,4-b][1,4] benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), abrazaderas G, fenoxazina citidina (9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido[3':2':4,5]pirrolo[2,3-d]pirimidin-2-ona), 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidroxouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido uracil-5oxiacético, wyutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido 5-oxaacético-uracilo metiléster, ácido 5-oxiacético-uracilo, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza con un heterociclo.

Un par de bases que comprende una o más nucleobases no naturales se ejemplifica mediante el par de bases no natural dTPT3PA-dNaM (es decir, el par formado entre dTPT3PA y dNaM; Figura 1A y 1B). Además, la reactividad ortogonal de los diferentes centros reactivos/enlazadores desarrollados (es decir, fosforotioatos, aminos y alquinos) permite el arreglo selectivo de diferentes fracciones en el mismo oligonucleótido (ADN o ARN). Otra composición se ejemplifica adicionalmente con el par de bases no natural dTPT3PA-dMMO2pCO, en el que, en diversas realizaciones, el grupo alquino de dMMO2pCO se usa para unir un grupo funcional mediante cicloadición de azida-alquino catalizada con cobre (I) (CuAAC) y, después de la desprotección de dTPT3PA, la amina libre se usa para unir un grupo funcional diferente a través del acoplamiento de N-hidroxisuccinimida (NHS).

Si bien se informan y se derivatizan con enlazadores otros varios pares de bases no naturales, los pares derivatizados con enlazadores descritos aquí, dTPT3PA-dNaM, d5SICSCO-dNaM, dTPT3PA-dMMO2pCO (Figura 2), se validan más y se replican mejor dentro del ADN y se transcriben mejor en ARN. En particular, dTPT3PA-dNaM, se replica bien y, por lo tanto, es adecuado para uso en la práctica de los procedimientos divulgados y reivindicados en la presente memoria.

5

En la presente memoria se divulgan pares de bases no naturales que comprenden dTPT3, en los que dTPT3 en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dTPT3 o dTPT3 derivatizados con el enlazador (por ejemplo, dTPT3PA) incluyen, sin limitación, 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-dDMMAN, 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, y dTPT3-CNMO, en los que la base complementaria dTPT3 está o no derivatizada con el enlazador (por ejemplo, dMMO2pCO). En la Figura 9 se ilustra dTPT3 como un análogo β . En la Figura 2 se ilustra un ejemplo de un dTPT3 derivatizado con el enlazador, en el que, en algunos casos, R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarias a dTPT3 incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, el dTPT3 o un dTPT3 derivatizado con el enlazador tiene las bases apareadas con dTPT3 o un dTPT3 derivatizado con el enlazador, para formar un par homo-nucleobase. En algunas realizaciones, el dTPT3 o un dTPT3 derivatizado con el enlazador tiene las bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12 y 13, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector, por ejemplo, dTPT3PA. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, por ejemplo, dTPT3A, en el que, en algunos casos, se eliminó un grupo protector.

25

30

En la presente memoria se divulgan pares de bases no naturales que comprenden dMMS, en los que dMMS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dMMS o dMMS derivatizado con el enlazador (por ejemplo, dMMSPA) incluyen, sin limitación, 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, y dFTPT3-dMMS, en los que la base complementaria dMMS se derivatiza o no del enlazador (por ejemplo pFTPT3pA). En la Figura 11 dMMS se ilustra como un análogo α 14a. En algunas realizaciones, un dMMS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dMMS incluyen, sin limitación, los análogos β o análogos β derivatizados con el enlazador ilustrados en las Figuras 9, 12 y 13. En algunas realizaciones, dMMS o un dMMS derivatizado con el enlazador tiene bases apareadas con dMMS o un dMMS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dMMS o un dMMS derivatizado con el enlazador tiene bases apareadas con una nucleobase α , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 8, 10, 11, y 15, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

35

40

45

50

En la presente memoria se divulgan pares de bases no naturales que comprenden dDMS, en los que el dDMS en algunos casos se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dDMS o dDMS derivatizado con el enlazador (por ejemplo, dDMSPA) incluyen, sin limitación, 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, en los que la base complementaria de dDMS se derivatiza con el enlazador o no (por ejemplo pFTPT3pA). En la Figura 11 se ilustra dDMS como un análogo α 14a. En algunas realizaciones, un dDMS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dDMS incluyen, sin limitación, análogos β o análogos β derivatizados con el enlazador ilustrados en las Figuras 9, 12 y 13. En algunas realizaciones, dDMS o un dDMS derivatizado con el enlazador tiene bases apareadas con dDMS o un dDMS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dMMS o un dMMS derivatizado con el enlazador tiene bases apareadas con una nucleobase α , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 8,

55

60

65

10, 11, y 15, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

5 En la presente memoria se divulgan pares de bases no naturales que comprenden dFEMS, en los que dFEMS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dFEMS o dFEMS derivatizado con el enlazador (por ejemplo, DFEMSPA) incluyen, sin limitación, 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, en los que la base complementaria de dFEMS derivatizado con el enlazador o no (por ejemplo, pFTPT3pA). En la Figura 11 se ilustra dFEMS como un análogo α 14a. En algunas realizaciones, un dFEMS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dFEMS, incluyen, sin limitación, análogos β o análogos β derivatizados con el enlazador ilustrados en las Figuras 9, 12 y 13. En algunas realizaciones, dFEMS o un dFEMS derivatizado con el enlazador tiene bases apareadas con dFEMS o un dFEMS derivatizado con el enlazador, para formar un par homo-nucleobase. En algunas realizaciones, dFEMS o un dFEMS derivatizado con el enlazador tiene bases apareadas con una nucleobase α , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 8, 10, 11 y 15, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

25 En la presente memoria se divulgan pares de bases no naturales que comprenden dBrMS, en los que dBrMS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dBrMS o dBrMS derivatizado con el enlazador (por ejemplo, dBrMSPA) incluyen, sin limitación, 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, en los que la base complementaria de dBrMS se derivatiza con el enlazador o no (por ejemplo, pFTPT3pA). En la Figura 11 se ilustra dBrMS como un análogo α 14a. En algunas realizaciones, un dBrMS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dBrMS incluyen, sin limitación, análogos β o análogos β derivatizados con el enlazador ilustrados en las Figuras 9, 12 y 13. En algunas realizaciones, dBrMS o un dBrMS derivatizada con el enlazador tiene bases apareadas con dBrMS o un dBrMS derivatizada con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dBrMS o un dBrMS derivatizada con el enlazador tiene bases apareadas con una nucleobase α , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 8, 10, 11 y 15, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

45 En la presente memoria se divulgan pares de bases no naturales que comprenden dIMS, en los que dIMS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dIMS o dIMS derivatizado con el enlazador (por ejemplo, dIMSPA) incluyen, sin limitación, 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, en los que base complementaria dIMS se derivatiza con el enlazador o no (por ejemplo, pFTPT3pA). En la Figura 11 se ilustra dIMS como un análogo α 14a. En algunas realizaciones, un dIMS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dIMS incluyen, sin limitación, análogos β o análogos β derivatizados con el enlazador ilustrados en las Figuras 9, 12 y 13. En algunas realizaciones, dIMS o un dIMS derivatizado con el enlazador tiene bases apareadas con dIMS o un dIMS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dIMS o un dIMS derivatizado con el enlazador tiene bases apareadas con una nucleobase α , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 8, 10, 11 y 15, o una nucleobase que de derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

60 En la presente memoria se divulgan pares de bases no naturales que comprenden dICS, en los que dICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dICS o dICS derivatizado con el enlazador (por ejemplo, dICSPA) incluyen, sin limitación, 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, en los que la base

complementaria de dICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dICS como un análogo β . En algunas realizaciones, un dICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, dICS o un dICS derivatizado con el enlazador tiene bases apareadas con dICS o un dICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dICS o un dICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12 y 13, o una nucleobase que se derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden dPICS, en los que dPICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dPICS o dPICS derivatizado con el enlazador (por ejemplo, dPICSPA) incluyen, sin limitación, 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, en los que la base complementaria dPICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dPICS como un análogo β . En algunas realizaciones, un dPICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dPICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, dPICS o un dPICS derivatizado con el enlazador tiene bases apareadas con dPICS o un dPICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dPICS o un dPICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12 y 13, o una nucleobase que se derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden dMICS, en los que dMICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dMICS o dMICS derivatizado con el enlazador (por ejemplo, dMICSPA) incluyen, sin limitación, 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, en los que la base complementaria dMICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dMICS como un análogo β . En algunas realizaciones, un dMICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dMICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, dMICS o un dMICS derivatizado con el enlazador tiene bases apareadas con dMICS o un dMICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dMICS o un dMICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12 y 13, o una nucleobase que se derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden d4MICS, en los que d4MICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden d4MICS o d4MICS derivatizado con el enlazador (por ejemplo, d4MICSPA) incluyen, sin limitación, 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, en los que la base complementaria d4MICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra d4MICS como un análogo β . En algunas realizaciones, un d4MICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a d4MICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, d4MICS o un d4MICS derivatizado con el enlazador tiene bases apareadas con d4MICS o un d4MICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, d4MICS o un d4MICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden d5MICS, en los que d5MICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden d5MICS o d5MICS derivatizado con el enlazador (por ejemplo, d5MICS-PA) incluyen, sin limitación, 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, en los que la base complementaria d5MICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra d5MICS como un análogo $\beta 2$. En algunas realizaciones, un d5MICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a d5MICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11 y 15. En algunas realizaciones, d5MICS o un d5MICS derivatizado con el enlazador tiene bases apareadas con d5MICS o un d5MICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, d5MICS o un d5MICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden dNICS, en los que dNICS en algunos casos se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dNICS o dNICS derivatizados con el enlazador (por ejemplo, dNICS-PA) incluyen, sin limitación, 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-dMMO2, dNICS-d2OMe, dNICS-dTfMO, dNICS-dCNMO, dNICS-dNAM, dNICS-dQMO, en los que la base complementaria dNICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dNICS como un análogo $\beta 3$. En algunas realizaciones, un dNICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dNICS, incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, dNICS o un dNICS derivatizado con el enlazador tiene bases apareadas con dNICS o un dNICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dNICS o un dNICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que se derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden dONICS, en los que dONICS en algunos casos se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dONICS o dONICS derivatizado con el enlazador (por ejemplo, dONICS-PA) incluyen, sin limitación, 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-dMMO2, dONICS-d2OMe, dONICS-dTfMO, dONICS-dCNMO, dONICS-dNAM, dONICS-dQMO, en los que la base complementaria dONICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dONICS como un análogo $\beta 3$. En algunas realizaciones, un dONICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dONICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, dONICS o un dONICS derivatizado con el enlazador tiene bases apareadas con dONICS o un dONICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dONICS o un dONICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que se derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

En la presente memoria se divulgan pares de bases no naturales que comprenden dSICS, en los que dSICS en algunos casos se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dSICS o dSICS derivatizado con el enlazador (por ejemplo, dSICS-PA) incluyen, sin limitación, 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, en los que la base complementaria de dSICS se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dSICS como un análogo $\beta 5$. En algunas realizaciones, un dSICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dSICS, incluyen, sin limitación, análogos α o análogos α derivatizados con el

5 enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, dSICS o un dSICS derivatizado con el enlazador tiene bases apareadas con dSICS o un dSICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dSICS o un dSICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que de derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

10 En la presente memoria se divulgan pares de bases no naturales que comprenden dSNICS, en los que dSNICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden dSNICS o dSNICS derivatizado con el enlazador (por ejemplo, dSNICSPA) incluyen, sin limitación, 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, en los que la base se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra dSNICS como un análogo β 5. En algunas realizaciones, un dSNICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a dSNICS, incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, dSNICS o un dSNICS derivatizado con el enlazador tiene bases apareadas con dSNICS o un dSNICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, dSNICS o un dSNICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que de derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

30 En la presente memoria se divulgan pares de bases no naturales que comprenden d4SICS, en los que d4SICS en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden d4SICS o d4SICS derivatizado con el enlazador (por ejemplo, D4SICSPA) incluyen, sin limitación, 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, en los que la base complementaria d4SICS se derivatiza del enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra d4SICS como un análogo β 5. En algunas realizaciones, un d4SICS derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a d4SICS incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, d4SICS o un d4SICS derivatizado con el enlazador tiene bases apareadas con d4SICS o un d4SICS derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, d4SICS o un d4SICS derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que de derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

45 En la presente memoria se divulgan pares de bases no naturales que comprenden d7OFP, en los que d7OFP en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden d7OFP o d7OFP derivatizado con el enlazador (por ejemplo, d7OFPFA) incluyen, sin limitación, 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, en los que la base complementaria d7OFP se derivatiza con el enlazador o no (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra d7OFP como un análogo β 5. En algunas realizaciones, un d7OFP derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a d7OFP incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, d7OFP o un d7OFP derivatizado con el enlazador tiene bases apareadas con d7OFP o un d7OFP derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, d7OFP o un d7OFP derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase que de derivatiza de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

65 En la presente memoria se divulgan pares de bases no naturales que comprenden d7OTP, en los que d7OTP en algunos casos, se derivatiza con el enlazador. Los pares de bases no naturales que comprenden d7OTP o d7OTP derivatizado con el enlazador (por ejemplo, d7OTPPA) incluyen, sin limitación, 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, en los que la base complementaria d7OTP se derivatiza con el enlazador (por ejemplo, dDMOpCO, dDMOpCC). En la Figura 9 se ilustra d7OTP como un análogo β 5. En algunas realizaciones, un d7OTP derivatizado con el enlazador comprende un grupo funcional R, en el que R es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga o R es un enlazador acoplado al que se une una carga. Los análogos de nucleobase que son complementarios a d7OTP incluyen, sin limitación, análogos α o análogos α derivatizados con el enlazador ilustrados en las Figuras 8, 10, 11, y 15. En algunas realizaciones, d7OTP o un d7OTP derivatizado con el enlazador tiene bases apareadas con d7OTP o un d7OTP derivatizado con el enlazador, para formar un par de homo-nucleobase. En algunas realizaciones, d7OTP o un d7OTP derivatizado con el enlazador tiene bases apareadas con una nucleobase β , que incluye pero no se limita a, cualquier nucleobase ilustrada en las Figuras 9, 12, y 13, o una nucleobase derivatizada de la misma. En algunas realizaciones, una fracción enlazadora se protege con un grupo protector. En algunas realizaciones, una fracción enlazadora no se protege con un grupo protector, en la que en algunos casos, se eliminó un grupo protector.

Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase TPT3 y una segunda nucleobase MMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase TPT3 y una segunda nucleobase DMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase TPT3 y una segunda nucleobase FEMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase TPT3 y una segunda nucleobase BrMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase TPT3 y una segunda nucleobase IMS.

Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase FTPT3 y una segunda nucleobase MMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase FTPT3 y una segunda nucleobase DMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase FTPT3 y una segunda nucleobase FEMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase FTPT3 y una segunda nucleobase BrMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase FTPT3 y una segunda nucleobase IMS.

Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase 5SICS y una segunda nucleobase MMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase 5SICS y una segunda nucleobase DMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase 5SICS y una segunda nucleobase FEMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase 5SICS y una segunda nucleobase BrMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase 5SICS y una segunda nucleobase IMS.

Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase SICS y una segunda nucleobase MMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase SICS y una segunda nucleobase DMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase SICS y una segunda nucleobase FEMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase SICS y una segunda nucleobase BrMS. Además, en la presente memoria se divulga, en diversas realizaciones, un par de nucleobase que comprende una primera nucleobase SICS y una segunda nucleobase IMS.

Además, en la presente memoria se divulga, en diversas realizaciones, un oligonucleótido bicatenario dúplex en el que una primera cadena oligonucleotídica incluye una nucleobase no natural (por ejemplo, un análogo de nucleobase) divulgada en la presente memoria, y una segunda cadena oligonucleotídica complementaria que comprende una nucleobase complementaria de emparejamiento de bases en un sitio complementario de emparejamiento de bases de la misma. Por ejemplo, para dTPT3, una segunda cadena oligonucleotídica complementaria comprende dNaM, dDMO o dMMO2, o un análogo de la misma derivatizado con el enlazador, en un sitio complementario de emparejamiento de bases. De esta manera, la interacción de emparejamiento entre la primera cadena oligonucleotídica y la segunda cadena oligonucleotídica incluye una interacción específica de emparejamiento de nucleobase entre una fracción de nucleobase no natural divulgada en la presente memoria y una nucleobase complementaria, que puede ser una nucleobase natural o no natural.

En la presente memoria se divulga, en algunas realizaciones, un oligonucleótido bicatenario dúplex en el que una primera cadena oligonucleotídica comprende un compuesto que tiene la fórmula β 8a o β 8b, y una segunda cadena oligonucleotídica complementaria comprende una nucleobase complementaria de emparejamiento de bases en un sitio complementario de emparejamiento de bases de la misma. En algunas realizaciones, la nucleobase complementaria de emparejamiento de bases es un compuesto que tiene la fórmula α 14a, α 14b, α 14c, α 14d, α 14e o α 14f. En algunas

realizaciones, la nucleobase complementaria de emparejamiento de bases incluye, sin limitación, citosina, guanina, adenina, timina, uracilo, 2-aminoadenina-9-ilo, 2-aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, 2-propilo y derivados de alquilo de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-deazaguanina, 3-deazaadenina, 4-tio-uracilo, 4-tio-timina, uracil-5-ilo, hipoxantina-9-ilo (I), 5-metil-citosina, 5-hidroximetil citosina, xantina, hipoxantina, 5-bromo y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5-sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxiaurea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo y 5-yodouracilo, 6-alquil derivados de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azo citosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-deazaguanina, 7-deazaguanosina, 7-deaza-adenina, 7-deaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y 8-hidroxilo sustituido adeninas y guaninas; N4-etilcitosina, purinas N-2 sustituidas, purinas N-6 sustituidas, purinas O-6 sustituidas, aquellas que aumentan la estabilidad de la formación dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos expandidos por tamaño, ácidos nucleicos fluorados, pirimidinas tricíclicas, fenoxazina citidina ([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazina-2(3H)-ona), abrazaderas G, fenoxazina citidina (9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido [3',2':4,5]pirrolo [2,3-d]pirimidin-2-ona), 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouacilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metil-tio-N6-isopenteniladenina, ácido uracil-5oxiacético, wibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido uracil-5-oxiacético metiléster, ácido uracil-5-oxiacético, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza con un heterociclo.

En la presente memoria se divulga, en algunas realizaciones, un oligonucleótido bicatenario dúplex en el que una primera cadena oligonucleotídica comprende un compuesto que tiene la fórmula $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$ o $\alpha 14f$, y una segunda cadena oligonucleotídica complementaria comprende una nucleobase complementaria de emparejamiento de bases en un sitio complementario de emparejamiento de bases de la misma. En algunas realizaciones, la nucleobase complementaria de emparejamiento de bases es un compuesto que tiene la fórmula $\beta 8a$ o $\beta 8b$. En algunas realizaciones, la nucleobase complementaria de emparejamiento de bases incluye, sin limitación, citosina, guanina, adenina, timina, uracilo, 2-aminoadenina-9-ilo, 2-aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, 2-propilo y derivados de alquilo de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-deazaguanina, 3-deazaadenina, 4-tio-uracilo, 4-tio-timina, uracil-5-ilo, hipoxantina-9-ilo (I), 5-metil-citosina, 5-hidroximetil citosina, xantina, hipoxantina, 5-bromo, y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5-sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxiaurea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo, y 5-yodouracilo, 6-alquil derivados de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azo citosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-deazaguanina, 7-deazaguanosina, 7-deaza-adenina, 7-deaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y 8-hidroxilo sustituido adeninas y guaninas; N4-etilcitosina, purinas N-2 sustituidas, purinas N-6 sustituidas, purinas O-6 sustituidas, aquellas que aumentan la estabilidad de la formación dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos expandidos por tamaño, ácidos nucleicos fluorados, pirimidinas tricíclicas, fenoxazina citidina ([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazina-2(3H)-ona), abrazaderas G, fenoxazina citidina (9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido [3',2':4,5]pirrolo [2,3-d]pirimidin-2-ona), 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouacilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metil-tio-N6-isopenteniladenina, ácido uracil-5oxiacético, wibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido uracil-5-oxiacético metiléster, ácido uracil-5-oxiacético, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza con un heterociclo.

En algunas realizaciones, al menos un R2 de una nucleobase en un oligonucleótido bicatenario dúplex es un enlazador acoplado que se une con una carga. En algunas realizaciones, la carga es un grupo indicador, proteína o compuesto que comprende funcionalidad catalítica.

5 En algunas realizaciones, una primera cadena oligonucleotídica que comprende un análogo de nucleobase divulgado en la presente memoria se prepara por síntesis con una nucleobase que comprende un enlazador reactivo, seguido de acoplamiento del reactivo de carga con la primera cadena oligonucleotídica, o en el que la primera cadena oligonucleotídica se prepara por síntesis con una nucleobase que comprende un enlazador acoplado que se une a una carga.

10 En algunas realizaciones, un oligonucleótido bicatenario dúplex tiene una primera cadena que comprende dTPT3 o un derivado del mismo, y una segunda cadena que comprende dNaM, dDMO, o dMMO2 o un derivado del mismo en un sitio complementario de emparejamiento de bases del mismo.

15 Además en la presente memoria se divulga, en diversas realizaciones, un procedimiento para llevar a cabo una funcionalización específica del sitio de un dúplex oligonucleotídico bicatenario, que comprende: incorporar una nucleobase no natural que comprende un enlazador reactivo que comprende un centro reactivo, la nucleobase que tiene cualquiera de las siguientes fórmulas $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$, $\beta 8a$, $\beta 8b$, $\beta 9a$, $\beta 9b$, $\alpha 15a$, $\alpha 15b$, $\alpha 16a$ o $\alpha 16b$, en una primera cadena de oligonucleótidos; después, sintetizando una segunda cadena complementaria a la primera cadena, la segunda cadena comprende una nucleobase complementaria a la nucleobase no natural en una posición complementaria específica del sitio, en condiciones tales que la primera cadena y la segunda cadena forman un dúplex oligonucleotídico de doble cadena; después, contactando el dúplex oligonucleotídico bicatenario que incorpora la nucleobase no natural que comprende la fracción enlazadora reactivo con un reactivo de carga que comprende un cargamento y un grupo de reactividad complementaria, en condiciones adecuadas para que se produzca la reacción del enlazador reactivo y el grupo de reactividad complementaria para producir un enlazador acoplado; para proporcionar el dúplex de oligonucleótidos bicatenario funcionalizado con la carga unida al mismo a través de un enlazador acoplado.

25 Además, en la presente memoria se divulga, en diversas realizaciones, un procedimiento para llevar a cabo una funcionalización específica del sitio de un oligonucleótido bicatenario dúplex, que comprende: incorporar una nucleobase no natural que comprende un enlazador reactivo que comprende un centro reactivo, la nucleobase se seleccionó del grupo que consiste en d5SICSCO, d5SICSCC, dDMOCO, dDMOCC, dMMO2pCO, dMMO2pCC, dTPT3, dTPT3A, dTPT3PA, dTPT3CO, y dTPT3CC, en una primera cadena oligonucleotídica; después, mediante la síntesis de una segunda cadena complementaria a la primera cadena, la segunda cadena comprende una nucleobase complementaria a la nucleobase no natural en una posición complementaria específica del sitio, en condiciones tales que la primera cadena y la segunda cadena forman un oligonucleótido bicatenario dúplex; después, tras contactar el oligonucleótido bicatenario dúplex que incorpora la nucleobase no natural que comprende la fracción enlazadora reactiva con un reactivo de carga que comprende una carga y un grupo de reactividad complementaria, en condiciones adecuadas para la reacción del enlazador reactivo y el grupo de reactividad complementaria para producir un enlazador acoplado; para proporcionar el oligonucleótido bicatenario dúplex funcionalizado con la carga unida al mismo a través de un enlazador acoplado.

40 En una realización, el enlazador se une con una carga después de que se incorpora el trifosfato 5' correspondiente en un oligonucleótido de ADN o ARN mediante el uso de una polimerasa de ADN o ARN (después de la desprotección, si es necesario). En otra realización, se sintetiza un segundo oligonucleótido que es complementario a la primera cadena, la segunda cadena contiene un nucleótido no natural en una posición complementaria al nucleótido no natural de la primera cadena; después, la reacción del oligonucleótido bicatenario resultante con un reactivo portador de carga que reacciona selectivamente con el centro reactivo del enlazador reactivo proporciona un oligonucleótido bicatenario funcionalizado (por ejemplo, ADN/ADN, ADN/ARN o ARN/ARN) que porta una carga.

45 Además, en la presente memoria se divulgan, en diversas realizaciones, estructuras que comprenden la fórmula: $N1 - Zx - N2$, en la que $N1$ es un nucleótido o análogo del mismo, o un grupo fosfato terminal; en la que $N2$ es un nucleótido o análogo de este, o un grupo hidroxilo terminal; en la que Z es un compuesto que tiene cualquiera de las fórmulas $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$, $\beta 8a$, $\beta 8b$, $\beta 9a$, $\beta 9b$, $\alpha 15a$, $\alpha 15b$, $\alpha 16a$ o $\alpha 16b$; y en las que x es un número entero de 1 a 20. En algunas realizaciones, la estructura es un oligonucleótido. En algunas realizaciones, el oligonucleótido es un ácido ribonucleico o un ácido desoxirribonucleico. En algunas realizaciones, el oligonucleótido es un aptámero o un sensor basado en ácido nucleico. En algunas realizaciones, el oligonucleótido es una baliza molecular. En algunas realizaciones, el oligonucleótido es un análogo de ARN o análogo de ADN.

55 Además, en la presente memoria se divulga, en diversas realizaciones, un procedimiento para identificar un aptámero de ácido nucleico que comprende al menos un compuesto divulgado en la presente memoria (por ejemplo, $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$, $\beta 8a$, $\beta 8b$, $\beta 9a$, $\beta 9b$, $\alpha 15a$, $\alpha 15b$, $\alpha 16a$, $\alpha 16b$), que tiene una propiedad conveniente mejorada con respecto a una molécula objetivo, el procedimiento comprende: a) preparar una mezcla candidata de aptámeros de ácido nucleico monocatenario, en la que cada aptámero de ácido nucleico de la mezcla candidata de aptámeros comprende al menos un compuesto divulgado en la presente memoria (por ejemplo, $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$, $\beta 8a$, $\beta 8b$, $\beta 9a$, $\beta 9b$, $\beta 15a$, $\alpha 15b$, $\alpha 16a$, $\alpha 16b$); después, b) poner en contacto la mezcla candidata con la molécula objetivo en condiciones adecuadas para que se produzca la unión a la molécula objetivo; entonces, c) dividir uno o más aptámeros de ácido nucleico que tienen la propiedad conveniente con respecto a la molécula objetivo entre los aptámeros de la mezcla candidata; y después, d) amplificar uno o más aptámeros de ácido nucleico con la propiedad conveniente, in vitro, para producir uno o más aptámeros de ácido nucleico, que tienen una propiedad conveniente

mejorada con respecto a la molécula objetivo. En algunas realizaciones, el procedimiento comprende además el paso e) repetir los pasos c) y d). En algunas realizaciones, los aptámeros de ácidos nucleicos monocatenarios se seleccionan del grupo que consiste en ADN monocatenario y ARN monocatenario. En algunas realizaciones, la propiedad conveniente es una afinidad de unión para un objetivo. En algunas realizaciones, la propiedad conveniente es una actividad inducida por unión al objetivo. En algunas realizaciones, la propiedad conveniente es una actividad catalítica. En algunas realizaciones, la propiedad conveniente es una actividad de inhibición, una actividad de activación o una modificación de una actividad de inhibición o una actividad de activación. En algunas realizaciones, la propiedad conveniente es una actividad de cambio de estructura o una modificación de una actividad de cambio de estructura. En algunas realizaciones, la propiedad conveniente es una actividad cooperativa. En algunas realizaciones, la actividad conveniente es una eficacia mejorada de actualización celular.

Además en la presente memoria se divulga, en ciertas realizaciones, un aptámero que comprende un compuesto que tiene cualquiera de las siguientes fórmulas: $\alpha 14a$, $\alpha 14b$, $\alpha 14c$, $\alpha 14d$, $\alpha 14e$, $\alpha 14f$, $\beta 8a$, $\beta 8b$, $\beta 9a$, $\beta 9b$, $\alpha 15a$, $\alpha 15b$, $\alpha 16a$, $\alpha 16b$.

15

Breve descripción de las figuras

La **Figura 1** ilustra el emparejamiento de dTPT3-dNaM, d5SICS-dNaM, d5SICS-dMMO2, d5SICS-dDMO en ADN o ARN.

La **Figura 2** ilustra los nucleótidos derivatizados con el enlazador dTPT3R, d5SICSR, dMMO2R, dMMO2pR, dDMOR, dNaMpR, dNaMpR, dFEMO y dEMO, donde R = 3-aminopropin-1-ilo (denotado como A, por ejemplo, dTPT3A); R = dicloroacetil-3-aminopropin-1-ilo (denotado como PA); R = 4-oxahepta-1,6-diin-1-ilo (denotado como CO); R = hepta-1,6-diin-1-ilo (denotado como CC).

La **Figura 3** muestra una vista general de la estrategia de marcaje específica del sitio de post-síntesis basado en fosforotioato.

La **Figura 4** muestra una vista general de la estrategia de marcaje específica del sitio de post-síntesis basado en amino. Los nucleótidos modificados con el enlazador también pueden incorporarse directamente en la plantilla de ADN mediante el uso de la síntesis estándar en fase sólida de oligonucleótidos y las fosfoamiditas correspondientes.

La **Figura 5** muestra una vista general de la estrategia de marcaje específica del sitio de post-síntesis basado en química de clics. Los nucleótidos modificados con el enlazador también pueden incorporarse directamente en la plantilla de ADN mediante el uso de la síntesis estándar en fase sólida de oligonucleótidos y las fosfoamiditas correspondientes.

La **Figura 6** muestra datos representativos que ilustran el marcaje post-amplificación del ADN analizada mediante el cambio de gel de estreptavidina (SA). La banda de migración más rápida corresponde al dsADN, mientras que la banda de migración más lenta corresponde al complejo 1: 1 entre dsADN y estreptavidina. (A) La eficiencia de marcaje es del 72 % con d5SICSPA-dNaM y del 80 % con dTPT3PA-dNaM. (B) La eficiencia de marcaje es del 6 %, 84 % y 56 % con d5SICSPA-dNaM en la primera posición (cebador 1: par de bases no natural en la posición 1), novena (cebador 2: par de bases no natural en la posición once) y undécima (cebador 3: par de bases no natural en la posición nueve). Las eficiencias de marcaje correspondientes con dTPT3PA-dNaM son 72 %, 94 % y 81 %.

La **Figura 7** muestra datos de electroforesis en gel que confirman la transcripción de longitud completa de ARN que contiene análogos derivatizados con el enlazador de 5SICS o MMO2.

La **Figura 8** ilustra 12 agrupaciones de análogos de nucleobase α , $\alpha 1$ - $\alpha 12$.

La **Figura 9** ilustra 6 agrupaciones de análogos de nucleobase β , $\beta 1$ - $\beta 6$.

La **Figura 10** ilustra 2 agrupaciones de análogos de nucleobase α , $\alpha 13$ y $\alpha 14$; en las que cada X es independientemente carbono o nitrógeno; en las que cada R1 es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en las que cada R2 es opcional y cuando está presente, es independientemente hidrógeno, alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, grupo azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en las que cada R es opcional y cuando está presente, es independientemente hidrógeno, alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, grupo azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno.

La **Figura 11** ilustra ejemplos de análogos de nucleobase $\alpha 14a$, que incluye análogos de nucleobase derivatizados con el enlazador, MMSpCO y MMSPA.

La **Figura 12** ilustra 2 agrupaciones de análogos de nucleobase β , $\beta 7$ y $\beta 8$; en las que cada X es independientemente carbono o nitrógeno; en las que cada R2 es opcional y cuando está presente es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno; y en las que R es opcional y cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinilo, metoxi, metanotiol, metanoseleno,

65

halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga.

La **Figura 13** ilustra ejemplos de análogos de nucleobase derivatizados con el enlazador $\beta 8$.

5 La **Figura 14** muestra porcentajes de pares de bases no naturales retenidos en el ADN después de la amplificación durante 6 rondas de cribado.

10 La **Figura 15** ilustra análogos de nucleobase α , $\alpha 15a$, $\alpha 15b$, $\alpha 16a$, y $\alpha 16b$; en los que cada X es independientemente carbono o nitrógeno; en los que cada R1 es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en los que cada R2 es opcional y cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinoilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en los que cada Y es independientemente azufre, oxígeno, selenio, o amina secundaria; en los que cada E es independientemente oxígeno, azufre o selenio.

15 La **Figura 16** ilustra análogos de nucleobase β , $\beta 9a$ y $\beta 9b$; y en los que cada X es independientemente carbono o nitrógeno; en los que cada R1 es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al cual se une una carga; en los que cada R2 es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenoilo, alquinoilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, nitro, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en los que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en los que cada E es independientemente oxígeno, azufre o selenio.

Descripción detallada

30 Las expresiones tales como "en condiciones adecuadas para proporcionar" o "en condiciones suficientes para producir" o similares, en el contexto de procedimientos de síntesis, como se usa en la presente memoria se refiere a condiciones de reacción, tales como tiempo, temperatura, disolvente, concentraciones de reactivo y similares, que están dentro de la habilidad ordinaria para que un experimentador varíe, que proporcionan una cantidad útil o producen un producto de reacción. No es necesario que el producto de reacción conveniente sea el único producto de reacción o que los materiales de partida se consuman por completo, siempre que el producto de reacción conveniente pueda aislarse o usarse de otro modo.

35 Por "químicamente factible" se entiende una disposición de unión o un compuesto donde no se violan las reglas generalmente entendidas de estructura orgánica; por ejemplo, una estructura dentro de una definición de una reivindicación que contendría en ciertas situaciones un átomo de carbono pentavalente que no existiría en la naturaleza se entendería que no está dentro de la reivindicación. Las estructuras divulgadas en la presente memoria, en todas sus realizaciones, incluyen solo estructuras "químicamente viables", y cualquier estructura mencionada que no sea químicamente factible, por ejemplo, en una estructura mostrada con átomos o grupos variables, no se divulga o reivindica en la presente memoria.

40 Un "análogo" de una estructura química, como se usa el término en la presente memoria, se refiere a una estructura química que conserva una similitud sustancial con la estructura parental, aunque puede no derivarse fácilmente sintéticamente de la estructura parental. En algunas realizaciones, un análogo de nucleótido es un nucleótido no natural. En algunas realizaciones, un análogo de nucleósido es un nucleósido no natural. Una estructura química relacionada que se obtiene fácilmente por síntesis a partir de una estructura química original se conoce como un "derivado."

45 Por consiguiente, un "análogo de ADN" o un "análogo de ARN", como se usan en la presente memoria, se refieren a polímeros similares a ADN o ARN tales como ácidos nucleicos peptídicos (PNA), ácidos nucleicos bloqueados (LNA), fosfortioatos y similares, que se conocen bien en la técnica. Los análogos de ADN y ARN, así como el ADN y el ARN, se pueden sintetizar en sintetizadores automáticos, por ejemplo, mediante el uso de química de fosfoamidita u otros enfoques químicos adaptados para el uso del sintetizador.

50 El ADN incluye, pero no se limita a, ADNc y ADN genómico. El ADN puede estar unido, por medios covalentes o no covalentes, a otra biomolécula, que incluye, pero no se limitan a, ARN y péptido. El ARN incluye ARN codificante, por ejemplo, ARN mensajero (ARNm). En algunas realizaciones, el ARN es ARNr, ARNi, ARNsno, ARNmicro, ARNsi, ARNsn, ARNex, ARNip, ARNc largo, o cualquier combinación o híbrido de los mismos. En algunos casos, el ARN es un componente de una ribozima. El ADN y el ARN pueden estar en cualquier forma, que incluye, pero no se limita a, lineal, circular, superenrollado, monocatenario y bicatenario.

60 La expresión "grupo protector de amino" o "protegido con amino" como se usa en la presente memoria se refiere a aquellos grupos destinados a proteger un grupo amino contra reacciones indeseables durante los procedimientos

sintéticos y que después se pueden eliminar para revelar la amina. Los grupos protectores de amino usados comúnmente se divulgan en *Protective Groups in Organic Synthesis*, Greene, T.W.; Wuts, P.G.M., John Wiley & Sons, Nueva York, NY, (3a edición, 1999). Los grupos protectores amino incluyen grupos acilo tales como formilo, acetilo, propionilo, pivaloilo, t-butilacetilo, 2-cloroacetilo, 2-bromoacetilo, trifluoroacetilo, tricloroacetilo, o-nitrofenoxiacetilo, α -clorobutirilo, benzoilo, 4-clorobenzoilo, 4-bromobenzoilo 4-nitrobenzoilo, y similares; grupos sulfonilo tales como bencenosulfonilo, p-toluenosulfonilo y similares; grupos alcóxido o ariloxycarbonilo (que forman uretanos con la amina protegida) tales como benciloxycarbonilo (Cbz), p-clorobenciloxycarbonilo, p-metoxibenciloxycarbonilo, p-nitrobenciloxycarbonilo, 2-nitrobenciloxycarbonilo, p-bromobenciloxycarbonilo, 3,4-dimetoxibenciloxycarbonilo, 3,5-dimetoxibenciloxycarbonilo, 2,4-dimetoxibenciloxycarbonilo, 4-metoxibenciloxycarbonilo, 2-nitro-4,5-dimetoxibenciloxycarbonilo, 3,4,5-trimetoxibenciloxycarbonilo, 1-(p-bifenilil)-1-metiletoxycarbonilo, α,α -dimetil-3,5-dimetoxibenciloxycarbonilo, benciliriloxycarbonilo, t-butiloxycarbonilo (Boc), diisopropilmetoxycarbonilo, isopropiloxycarbonilo, entoxycarbonilo, mentoxycarbonilo, aliloxycarbonilo (Alloc), 2,2,2-tricloroetoxycarbonilo, 2-trimetilsililetoxycarbonilo (Teoc), fenoxycarbonilo, 4-nitrofenoxycarbonilo, fluorenil-9-metoxycarbonilo (Fmoc), ciclopentiloxycarbonilo, adamantiloxycarbonilo, ciclohexiloxycarbonilo, feniltiocarbonilo y similares; grupos aralquilo tales como bencilo, trifenilmetilo, benciloximetilo y similares; y grupos sililo tales como trimetilsililo y similares. Los grupos protectores de amina también incluyen grupos protectores de aminocíclicos tales como ftaloilo y ditiosuccinimidilo, que incorporan el nitrógeno amino en un heterociclo. Típicamente, los grupos protectores de amino incluyen formilo, acetilo, benzoilo, pivaloilo, t-butilacetilo, fenilsulfonilo, Alloc, Teoc, bencilo, Fmoc, Boc y Cbz. Los grupos protectores también incluyen carbamato de metilo, carbamato de 9-fluorenilmetilo, carbamato de 2,2,2-tricloroetilo, carbamato de t-butilo, carbamato de 2-(trimetilsilil)etilo, carbamato de alilo, carbamato de bencilo, carbamato de m-nitrofenilo, trifluoroacetamida, bencilamina, allamina, y tritilamina. Los grupos protectores también incluyen formamidas, acetamidas, trifluoroacetamidas, p-toluenosulfonilo, trifluorometanosulfonilo, trimetilsililetanosulfonamida y t-terc-butilfulfonilo. El experto en la técnica ordinario selecciona y usa el grupo protector de amino apropiado para la tarea sintética en cuestión.

Los análogos de ADN y ARN incluyen análogos de PNA (ácido nucleico peptídico) y LNA (ácido nucleico bloqueado).

Un ácido nucleico peptídico (PNA) es un análogo sintético de ADN/ARN en el que una cadena principal similar a un péptido reemplaza al esqueleto de azúcar-fosfato de ADN o ARN. Los oligómeros de PNA muestran una fuerza de unión mayor y una especificidad mayor en la unión a los ADN complementarios, con un desajuste de base de PNA/ADN que es más desestabilizador que un desajuste similar en un dúplex de ADN/ADN. Esta fuerza y especificidad de unión también se aplica a los dúplex de PNA/ARN. Los PNA no se reconocen fácilmente ni por las nucleasas ni por las proteasas, lo que los hace resistentes a la degradación enzimática. Los PNA también son estables en un rango de pH amplio. Véase también Nielsen PE, Egholm M, Berg RH, Buchardt O (diciembre de 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; y, Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Nordén B y Nielsen PE (1993), "PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen Bonding Rules". *Nature* 365 (6446): 566-8. doi: 10.1038/365566a0. PMID 7692304

Un ácido nucleico bloqueado (LNA) es un nucleótido de ARN modificado, en el que la fracción ribosa de un nucleótido de LNA se modifica con un puente adicional que conecta el oxígeno 2' y el carbono 4'. El puente "bloquea" la ribosa en el 3'-endo conformación (norte), que a menudo se encuentra en los dúplex en forma de A. Los nucleótidos de LNA pueden mezclarse con residuos de ADN o ARN en el oligonucleótido cuando se desea. Tales oligómeros pueden sintetizarse químicamente y están disponibles comercialmente. La conformación de ribosa bloqueada mejora el apilamiento de bases y la organización previa de la cadena principal. Véase, por ejemplo, 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; y, Satoshi Obika; Daishu Nanbu, Yoshiyuki Hari, Ken-ichiro Morio, Yasuko In, Toshimasa Ishida, Takeshi Imanishi (1997), "Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3'-endo sugar puckering", *Tetrahedron Lett.* 38 (50): 8735-8. doi:10.1016/S0040-4039(97)10322-7.

Una sonda de baliza molecular o baliza molecular es una sonda de hibridación de oligonucleótidos que puede detectar la presencia de una secuencia específica de ácido nucleico en una solución homogénea. Las balizas moleculares son moléculas en forma de horquilla con un fluoróforo inactivado internamente cuya fluorescencia se restaura cuando se unen a una secuencia de ácido nucleico objetivo. Véase, por ejemplo, Tyagi S, Kramer FR (1996), "Molecular beacons: probes that fluoresce upon hybridization", *Nat Biotechnol.* 14 (3): 303-8. PMID 9630890; Täpp I, Malmberg L, Rennel E, Wik M, Syvänen AC (2000 Apr), "Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan assay and Molecular Beacon probes", *Biotechniques* 28 (4): 732-8. PMID 10769752; y, Akimitsu Okamoto (2011), "ECHO probes: a concept of fluorescence control for practical nucleic acid sensing", *Chem. Soc. Rev.* 40: 5815-5828.

En algunas realizaciones, una nucleobase es generalmente la porción de base heterocíclica de un nucleósido. Las nucleobases pueden ser de origen natural, pueden modificarse, pueden no tener similitud con las bases naturales y pueden sintetizarse, por ejemplo, mediante síntesis orgánica. En ciertas realizaciones, una nucleobase comprende cualquier átomo o grupo de átomos capaces de interactuar con una base de otro ácido nucleico con o sin el uso de enlaces de hidrógeno. En ciertas realizaciones, una nucleobase no natural no se deriva de una nucleobase natural. Se debe notar que las nucleobases no naturales no poseen necesariamente propiedades básicas, sin embargo, se denominan nucleobases por simplicidad. En algunas realizaciones, cuando se refiere a una nucleobase, una "(d)" indica que la nucleobase puede unirse a una desoxirribosa o una ribosa.

En algunas realizaciones, un nucleósido es un compuesto que comprende una fracción nucleobase y una fracción azúcar. Los nucleósidos incluyen, pero no se limitan a, nucleósidos de origen natural (como se encuentra en el ADN y el ARN), nucleósidos abásicos, nucleósidos modificados, y nucleósidos que tienen bases miméticas y/o grupos de azúcar. Los nucleósidos incluyen nucleósidos que comprenden cualquier variedad de sustituyentes. Un nucleósido puede ser un compuesto glucósido formado a través de la unión glucosídica entre una base de ácido nucleico y un grupo reductor de un azúcar.

En algunas realizaciones, un nucleótido es un compuesto en el cual la fracción de azúcar de un nucleósido forma un éster con ácido fosfórico, más preferentemente un éster mono, di o trifosfato. La fracción de azúcar de dicho nucleósido o nucleótido puede ser ribofuranosilo, 2'-desoxirribofuranosilo o ribofuranosilo 2'-sustituido que tiene un sustituyente en la posición 2'. Del mismo modo, la fracción de ácido fosfórico puede ser ácido tiofosfórico. A saber, las fracciones de azúcar y ácido fosfórico pueden estar en la misma forma encontrada en nucleósidos, nucleótidos o derivados conocidos de los mismos. Un ribonucleótido cuya fracción de azúcar es ribofuranosilo puede usarse como un miembro constituyente del ARN. Un desoxirribonucleótido cuya fracción de azúcar es desoxirribofuranosilo puede usarse como un miembro constituyente del ADN. Un nucleótido puede ser un nucleósido que comprende además un grupo de enlace fosfato. Los nucleótidos pueden incluir nucleósidos que contienen una fracción fosfato.

Se desarrolló y mostró por nosotros una clase de pares de bases no naturales, ejemplificada por d5SICS-dNaM y d5SICS-dMMO2 (Figura 1), para ser replicada (incluso a través de PCR) y transcrita por una amplia gama de polimerasas naturales con eficiencias y fidelidades cercanas a las de pares de bases naturales (Véase 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), 2009, 5046-5047; Lavergne T.; Degardin M.; Malyshev D.A.; Quach H.T.; Dhami K.; Ordoukhanian P.; Romesberg, F.E.; Expanding the scope of replicable unnatural ADN: Stepwise optimization of a predominantly hydrophobic base pair. *J. Am. Chem. Soc.* 2013, 135, 5408-5419; Seo, Y.J., Malyshev D.A., Lavergne T., Ordoukhanian P. y Romesberg, F.E., *J. Am. Chem. Soc.* 2011, 133, 19878.). Estos pares de bases no naturales se forman entre análogos de nucleótidos que tienen nucleobases no naturales, predominantemente hidrófobas. Los pares de bases se muestran en la Figura 1, con cada nucleótido no natural que se incorpora en oligonucleótidos en posiciones complementarias (es decir, apareadas); se entiende que las nucleobases se unen en la posición que indica la línea ondulada, a la posición 1' de una fracción ribosilo o 2'-desoxirribosilo, que se incorpora en el ARN o el ADN mediante grupos fosfato o fosforotioato que unen los grupos hidroxilo 3' y 5' de los grupos ribosilo o desoxirribosilo, respectivamente, como en el caso de los ácidos nucleicos completamente naturales. El emparejamiento de bases tiene lugar como parte de una estructura complementaria de bases apareadas, como se conoce bien en la formación de estructuras dúplex de oligonucleótidos. En la presente memoria se divulga, en diversas realizaciones, una nucleobase no natural dTPT3 y sus variantes derivatizadas con el enlazador, que se cree que tiene bases apareadas con las nucleobases no naturales dNaM, dMMO2, y dDMO (o sus variantes derivatizadas con el enlazador) de manera similar.

Se ha demostrado que los nucleótidos no naturales dTPT3 y dTPT3^{PA} se incorporan eficientemente en el ADN por las ADN polimerasas opuestas a dNaM (Figura 2). Tanto dTPT3 como dTPT3^{PA} (PA = dicloroacetil-3-aminopropin-1-ilo) así como otras variantes derivatizadas con el enlazador de dTPT3 (Figura 3) que incluyen, cuando R = 3-aminopropin-1-ilo (dTPT3^A), R = 4-oxahepat-1,6-diyn-1-ilo (dTPT3^{CO}), o R = hepta-1,6-diyn-1-ilo (dTPT3^{CC})) también se espera que se apareen con dNaM, dDMO o dMMO2, o análogos derivatizados con enlazadores de los mismos. Las tasas de incorporación de dTPT3 y las variantes del mismo derivatizados con el enlazador, se acercan a las de un par de bases naturales opuestas a dNaM. Los pares de bases no naturales adicionales identificados con tasas de incorporación eficientes incluyen dTPT3-dFEMO, dTPT3-dFIMO, dTPT3-dIMO, dFTPT3-dNaM, dFTPT3-dFEMO, dFTPT3-dFIMO y dFTPT3-dIMO.

Además, en la presente memoria se divulgan, en diversas realizaciones, nucleótidos no naturales con análogos de nucleobase que incluyen, análogos α (por ejemplo, cualquiera de las Figuras 8, 10, 11, 15 y derivados de los mismos); análogos β (por ejemplo, cualquiera de las Figuras 9, 12, 13, 16 y derivados de los mismos); d5SICS^{CO}, d5SICS^{CC}, dDMO^{CO}, dDMO^{CC}, dMMO2^{pCO}, dMMO2^{pCC}, dTPT3, dTPT3^{PA}, dTPT3^A, dTPT3^{CO}, dTPT3^{CC}, y las formas ribosilo de estos, y los análogos de los mismos (véase la Figura 2); en forma de nucleósidos, nucleósidos 5' trifosfatos, y análogos de los mismos (por ejemplo, ribosilo y 2'-desoxirribosilo), nucleótidos y análogos de los mismos (por ejemplo, ribosilo y 2'-desoxirribosilo, fosfato y fosforotioato), que incluye los reactivos de nucleótidos derivados de ellos para su uso en la síntesis de ARN/ADN (fosforamiditas protegidas con DMT) y para su uso en la incorporación enzimática de oligonucleótidos como por PCR o transcripción mediada por la ARN polimerasa de T7, e incorporada en ácidos nucleicos (oligonucleótidos) como el ADN y el ARN. Los compuestos que comprenden los análogos de nucleobase no

naturales también pueden incorporarse en análogos de ADN o en análogos de ARN, tales como PNA, LNA, y otros polímeros análogos de polinucleótidos similares. Los análogos de nucleobase ejemplares divulgados en la presente memoria incluyen análogos $\beta 8$ que comprenden las fórmulas $\beta 8a$ y $\beta 8b$, como se muestra en la Figura 12, en las que cada X es independientemente carbono o nitrógeno; en las que cada R_2 es opcional y cuando está presente es independientemente hidrógeno, un grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno; y en las que R es opcional y cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo, adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga. Los ejemplos de análogos $\beta 8$ incluyen dTPT3 y análogos del mismo derivatizados con el enlazador. Los análogos de nucleobase ejemplares divulgados en la presente memoria incluyen análogos de $\alpha 14$ que comprenden las fórmulas $\alpha 14a$ - $\alpha 14f$, como se muestra en la Figura 10, en las que cada X es independientemente carbono o nitrógeno; en las que cada R_1 es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en donde cada R_2 es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga; en las que cada R es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio, o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno. Los ejemplos de análogos $\alpha 14$ incluyen dMMS, dDMS, dFEMS, dBrMS, dIMS y análogos de los mismos derivatizados con el enlazador.

Además, en la presente memoria se divulgan, en diversas realizaciones, pares de bases no naturales que comprenden cualquier análogo α o derivado del mismo divulgado en la presente memoria, y cualquier análogo β o derivado del mismo divulgado en la presente memoria. Los derivados incluyen, pero no se limitan a, sustituciones de átomos y adiciones de fracciones enlazadoras. Las fracciones enlazadoras pueden unirse a los análogos durante la síntesis o después de la incorporación de la nucleobase en un ácido nucleico. Los pares de bases no naturales ejemplares incluyen, pero no se limitan a, dTPT3-dNaM, dTPT3-dFEMO, dTPT3-dFIMO, dTPT3-dIMO, dFTPT3-dNaM, dFTPT3-dFEMO, dFTPT3-dFIMO, dFTPT3-dIMO. Los pares de bases no naturales incluyen, pero no se limitan a, 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, y dFTPT3-dIMS; en los que una o dos nucleobases no naturales del par de bases no natural pueden derivatizarse con un enlazador. Los ejemplos de pares de bases no naturales de esta divulgación incluyen además cualquier par descrito en el Ejemplo 1. Los análogos β ejemplares incluyen los que se presentan en las Figuras. 9, 12 y 13. Los análogos de nucleobase β ejemplares incluyen análogos $\beta 8$ que comprenden las fórmulas $\beta 8a$ y $\beta 8b$, como se muestra en la Figura 12, en las que cada X es independientemente carbono o nitrógeno; en las que cada R_2 es opcional y cuando está presente es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio, o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno; y en las que R es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga. Los ejemplos de análogos β incluyen dTPT3, d5SICS, dFTPT3 y derivados o análogos de los mismos. Los análogos α ejemplares incluyen los que se presentan en las Figs. 8, 10 y 11. Los análogos α ejemplares incluyen análogos $\alpha 14$ que comprenden las fórmulas $\alpha 14a$ - $\alpha 14f$, como se muestra en la Figura 10, en las que cada X es independientemente carbono o nitrógeno; en las que cada R_1 es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga; en las que cada R_2 es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga; en las que cada R es opcional y, cuando está presente, es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoseleno, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, enlazador acoplado al que se une una carga; en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria; en las que cada E es independientemente azufre, selenio u oxígeno. Los ejemplos de análogos α incluyen dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO y derivados o análogos de los mismos. En algunas realizaciones, un par de bases no natural incluye un análogo α y una base natural. En algunas realizaciones, un par de bases no natural incluye un análogo β y una base natural. Además, en la presente memoria se divulgan, en algunos aspectos, pares de bases no naturales que comprenden los mismos dos análogos de nucleósidos no naturales o derivados de los mismos.

Un par de bases no natural, en varios aspectos, comprende una nucleobase no natural divulgada en la presente memoria (por ejemplo, un análogo α o derivado del mismo, un análogo β o derivado del mismo) y otra nucleobase no natural que incluye, pero no se limita a, 2-aminoadenina-9-ilo, 2- aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, 2-propilo y derivados de alquilo de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-deazaguanina, 3-deazaadenina, 4-tio-uracilo, 4-tio-timina, uracil-5-ilo, hipoxantina-9-ilo (I), 5-metil-citosina, 5-hidroxiometil citosina, xantina, hipoxantina, 5-bromo y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5-sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-bromocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxiourea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clouracilo, 5-fluouracilo, y 5-yodouracilo, 6-alquil derivados de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azo citosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-deazaguanina, 7-deazaguanosina, 7-deaza-adenina, 7-deaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y adeninas y guaninas 8-hidroxilo sustituidas; N4-etilcitosina, purinas N-2 sustituidas, purinas N-6 sustituidas, purinas O-6 sustituidas, aquellas que aumentan la estabilidad de la formación de dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos expandidos por tamaño, ácidos nucleicos fluorados, pirimidinas tricíclicas, fenoxazina citidina ([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazina-2(3H)-ona), abrazaderas G, fenoxazina citidina (9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4] benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido[3', 2':4,5]pirrolo [2,3-d] pirimidin-2-ona), 5-fluouracilo, 5-bromouracilo, 5-clouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5- (carboxihidroxi)metil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metil-tio-N6-isopenteniladenina, ácido oxiacético 5-uracilo, wibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido oxiacético 5-uracilo metiléster, ácido oxiacético 5-uracilo, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza con un heterociclo. Los análogos α del par de bases no natural incluyen, sin limitación, dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO y derivados o análogos de los mismos. Los análogos β del par de bases no natural incluyen, sin limitación, dTPT3, d5SICS y dFTPT3.

En algunas realizaciones, las nucleobases no naturales y los pares de bases no naturales divulgados en esta memoria tienen incorporación y extensión eficientes con polimerasas naturales. En algunas realizaciones, las nucleobases no naturales y los pares de bases no naturales divulgados en la presente memoria tienen incorporación y extensión eficientes con polimerasas modificadas. Se evalúa el efecto de una nucleobase no natural o un derivado de nucleobase no natural sobre el reconocimiento de la polimerasa, en realizaciones ejemplares, mediante la determinación de la eficiencia en estado estacionario (por ejemplo, constante de velocidad de segundo orden k_{cat}/K_M) con el que la polimerasa sintetiza un par de bases no natural, mediante la inserción del nucleótido no natural opuesto a su base complementaria en una plantilla, y extiende el terminal cebador no natural resultante, mediante la inserción del siguiente nucleótido natural correcto. Las tasas de síntesis y extensión correspondientes para emparejamientos erróneos con nucleótidos naturales también pueden medirse para determinar la fidelidad. En algunas realizaciones, las polimerasas no necesitan modificarse para mejorar las tasas de incorporación o extensión. Las realizaciones y ejemplos divulgados en esta memoria pueden realizarse con cualquier polimerasa conocida. Las polimerasas incluyen polimerasas de origen natural y cualquier variación modificada de las mismas, que incluye, pero no se limita a, mutantes, recombinantes, fusiones, modificaciones genéticas, modificaciones químicas, sintéticos y análogos. Las polimerasas naturales y las variaciones modificadas de las mismas no se limitan a las polimerasas que retienen la capacidad de catalizar una reacción de polimerización. En algunos casos, las variaciones naturales y/o modificadas de las mismas conservan la capacidad de catalizar una reacción de polimerización. Las polimerasas mutantes incluyen polimerasas en las que uno o más aminoácidos se reemplazan con otros aminoácidos (naturales o no naturales) y polimerasas que tienen una o más inserciones o deleciones de aminoácidos. En algunas realizaciones una polimerasa se refiere a proteínas de fusión que comprenden al menos dos porciones unidas entre sí, por ejemplo, donde una porción comprende un péptido que puede catalizar la polimerización de nucleótidos en una cadena de ácido nucleico se une a otra porción que comprende una segunda fracción, tal como, una enzima indicadora o un dominio modificador de la procesividad. Una realización ejemplar de tal polimerasa es la ADN polimerasa T7, que comprende un dominio de polimerización de ácido nucleico y un dominio de unión a tiorredoxina, en los que la unión a tiorredoxina mejora la procesividad de la polimerasa. En ausencia de la unión a la tiorredoxina, la ADN polimerasa T7 es una polimerasa distributiva con procesividad de solo una a algunas bases. Las ADN polimerasas incluyen, pero no se limitan a, ADN polimerasas bacterianas, ADN polimerasas eucariotas, ADN polimerasas arqueales, ADN polimerasas virales y ADN polimerasas fágicas. Las ADN polimerasas bacterianas incluyen las ADN polimerasas I, II y III, IV y V de *E. coli*, el fragmento Klenow de la ADN polimerasa de *E. coli*, la ADN polimerasa de *Clostridium stercorarium* (Cst), la ADN polimerasa de *Clostridium thermocellum* (Cth) y la ADN polimerasa de *Sulfolobus solfataricus* (Sso). Las ADN polimerasas eucariotas incluyen las ADN polimerasas α , β , γ , δ , ϵ , η , ζ , σ , λ , μ , ν y κ , así como la polimerasa Rev1 (desoxicitidil transferasa terminal) y la desoxinucleotidil transferasa terminal (TdT). Las ADN polimerasas virales incluyen la ADN polimerasa T4, ADN polimerasa phi-29, GA-1, las ADN polimerasas similares a phi-29, ADN polimerasa PZA, ADN polimerasa phi-15, ADN polimerasa Cp1, ADN polimerasa Cp7, ADN polimerasa T7 y polimerasa T4. Las ADN polimerasas arqueales incluyen las ADN polimerasas termoestables y/o termofílicas tales como ADN polimerasas aisladas de la ADN polimerasa de *Thermus aquaticus* (Tag), ADN polimerasa de *Thermus filiformis* (Tfi), ADN polimerasa de *Thermococcus zilligii* (Tzi), ADN polimerasa de *Thermus thermophilus* (Tth), ADN polimerasa de *Thermus flavus* (Tfl), ADN polimerasa de *Pyrococcus woesei* (Pwo), ADN polimerasa de *Pyrococcus furiosus* (Pfu) y ADN polimerasa de Turbo Pfu, ADN polimerasa de *Thermococcus litoralis* (Tli), *Pyrococcus* sp. La polimerasa GB-D, ADN polimerasa de *Thermotoga maritima* (Tma), ADN polimerasa de *Bacillus stearothermophilus* (Bst), ADN polimerasa de *Pyrococcus kodakaraensis* (KOD), ADN polimerasa de Pfx, *Thermococcus* sp. ADN polimerasa JDF-3 (JDF-3), ADN polimerasa de *Thermococcus gorgonarius* (Tgo), ADN polimerasa de *Thermococcus acidophilium*; ADN polimerasa de *Sulfolobus acidocaldarius*; *Thermococcus* sp. 9 ° ADN polimerasa N-7; ADN polimerasa de *Pyrodictium occultum*; ADN polimerasa de *Methanococcus voltae*; ADN polimerasa de *Methanococcus thermoautotrophicum*; ADN polimerasa de *Methanococcus jannaschii*; ADN polimerasa de *Desulfurococcus cepa* TOK (D. Tok Pol); ADN polimerasa de *Pyrococcus abyssi*; ADN polimerasa de *Pyrococcus horikoshii*; ADN polimerasa de *Pyrococcus islandicum*; ADN polimerasa de *Thermococcus fumicolans*; ADN polimerasa de *Aeropyrum pernix*; y la ADN polimerasa heterodimérica DP1/DP2. Las ARN polimerasas incluyen, pero no se limitan a, ARN polimerasas virales tales como ARN polimerasa T7, polimerasa T3, polimerasa SP6, y polimerasa K11; ARN polimerasas eucariotas tales como ARN polimerasa I, ARN polimerasa II, ARN polimerasa III, ARN polimerasa IV y ARN polimerasa V; y ARN polimerasa de Archaea.

En algunas realizaciones, una polimerasa tiene una especificidad por un nucleótido no natural que comprende un análogo de nucleobase α o β que es al menos aproximadamente 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa hacia un nucleótido natural. En algunas realizaciones, una polimerasa tiene una especificidad para un nucleótido no natural que comprende un análogo de nucleobase α o β y un azúcar modificado que es al menos aproximadamente 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa hacia un nucleótido natural y/o el nucleótido no natural sin el azúcar modificado. En algunas realizaciones, una polimerasa tiene una especificidad para un nucleótido no natural que comprende un análogo de nucleobase α o β derivatizado con el enlazador que es al menos aproximadamente 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 97 %, 98 %, 99 %, 99,5 %, 99,99 % de la especificidad de la polimerasa hacia un nucleótido natural y/o el nucleótido no natural sin el enlazador. En algunas realizaciones, la nucleobase no natural es dTPT3. En algunas realizaciones, la nucleobase no natural es dMMS. En algunas realizaciones, la nucleobase no natural es dDMS. En algunas realizaciones, la nucleobase no natural es dBrMS. En algunas realizaciones, la nucleobase no natural es IMS. En algunas realizaciones, la nucleobase no natural es dFEMS. En algunas realizaciones, la nucleobase no natural es MMS^{pCO}. En algunas realizaciones, la nucleobase no natural es dMMS^{PA}. En algunas realizaciones, la nucleobase no natural es dFTPT3. En

5 algunas realizaciones, la nucleobase no natural es dTPT^{PA}. En algunas realizaciones, la nucleobase no natural es dTPT3^{CO}. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14a$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14b$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14c$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14d$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14e$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\alpha 14f$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\beta 8a$ o un derivado o análogo de la misma. En algunas realizaciones, la nucleobase no natural comprende la fórmula $\beta 8b$ o un derivado o análogo de la misma.

15 Las polimerasas pueden caracterizarse de acuerdo con su fidelidad cuando se usan con un nucleótido natural y/o no natural particular o colecciones de nucleótidos naturales y/o no naturales, en los que el nucleótido no natural comprende un análogo de nucleobase α o β divulgado en la presente memoria. En diversas realizaciones, la fidelidad generalmente se refiere a la precisión con la que una polimerasa incorpora nucleótidos correctos en un oligonucleótido en crecimiento cuando se hace una copia de una plantilla de oligonucleótidos. La fidelidad de la polimerasa puede medirse como la relación de incorporaciones correctas e incorrectas de nucleótidos naturales y no naturales cuando están presentes los nucleótidos naturales y no naturales, por ejemplo, en concentraciones iguales, para competir por la síntesis de la cadena en el mismo sitio el complejo binario de la plantilla de ácido nucleico-cadena de polimerasa. La fidelidad de la ADN polimerasa puede calcularse como la relación de (k_{cat}/K_M) para el nucleótido natural y no natural y (k_{cat}/K_M) para el nucleótido natural y antinatural incorrecto; donde k_{cat} y K_M son parámetros de Michaelis-Menten en cinética enzimática en estado estacionario. En algunas realizaciones, una polimerasa tiene un valor de fidelidad de al menos aproximadamente 100, 1.000, 10.000, 100.000 o 1×10^6 , con o sin actividad de corrección de lectura. En algunas realizaciones, una polimerasa tiene un valor de fidelidad de al menos aproximadamente 100, 1.000, 10.000, 100.000 o 1×10^6 para la incorporación de nucleótidos no naturales. En algunas realizaciones, el nucleótido no natural es dTPT3TP o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido es dNAM o un derivado del mismo. En algunas realizaciones, el nucleótido no natural es dNaMTP o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido es dTPT3 o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14a$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14b$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14c$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14d$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14e$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8a$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14f$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8b$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14a$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8b$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14b$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8b$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14c$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8b$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14d$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8e$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14e$ o un derivado del mismo. En algunas realizaciones, el nucleótido no natural comprende $\beta 8f$ o un derivado del mismo, y su nucleobase correspondiente en la plantilla de oligonucleótido comprende $\alpha 14f$ o un derivado del mismo.

50 Los pares de bases no naturales ejemplificadas en la presente memoria, en algunas realizaciones, se sintetizan/amplifican con eficiencia y fidelidad parecidas a pares de bases naturales. Los pares de bases no naturales comprenden, en diversas realizaciones, cualquier análogo de nucleobase α o derivado del mismo, y/o cualquier análogo de nucleobase β o derivado del mismo. Los ejemplos de análogos β incluyen dTPT3, d5SICS, dFTPT3 y derivados o análogos de los mismos. Los ejemplos de análogos α incluyen dMMS, dDMS, dBrMS, dIMS, dFEMS, dNAM, dMMO2, dDMO, dEMO, dFEMO y derivados o análogos de los mismos. En algunas realizaciones, un par de bases no natural se amplifica eficientemente en una variedad de contextos de secuencia diferentes, que incluyen secuencias ricas en GC y AT, secuencias aleatorias y secuencias que comprenden múltiples pares de bases nucleares no naturales, con más del 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 o mayor fidelidad por duplicación. Por ejemplo, un par de nucleobase no natural que comprende una o más nucleobases no naturales tiene una eficiencia de síntesis y/o fidelidad que es al menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % similar a una eficiencia de amplificación y/o fidelidad de un par de bases naturales. Como otro ejemplo, un par de nucleobases no naturales que comprende una o más nucleobases no naturales tiene una eficiencia de síntesis y/o fidelidad que es como máximo 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 % o 0,5 % menos eficiente y/o preciso que el de un par de bases natural. En algunas realizaciones, se transcribe un par de nucleobase no natural con

buena eficiencia y selectividad en ambos contextos de cadena (por ejemplo, dX tiene plantilla de inserción YTP y dY tiene plantilla de inserción XTP). En algunas realizaciones, en relación con la velocidad a la que se transcribe una secuencia completamente natural, la incorporación de un nucleótido no natural no reduce la velocidad de la transcripción de longitud completa. En algunas realizaciones, en relación con la velocidad a la que se transcribe una secuencia completamente natural, la incorporación de un nucleótido no natural reduce la velocidad de transcripción de longitud completa por un factor menor que 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35 o 40. En algunas realizaciones, el par de bases no natural comprende dTPT3 o un derivado o análogo del mismo, y dNaM o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende dTPT3 o un derivado o análogo del mismo, y dNaM o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende dTPT3 o un derivado o análogo del mismo, y dNaM o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende dTPT3 o un derivado o análogo del mismo, y dNaM o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14a o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14b o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14c o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14d o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14e o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8a o un derivado o análogo del mismo, y α 14f o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14a o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14b o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14c o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14d o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14e o un derivado o análogo del mismo. En algunas realizaciones, el par de bases no natural comprende β 8b o un derivado o análogo del mismo, y α 14f o un derivado o análogo del mismo.

Además en la presente memoria se divulgan, en diversas realizaciones, pares de bases no naturales que comprenden una o más nucleobases no naturales (por ejemplo, nucleobase α , nucleobase β o nucleobase α y nucleobase β), en las que una o dos nucleobases comprenden un enlazador. Un enlazador comprende un centro reactivo. Los ejemplos de centros reactivos incluyen, pero no se limitan a, alquilo, alqueno, alquino, fenilo, bencilo, halo, hidroxilo, carbonilo, aldehído, haloformilo, éster de carbonato, carboxilato, carboxilo, éster, metoxi, hidroperoxi, peroxi, éter, hemiacetal, hemicetal, acetal, cetil, ortoéster, metilendioxi, éster de ortocarbonato, carboxamida, amina primaria, amina secundaria, imida, azida, azo, cianato, isocianato, nitrato, nitrilo, isonitrilo, nitrosooxi, nitro, nitroso, piridil, sulfhidril, sulfuro, disulfuro, sulfinilo, sulfo, tiocianato, isotiocianato, carbonotioilo, fosfino, fosfeno, fosfato, borono, boronato, borino, borinato y una combinación de los mismos. Un ejemplo de una nucleobase derivatizada con el enlazador es TPT3^R mostrado en la Figura 2, en la que el superíndice R indica el enlazador. En algunas realizaciones, un enlazador se modifica con un grupo protector, por ejemplo, TPT3^{PA}, donde el enlazador es un enlazador de propargilo protegido.

En algunas realizaciones, un análogo de nucleobase divulgado en la presente memoria comprende un enlazador con función amino o un enlazador con función amino protegido (por ejemplo, dX^{PA}). En ciertas realizaciones, el enlazador con función amino es 3-aminopropin-1-ilo. En algunas realizaciones, un análogo de nucleobase divulgado en la presente memoria comprende un enlazador de éter alquino-azida para derivatización mediante química de clic o un enlazador de éter alquino-azida protegido para derivatización mediante química de clic. En ciertas realizaciones, el enlazador éter alquino-azida es 4-oxahepat-1,6-diin-1-ilo. En algunas realizaciones, un análogo de nucleobase divulgado en la presente memoria comprende un enlazador de trimetileno alquino-azida para derivatización mediante química de clic o un enlazador de trimetileno alquino-azida protegido para derivatización mediante química de clic. En ciertas realizaciones, el enlazador trimetileno alquino-azida es hepta-1,6-diin-1-ilo. En algunas realizaciones, X es un análogo de nucleósido β que tiene cualquiera de las fórmulas de las Figuras 9, 12, 13 y 16. En algunas realizaciones, X es ICS, PICS, MICS, 4MICS, 5MICS, NICS, ONICS, SICS, SNICS, 5SICS, 4SICS, 7OFF, 7OTP, TPT2, TPT3 o FTPT3. En algunas realizaciones, X es un análogo de nucleósido α que tiene cualquiera de las fórmulas de las Figuras 8, 10, 11 y 15. En algunas realizaciones, X es 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 o QMO.

En algunas realizaciones, un enlazador es un enlazador de propinilo, tal como los usados con nucleótidos naturales. Estos enlazadores comprenden propargilaminas, y la amina sirve como un sitio reactivo para unir otras funcionalidades.

En diversas realizaciones, una nucleobase derivatizada con el enlazador comprende un espaciador. Un espaciador ejemplar es la acetamidohexanamida. Un espaciador puede ser hidrófilo. Un espaciador puede conectar un enlazador a un grupo funcional. Los espaciadores incluyen, pero no se limitan a, Espaciador C3 (espaciador de 3 carbonos), Espaciador C6 (espaciador de 6 carbonos), espaciador fotosensible, espaciador de hexanodiol, Espaciador 9 (espaciador de trietilenglicol), Espaciador C12 (espaciador de 12 carbonos), Espaciador 18 (espaciador de hexaetilenglicol de 18 átomos) y Espaciador 1',2'-didesoxirribosa.

Un par de nucleobase no natural que comprende una o dos nucleobases derivatizados con el enlazador, en algunos casos, se amplifica con una eficiencia y fidelidad que es similar a la de un par de bases natural o un par de bases no natural derivatizadas sin enlazador. Por ejemplo, un par de nucleobase no natural que comprende una o dos nucleobases no naturales derivadas del enlazador tiene una eficiencia de síntesis y/o fidelidad que tiene al menos 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, o 99 % similar a una eficiencia de síntesis y/o fidelidad de un par de bases natural o un par de bases no natural derivatizadas sin enlazador. Como otro ejemplo, un par de nucleobases no naturales que comprende una o dos nucleobases no naturales derivatizados con el enlazador tiene una eficiencia de síntesis y/o fidelidad que es como máximo 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 % o 0,5 % menos eficiente y/o preciso que el de un par de bases natural o un par de bases no natural derivatizadas sin enlazador. En algunas realizaciones, un par de nucleobases no naturales comprende dTPT3^{PA}. En algunas realizaciones, un par de nucleobases no naturales comprende dTPT3^{CO}. En algunas realizaciones, un par de nucleobases no naturales comprende dMMS^{PCO}. En algunas realizaciones, un par de nucleobases no naturales comprende dMMS^{PA}. En algunas realizaciones, un par de nucleobases no naturales comprende dNaM^R. En algunas realizaciones, un par de nucleobases no naturales comprende dMMO2^R. En algunas realizaciones, un par de nucleobases no naturales comprende dDMO^R. En algunas realizaciones, un par de nucleobases no naturales comprende dSSICS^R. En algunas realizaciones, un par de nucleobases no naturales comprende dMMS^R. En algunas realizaciones, un par de nucleobases no naturales comprende dDMS^R. En algunas realizaciones, un par de nucleobases no naturales comprende dFEMS^R. En algunas realizaciones, un par de nucleobases no naturales comprende dBrMS^R. En algunas realizaciones, un par de nucleobases no naturales comprende dIMS^R.

En algunas realizaciones, una nucleobase no natural derivatizada con el enlazador tiene una eficacia de inserción aumentada durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador. En algunas realizaciones, una nucleobase no natural derivatizada con el enlazador tiene una eficiencia de inserción disminuida durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador. En algunos casos, una nucleobase no natural derivatizada con el enlazador tiene aproximadamente la misma eficacia de inserción durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador. En algunas realizaciones, una nucleobase no natural derivatizada con el enlazador protegido tiene una mayor eficiencia de inserción durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador protegido. En algunas realizaciones, una nucleobase no natural derivatizada con el enlazador protegido tiene una eficacia de inserción disminuida durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador protegido. En algunos casos, una nucleobase no natural derivatizada con el enlazador protegido tiene aproximadamente la misma eficacia durante la síntesis de oligonucleótidos, en comparación con la misma nucleobase no natural que no comprende un enlazador protegido.

Los procedimientos ejemplares para analizar la eficiencia de síntesis de pares de bases no naturales (inserción de una nucleobase no natural opuesta a su compañero en una plantilla) y la extensión (alargamiento continuo del cebador) se divulgan en la presente memoria. Una o ambas de las nucleobases en un par de bases no natural, en diversas realizaciones, puede ser una nucleobase no natural derivatizada con el enlazador. Un procedimiento utiliza un ensayo de estado estacionario. El ensayo se basa en determinar, bajo un conjunto fijo de condiciones, la cantidad de un cebador (por ejemplo, 23-mer) que se extiende mediante la adición del nucleótido no natural opuesto a su nucleótido complementario en una plantilla (por ejemplo, 45-mer) por una polimerasa (por ejemplo, el fragmento Klenow de *E. coli* ADN polimerasa I). En este ensayo, la eficiencia de la síntesis de pares de bases no naturales se caracteriza por medir el por ciento de incorporación (% inc) a una concentración dada del trifosfato no natural y el siguiente correcto, por ejemplo mediante el uso de una relación como $[24\text{-mer} + 25\text{-mer}]/[23\text{-mer} + 24\text{-mer} + 25\text{-mer}]$. En este ensayo, la eficiencia de la extensión se caracteriza por medir el porcentaje de extensión (% ext) a una concentración dada del nucleótido correcto próximo y las concentraciones saturantes de nucleótido no natural, por ejemplo mediante el uso de una relación $[25\text{-mer}]/[24\text{-mer} + 25\text{-mer}]$. Los resultados de un ensayo ejemplar de pre-estado estacionario se muestran en la Tabla 1, en la que el trifosfato no natural es 5SICS, FPT1, TPT1, TPT2, TPT3, FTPT3, TPT3^{PA} o 5SICS^{PA}. En algunas realizaciones, el por ciento de incorporación de una nucleobase antinatural es al menos 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 %, o 99 %. En algunas realizaciones, el por ciento de extensión de un nucleótido correcto próximo después de la inserción de una nucleobase no natural es al menos 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 % o más. En algunas realizaciones, la eficiencia de síntesis se incrementa mediante derivatización de la nucleobase no natural. Por ejemplo, mediante la adición de un enlazador, un enlazador protegido y/o un enlazador conjugado con una molécula de carga. Como otro ejemplo, la derivatización incluye sustituciones, adiciones o deleciones de átomos. En algunas realizaciones, el por ciento de extensión se incrementa mediante derivatización de la nucleobase no natural. La derivatización de una nucleobase no natural, en algunos casos, aumenta en al menos 1-2 órdenes de magnitud la eficiencia de inserción del nucleótido complementario al par de bases en la plantilla. Este aumento en la eficiencia puede deberse a un aumento k_{cat} y una disminución K_M .

65

Tabla 1. Cinética de estado pre-estacionario.

dXTP	% de incorporación ^a	% de extensión ^b
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 ^{PA}	68,3 ± 0,4	31,5 ± 0,7
5SICS ^{PA}	7,0 ± 0,2	5,5 ± 0,1

^aCondiciones de ensayo de incorporación: Trifosfato no natural 40 nM, dCTP 2 μM, 10s. ^bCondiciones de ensayo de extensión: Trifosfato no natural 10 μM, dCTP 2 μM, 10s. Los dXTP están pareados con dNaM.

Además, en esta memoria se divulgan los procedimientos de evaluación de la replicación. En un procedimiento, una plantilla de ácido nucleico dúplex que comprende un par de bases no natural (por ejemplo, dTPT3-dNaM o análogos de los mismos), se amplifica por PCR. En un ejemplo, un conjunto de reacciones de PCR emplea 48 ciclos con la polimerasa OneTaq. En otro ejemplo, un conjunto de reacciones de PCR emplea 20 ciclos de amplificación con Taq exonucleasa negativa. Se determina la eficiencia mediante el control del nivel de amplificación. La fidelidad, que generalmente se define como la extensión de pares de bases no naturales por duplicación, se determina a partir del porcentaje del ADN amplificado que retiene el par de bases no natural. El porcentaje de ADN amplificado que retiene el par de bases no natural puede determinarse a partir de las intensidades de pico relativas de un cromatograma de secuenciación. En algunas realizaciones, la fidelidad de la replicación de pares de bases no naturales es al menos 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 % o 99,99 %. La replicación de un par de bases no natural puede proceder con poco o ningún sesgo de la secuencia, en la que un pequeño sesgo de la secuencia indica que una secuencia desfavorable disminuye la fidelidad en menos del 1 %. Las fidelidades ejemplares se describen en el Ejemplo 1 y se muestran en las Tablas 4, 5 y 6.

Además, en la presente memoria, en diversas realizaciones, se encuentran oligonucleótidos, que incluyen ADN y/o ARN monocatenario y bicatenario (por ejemplo, dúplex), que comprenden una o más nucleobases no naturales descritas en la presente memoria (por ejemplo, cualquier nucleobase α o análogo o derivado del mismo y/o cualquier nucleobase β o análogo o derivado del mismo). La nucleobase puede ser cualquier nucleobase α o β nucleobase descrita en este documento, que incluye las de las Figuras 2, 8, 9, 10, 11, 12, 13, 15, y 16. Un oligonucleótido bicatenario incluye un ADN-ADN dúplex, un ADN-ARN dúplex híbrido y un ARN-ARN dúplex. En algunas realizaciones, el oligonucleótido comprende una nucleobase derivatizada con el enlazador.

En algunas realizaciones, un oligonucleótido comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50 o más nucleobases no naturales. En algunas realizaciones, el porcentaje de nucleobases no naturales en un oligonucleótido está entre aproximadamente 0 % y aproximadamente 1 %, entre aproximadamente 0 % y aproximadamente 2 %, entre aproximadamente 0 % y aproximadamente 3 %, entre aproximadamente 0 % y aproximadamente 4 %, entre aproximadamente 0 % y aproximadamente 5 %, entre aproximadamente 0 % y aproximadamente 10 %, entre aproximadamente 1 % y aproximadamente 10 %, entre aproximadamente 1 % y aproximadamente 15 %, entre aproximadamente 1 % y aproximadamente 20 %, entre aproximadamente 5 % y aproximadamente 10 %, entre aproximadamente 5 % y aproximadamente 20 %, entre aproximadamente 10 % y aproximadamente 30 %, entre aproximadamente 1 % y aproximadamente 50 %, o entre aproximadamente 1 % y aproximadamente 100 %.

Los ejemplos de oligonucleótidos que comprenden una o más nucleobases no naturales incluyen, pero no se limitan a, aptámeros de ADN y aptámeros de ARN. Los aptámeros de ADN y ARN incluyen, pero no se limitan a, cebadores y balizas moleculares. Un aptámero de ADN puede incluir un código de barras.

En algunas realizaciones, un oligonucleótido comprende dTPT3 o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende d5SICS o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende dNaM o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende dMMS o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende dDMS o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende dBrMS o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende dIMS o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende β8a o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende β8b o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido

comprende $\alpha 14a$ o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende $\alpha 14b$ o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende $\alpha 14c$ o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende $\alpha 14d$ o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende $\alpha 14e$ o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende $\alpha 14f$ o un derivado o análogo del mismo.

En algunas realizaciones, un oligonucleótido comprende dTPT3 o un derivado o análogo del mismo, y dNaM o un derivado o análogo del mismo. En algunas realizaciones, un oligonucleótido comprende un par de bases dTPT3-dNaM. En algunas realizaciones, un oligonucleótido comprende uno o más pares de bases seleccionadas de dTPT3-dFEMO, dTPT3-dFIMO, dTPT3-dIMO, dFTPT3-dNaM, dFTPT3-dFEMO, dFTPT3-dFIMO y dFTPT3-dIMO. En algunas realizaciones, un oligonucleótido comprende uno o más pares de bases seleccionados de dTPT3-MMS, dTPT3-DMS, dTPT3-FEMS, dTPT3-BrMS, dTPT3-IMS, dTPT3-DMN, 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-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-dCNMO, 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; en los que una o dos nucleobases no naturales del par de bases no natural pueden derivatizarse con un enlazador.

Un oligonucleótido que comprende una nucleobase no natural divulgada en la presente memoria puede comprender además una o más bases no naturales adicionales, que incluyen, pero no se limitan a, 2-aminoadenin-9-ilo, 2-aminoadenina, 2-F-adenina, 2-tiouracilo, 2-tio-timina, 2-tiocitosina, derivados 2-propil y alquil de adenina y guanina, 2-amino-adenina, 2-amino-propil-adenina, 2-aminopiridina, 2-piridona, 2'-desoxiuridina, 2-amino-2'-desoxiadenosina 3-deazaguanina, 3-deazaadenina, 4-tio-uracilo, 4-tio-timina, 5-il-uracilo, 9-il-hipoxantina (I), 5-metilcitosina, 5-hidroximetilcitosina, xantina, hipoxantina, 5-bromo, y 5-trifluorometil uracilos y citosinas; 5-halouracilo, 5-halocitosina, 5-propinil-uracilo, 5-propinil citosina, 5-uracilo, 5 sustituido, 5-halo, pirimidinas 5-sustituidas, 5-hidroxicitosina, 5-bromocitosina, 5-bromouracilo, 5-clorocitosina, citosina clorada, ciclocitosina, arabinósido de citosina, 5-fluorocitosina, fluoropirimidina, fluorouracilo, 5,6-dihidrocitosina, 5-yodocitosina, hidroxio-urea, yodouracilo, 5-nitrocitosina, 5-bromouracilo, 5-clorouracilo, 5-fluorouracilo, y 5-yodouracilo, 6-alquil derivados de adenina y guanina, 6-azapirimidinas, 6-azo-uracilo, 6-azo citosina, azacitosina, 6-azo-timina, 6-tio-guanina, 7-metilguanina, 7-metiladenina, 7-deazaguanina, 7-deazaguanosina, 7-deaza-adenina, 7-deaza-8-azaguanina, 8-azaguanina, 8-azaadenina, 8-halo, 8-amino, 8-tiol, 8-tioalquilo y adeninas y guaninas 8-hidroxil sustituidas; N4-etilcitosina, purinas N-2 sustituidas, purinas N-6 sustituidas, purinas O-6 sustituidas, aquellas que aumentan la estabilidad de la formación dúplex, ácidos nucleicos universales, ácidos nucleicos hidrófobos, ácidos nucleicos promiscuos, ácidos nucleicos expandidos por tamaño, ácidos nucleicos fluorados, pirimidinas tricíclicas, fenoxazina citidina([5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazina-2(3H)-ona), abrazaderas G, fenoxazina citidina (9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido[4,5-b]indol-2-ona), piridoindol citidina (H-pirido [3',2':4,5] pirrolo [2,3-d] pirimidin-2-ona), 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxilmetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo,

5 beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido oxiacético 5-uracilo, wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, ácido oxiacético 5-uracil-metiléster, ácido oxiacético 5-uracilo, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)w, y 2,6-diaminopurina y aquellos en los que la base de purina o pirimidina se reemplaza con un heterociclo.

10 Un oligonucleótido que comprende una nucleobase no natural divulgada en la presente memoria, puede comprender además una fracción de azúcar no natural, que incluye, pero no se limita a, una modificación en la posición 2': OH; alquilo inferior sustituido, alcarilo, aralquilo, O-alcarilo u O-aralquilo, SH, SCH3, OCN, Cl, Br, CN, CF3, OCF3, SOCH3, SO2 CH3, ONO2, NO2, N3, NH2F; O-alquilo, S-alquilo, N-alquilo; O-alquenilo, S-alquenilo, N-alquenilo; O-alquinilo, S-alquinilo, N-alquinilo; O-alquil-O-alquilo, 2'-F, 2'-OCH3, 2'-O(CH2) 2OCH3 en las que el alquilo, alquenilo y alquinilo pueden estar sustituidos o sustituidos en C1-C10, alquilo, alquenilo C2-C10, alquinilo C2 -C10, -O[(CH2)n O] mCH3, -O(CH2)nOCH3, -O(CH2)n NH2, -O(CH2)n CH3, -O(CH2)n-OH2 y -O(CH2)nON[(CH2)n CH3)]2, donde n y m son de 1 a aproximadamente 10; y/o una modificación en la posición 5': 5'-vinilo, 5'-metilo (R o S), una modificación en la posición 4', 4'-S, heterocicloalquilo, heterocicloalcarilo, aminoalquilamino, polialquilamino, sililo sustituido, un grupo de escisión de ARN, un grupo indicador, un intercalador, un grupo para mejorar las propiedades farmacocinéticas de un oligonucleótido, o un grupo para mejorar las propiedades farmacodinámicas de un oligonucleótido, y cualquier combinación de los mismos.

20 En algunas realizaciones, el oligonucleótido que comprende una nucleobase no natural divulgada en la presente memoria, comprende además una cadena principal no natural. Una cadena principal no natural incluye, pero no se limita a, fosforotioato, fosforotioato quirral, fosforoditioato, fosfotriéster, aminoalquilfosfotriéster, fosfonatos C1-C10, fosfonato 3'-alquileo, fosfonatos quirales, fosfinatos, fosforamidatos, 3'-amino fosforamidato y aminoalquilfosforamidatos, tionofosforamidatos, tionalquilfosfonatos, tionalquilfosfotriésteres, y boranofosfonatos.

25 Los procedimientos para determinar la estabilidad de los dúplex de oligonucleótidos que comprenden pares de bases no naturales (con o sin enlazadores) incluyen análisis termodinámico mediante mediciones de dicroísmo circular (CD) y experimentos de fusión UV. En algunas realizaciones, se emplean estudios de estabilidad de ADN dúplex para facilitar la selección de un par de bases de nucleótidos no natural adecuado, nucleobase no natural o derivados o sustituciones de nucleobase no natural. Los pares de bases no naturales seleccionados adecuadamente incluyen aquellos que aumentan la fidelidad de hibridación de oligonucleótidos en otras posiciones dentro del dúplex. Los pares de bases no naturales seleccionadas adecuadamente incluyen aquellos que aumentan la estabilidad del oligonucleótido dúplex. Pueden usarse nucleobases seleccionadas adecuadamente para optimizar los oligonucleótidos para aplicaciones biotecnológicas o terapéuticas donde la hibridación y la discriminación de alta fidelidad son críticas. En algunos casos, un par de bases no natural es al menos 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más, tan estable como un par de bases natural en un oligonucleótido dúplex. En algunos casos, la Tm de un dúplex que comprende uno o más pares de bases no naturales es inferior a 10 °C, 9 °C, 8 °C, 7 °C, 6 °C, 5 °C, 4,5 °C, 4 °C, 3,5 °C, 3 °C, 2,9 °C, 2,8 °C, 2,7 °C, 2,6 °C, 2,5 °C, 2,4 °C, 2,3 °C, 2,2 °C, 2,1 °C, 2 °C, 1,9 °C, 1,8 °C, 1,7 °C, 1,6 °C, 1,5 °C, 1,4 °C, 1,3 °C, 1,2 °C, 1,1 °C, 1 °C, 0,9 °C, 0,8 °C, 0,7 °C, 0,6 °C, 0,5 °C, 0,4 °C, 0,3 °C, 0,2 °C, 0,1 °C por debajo de la Tm del mismo dúplex en el que una o más nucleobases no naturales se reemplazan con una o más nucleobases naturales. En algunas realizaciones, la presencia de un par de bases no natural en un oligonucleótido dúplex no perturba significativamente la estructura dúplex.

45 En algunas realizaciones, un oligonucleótido que comprende una nucleobase derivatizada con el enlazador permite la modificación específica del sitio de ese ADN o ARN durante o después de la síntesis enzimática. Un nucleótido no natural divulgado en la presente memoria (por ejemplo, un nucleótido que comprende un análogo de nucleobase α o β no natural), en algunos casos, se modifica con un enlazador que permite la unión de diferentes grupos funcionales (por ejemplo, carga) sin ablación del reconocimiento de la polimerasa. Las funcionalidades específicas del sitio incluyen, pero no se limitan a fluoróforos, asas de NMR para caracterización (por ejemplo, F19), sondas IR (por ejemplo, Grupos azido y ciano), biotina (por ejemplo Para facilitar la identificación y/o purificación), etiquetas de afinidad, liposomas, y nanopartículas. En una realización, un enlazador proporciona bioconjugación mediante acoplamiento cruzado (por ejemplo, el grupo yodo). En una realización, un enlazador proporciona un asa para la bioconjugación mediante química de clic (por ejemplo, sustituyentes azido y alquino). En una realización, un oligonucleótido que comprende una nucleobase derivatizada con el enlazador es útil como cebador y/o baliza molecular.

55 Además, en la presente memoria se divulga, en diversas realizaciones, el uso de cualquier análogo de nucleósido divulgado en la presente memoria (α o β), o análogos o derivados de los mismos, en la escisión o funcionalización específica de un sitio de un oligonucleótido. En algunas realizaciones, un análogo de nucleósido comprende uno o más enlazadores configurados para modificación específica del sitio. Los ejemplos de análogos de nucleótidos que comprenden una fracción enlazadora incluyen, pero no se limitan a, d5SICSCO, d5SICSCC, dDMOCO, dDMOCC, dMMO2pCO, dMMO2pCC, dTPT3, dTPT3A, dTPT3PA, dTPT3CO, dMMSpCO, dMMSPA y dTPT3 o formas de ribosilo de los mismos, o análogos de los mismos. En esta memoria se divulgan, en diversas realizaciones, composiciones de materia per se de los oligonucleótidos funcionalizados, procedimientos de preparación de los oligonucleótidos funcionalizados, y procedimientos de uso de los oligonucleótidos funcionalizados. Diversas realizaciones proporcionan dTPT3, dTPT3PA, dTPT3A, dTPT3CO y dTPT3CC, u otros análogos derivatizados con el enlazador de dTPT3, incorporados en oligonucleótidos y la reacción o derivatización adicional de estos análogos de nucleobase no naturales

incorporados en un oligonucleótido con diversos reactivos para reacción selectiva con los análogos de nucleobase no naturales en un oligonucleótido en el que las nucleobases naturales (A, T, G, C, U) no reaccionan con estos reactivos en ningún grado apreciable. La familia en base a dTPT3 de los nucleótidos no naturales que portan enlazador es especialmente central, ya que hemos descubierto que se replican más eficientemente por las ADN polimerasas que los pares de bases que incluyen d5SICS o sus variantes derivatizados con el enlazador, lo que facilitará significativamente muchas de las aplicaciones potenciales. En algunas realizaciones, el por ciento de incorporación de un nucleótido no natural que comprende un enlazador en un oligonucleótido es al menos 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 %, o 99 %. En algunas realizaciones, el por ciento de extensión de un nucleótido correcto próximo en un oligonucleótido, en el que el nucleótido correcto próximo sigue a la incorporación de un nucleótido no natural que comprende un enlazador, es al menos 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 % o más. En algunas realizaciones, la adición de una funcionalidad específica del sitio disminuye el por ciento de incorporación de un nucleótido no natural en un oligonucleótido por al menos aproximadamente 50 %, 45 %, 40 %, 35 %, 30 %, 25 %, 20 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 %. En algunas realizaciones, la fidelidad de un nucleótido no natural derivatizado con el enlazador es al menos 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 % o 99,99 %. Por consiguiente, en diversas realizaciones, en la presente memoria se divulgan procedimientos para usar los nucleótidos no naturales derivatizados con el enlazador para producir ADN o ARN que se modifica específicamente en el sitio con otra molécula de interés. En algunas realizaciones, la inclusión específica del sitio de diferentes funcionalidades ocurre antes o después de la amplificación. En algunas realizaciones, la funcionalización específica del sitio se emplea para aplicaciones SELEX.

Una estrategia ejemplar para producir ADN o ARN que se modifica específicamente en el sitio con otra molécula de interés se denomina estrategia de fosforotioato (Figura 3), que se basa en la incorporación específica del sitio de un grupo fosforotioato en un ADN o ARN a través de un ribo o desoxirribo α -tiotriofosfato de uno de los nucleósidos no naturales de la Figura 1 o 2. Después de la incorporación en el ADN o el ARN, el fosforotioato puede usarse para acoplar reactivos que portan fracciones de γ -bromo- α,β -carbonil-, yodo (o bromo)acetil- o aziridinilsulfonamida para producir ADN o ARN funcionalizado específicamente para el sitio. Alternativamente, después de la incorporación en el ADN o ARN, el fosforotioato puede usarse para escindir el ADN o ARN de forma específica del sitio mediante el uso de yodo en una solución alcalina o yodoetanol, respectivamente. Por lo tanto, la estrategia del fosforotioato proporciona una modificación específica del sitio de la cadena principal de ácido nucleico, y proporciona un procedimiento de escisión específica del sitio de la cadena de oligonucleótidos.

Otra estrategia para producir ADN o ARN que se modifica específicamente en el sitio con otra molécula o interés, denominada *estrategia de enlace* (Figuras 4 y 5), hace uso de la derivatización de nucleobases no naturales con un enlazador (Figura 2) que puede usarse para unir grupos funcionales de interés, ya sea antes de la polimerización (mediante PCR o transcripción mediada por ARN polimerasa de T7 mediante el uso de un reactivo de trifosfato de nucleobase funcionalizado apropiado que se incorpora a la cadena de ADN o ARN que se sintetiza), o por reacción del enlazador de la nucleobase no natural después de la incorporación a la cadena de oligonucleótidos con un reactivo de funcionalización apropiado, por ejemplo, un reactivo que contiene NHS que también comprende el grupo funcional conveniente, en el que el NHS reacciona con el grupo amino libre de una nucleobase no natural funcionalizada con amino tal como d5SICS^A, dMMO2^A o dTPT3^A. La Figura 4 también muestra la estrategia del enlazador de funcionalización de amino mediante el uso de d5SICS^A y dMMO2^{PA} que permite el doble marcaje específico del sitio del ADN dúplex.

Por ejemplo, la funcionalización puede lograrse después de la incorporación en el oligonucleótido de la nucleobase no natural con el enlazador que porta un grupo amino primario (por ejemplo, dTPT3^A). Más específicamente, la funcionalización puede llevarse a cabo mediante la reacción del amino primario (por ejemplo, grupo propargilamino) y un reactivo portador de carga que incluye un éster de N-hidroxisuccinimida (NHS) (Figura 4). Los análogos desarrollados para esta aplicación incluyen d5SICS^A, d5SICS^{PA}, dMMO2^A, dMMO2^{PA}, dTPT3^{PA}, y dTPT3^A (recuerde que "A" se refiere al nucleótido con una propargilamina y "PA" se refiere al mismo enlazador con un grupo protector, véase la Figura 2 y su leyenda). El uso de dTPT3^{PA}, que porta un grupo propargilamino primario protegido, y dTPT3^A, que porta el grupo propargilamino primario, para la funcionalización de oligonucleótidos de secuencia específica, se divulga y reivindica en la presente memoria.

La funcionalización específica del sitio de un oligonucleótido también puede lograrse mediante el uso de la cicloadición de azida-alquino catalizada con cobre (I) (CuAAC) (es decir Estrategia de enlace "Química de clics"; Figura 5), y para estas aplicaciones d5SICS^{CO}, d5SICS^{CC}, dDMO^{CO}, dDMO^{CC}, dMMO2^{pCO}, dMMO2^{pCC}, dTPT3^{CO} y dTPT3^{CC} (Figura 2), puede utilizarse. En cada caso, los ribosiltrifosfatos de las nucleobases no naturales pueden emplearse para la transcripción para producir ARN marcado específicamente para el sitio, y los desoxirribosiltrifosfatos de las nucleobases no naturales pueden usarse, por ejemplo, en PCR, para producir ADN marcado específicamente para el sitio. Las nucleobases no naturales que comprenden un grupo enlazador acetilénico (alquinilo) adecuado para usar en la conjugación de CuAAC, d5SICS^{CO}, d5SICS^{CC}, dDMO^{CO}, dDMO^{CC}, dMMO2^{pCO}, dMMO2^{pCC}, dTPT3^{CO} y dTPT3^{CC}, los

procedimientos de su preparación, y los procedimientos de su uso en la preparación de tales oligonucleótidos marcados específicamente para el sitio, se divulgan y reivindican en la presente memoria.

5 *Demostración de la estrategia general de fosforotioato* (Figura 3). Para demostrar la viabilidad de nuestro sistema, hemos preparado el α -tiorifosfato del nucleótido no natural, d5SICS (d5SICS- α S), y lo hemos incorporado en el ADN opuesto a su nucleótido no natural afín dNaM, mediante el uso de PCR estándar. La eficiencia de amplificación y la fidelidad de incorporación de d(5SICS- α S)TP es superior al 99 % y prácticamente idéntica a los resultados obtenidos con d5SICS. Para funcionalizar este par de bases no natural, reaccionamos el enlace de fosforotioato incorporado específicamente en el sitio con yodoacetil-PEG₂-biotina para marcar el ADN dúplex con la funcionalidad de biotina.⁸
10 Para caracterizar este aducto específico del sitio, lo incubamos en presencia de estreptavidina y después cuantificamos la funcionalización mediante el ensayo de cambio de gel. Pudimos convertir 60-70 % del enlace de fosforotioato en el derivado funcionalizado, que es una eficiencia estándar (70 %) para los protocolos de marcaje informados previamente en la literatura (Véase 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.). Estos derivados conjugados muestran una
15 alta estabilidad en condiciones típicas para la desnaturalización por calor de los dúplex de ADN, *es decir* a 50 °C durante una noche dentro del rango de pH 6,0-8,3 (< 10 % de descomposición), así como a 95 °C durante 3 minutos a pH 8,3 (<5 % de descomposición). Prevedemos que la estrategia del fosforotioato puede emplearse igualmente bien con otros pares de bases no naturales, que incluyen d5SICS-dMMO2 y d5SICS-dNaM.

20 En la presente memoria se divulga, en diversas realizaciones, una estrategia de fosforotioato que usa un par de bases no natural dTPT3-dNaM, dTPT3-dMMO2, o dTPT3-dDMO, y variantes de las mismas derivatizadas con el enlazador.

Las estrategias en base a fosforotioato y enlazador no son mutuamente excluyentes y cuando se combinan deberían permitir que un sitio determinado se modifique simultáneamente con hasta tres grupos funcionales diferentes, uno se
25 une a una primera nucleobase de un par de nucleobases, un segundo se une a una segunda nucleobase de un par de nucleobases, y un tercero se une a la cadena principal 5' inmediatamente a un nucleótido no natural.

Demostración de la estrategia de enlace con la amina primaria (Figura 4). Para demostrar aún más la viabilidad de nuestro sistema, se sintetizaron y caracterizaron las variantes amino y amino protegidas derivatizadas con el enlazador de d5SICS y dMMO2 (Figura 2). Cuando los compañeros no naturales afines se emparejaron en el ADN opuesto, demostramos que cada uno se amplificó bien por PCR y se transcribió en ARN. El acoplamiento de ADN que contiene dMMO2^A o d5SICS^A preparado mediante amplificación por PCR con NHS-éster biotina procede con 55 % y 70 % de eficiencia, respectivamente.

35 Hemos demostrado que los ribonucleótidos trifosfatos de 5SICS^{PA}, 5SICS^A, MMO2^{PA} o MMO2^A, se transcriben en ARN mediante la ARN polimerasa T7 con alta eficiencia y fidelidad (Figura 7).

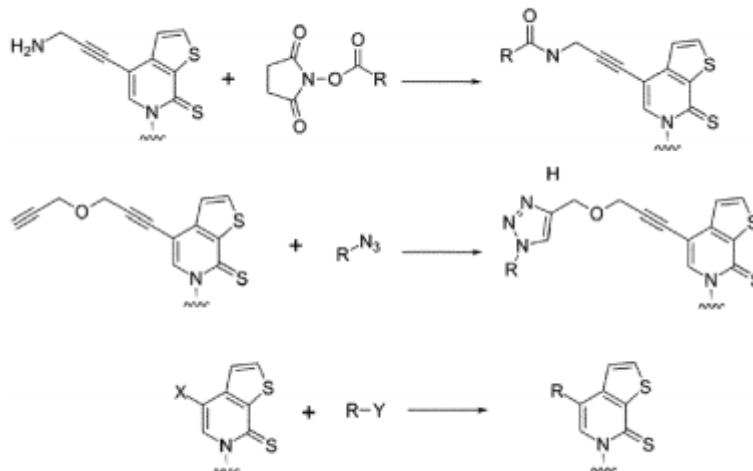
En esta memoria se divulga, en diversas realizaciones, la modificación específica del sitio de ADN o ARN mediante el uso de dTPT3^L-dNaM, dTPT3^L-dMMO2 o dTPT3^L-(d)DMO (donde R es un enlazador, *por ejemplo* R = H para dTPT3, R = 3-aminopropin-1-ilo para dTPT3^A, R = dicloroacetil-3-aminopropin-1-ilo para dTPT3^{PA}, R = 4-oxahepat-1,6-diin-1-ilo para dTPT3^{CO}, R = hepta-1,6-diin-1-ilo para dTPT3^{CC}).

45 *Demostración de la estrategia general de enlace con alquinos*. (Figura 5). Para demostrar aún más la viabilidad de nuestro sistema, hemos sintetizado y caracterizado las variantes funcionalizadas con alquino de d5SICS, dDMO y dMMO2, que incluye d5SICS^{CO}, d5SICS^{CC}, dDMO^{CO}, dDMO^{CC}, dMMO2^{PCO}, dMMO2^{PCC} (Figura 2) cada uno de estos nucleótidos no naturales funcionalizados con alquino debe amplificarse de manera eficiente mediante PCR cuando está presente en el ADN. Una vez amplificado, el ADN que contiene, por ejemplo, el par de bases d5SICS^{CO}-dNaM, puede modificarse eficazmente específicamente en el sitio con moléculas pequeñas o una o más proteínas que poseen grupos azida mediante el uso de química de Clic, por ejemplo, reacciones click catalizadas por cobre. También hemos
50 demostrado la utilidad de dEMO y dFEMO (Figura 2) para la incorporación en oligonucleótidos y el uso de estos oligonucleótidos funcionalizados en reacciones químicas de clic con azidas para funcionalizar los oligonucleótidos de una manera específica del sitio.

Demostración de la estrategia de enlace con dTPT3^{PA} (Figura 6). La adición de un enlazador a los andamios d5SICS y dMMO2 reduce significativamente la eficiencia con la que los nucleótidos no naturales se incorporan enzimáticamente en el ADN, lo que se espera que limite sus aplicaciones prácticas. Sin embargo, hemos encontrado que el andamio dTPT3 es mucho más tolerante a la adición del enlazador (Figura 6). Por ejemplo, dTPT3^{PA}TP se incorpora en un cebador opuesto a dNaM en un templo por las ADN polimerasas con prácticamente la misma eficiencia y fidelidad que un par de bases naturales. En consecuencia, en la presente memoria se divulga, en diversas realizaciones, el uso de nucleobases no naturales en base al andamio dTPT3, que incluye dTPT3^{PA} (enlazador aminofuncional protegido), dTPT3^A (enlazador aminofuncional), y dTPT3^{CO} (enlazador éter alquino-azida para la derivatización mediante química de clics), y dTPT3^{CC} (enlazador de trimetileno alquino-azida para la derivatización mediante química de clic) en la
60 síntesis de oligonucleótidos funcionalizados específicos del sitio.

65 El esquema 1 ilustra ejemplos de dTPT3 con diferentes enlazadores que podrían usarse para modificar el ADN o ARN específicamente en el sitio.

Esquema 1:



5
10
15
20

Para mayor claridad solo se muestran las fracciones de nucleobase de andamio dTPT3, pero se entiende que se usan como nucleótidos. Las reacciones de funcionalización pueden llevarse a cabo antes o después de la incorporación de las nucleobases no naturales en un oligonucleótido. El esquema 1, reacción superior, ilustra el uso de dTPT3^A que comprende un enlazador que porta una amina primaria que se acila mediante el uso de un éster activado para formar una amida, en la que el grupo R comprende la carga. La reacción intermedia del Esquema 1 ilustra el uso de dTPT3^{CO} (dTPT3^{CC} podría usarse también) que comprende un enlazador que porta alquínilo que reacciona con una azida para producir un triazol mediante química de Clic, en el que el grupo R del triazol que se forma comprende la carga. La reacción inferior ilustra el caso más general para un derivado de andamio dTPT3 que porta un grupo enlazador R₁ con una fracción reactiva que puede formar selectivamente un enlace covalente con un grupo R₂ que incluye una fracción reactiva complementaria a la fracción reactiva del enlazador, por ejemplo, tiol-maleimida, hidrazina-aldehído, etc.

En una realización, un enlazador que comprende un grupo reactivo azida es útil para unir una carga que comprende alquino a través de una reacción de clic. En una realización, un enlazador que comprende un grupo tiol puede formar enlaces disulfuro reversibles o enlaces irreversibles con una variedad de grupos de aceptación de carga, que incluyen, pero no se limitan a, maleimida, bromuro, yoduro, derivados de sulfonilo, ésteres activos y derivados de isotiocianato. En una realización, un enlazador que comprende un grupo azida es reactivo con una molécula de carga que comprende un grupo fosfina.

En una realización, un oligonucleótido que comprende una o más nucleobases no naturales derivatizadas con el enlazador se configura para su uso como baliza molecular. El fluoróforo de la baliza molecular es una molécula de carga unida a un centro reactivo de la nucleobase no natural derivatizada con el enlazador. Las moléculas de carga de fluoróforo ejemplares incluyen, pero no se limitan a, 6-FAM, Fluoresceína, Cy3TM, JOE (6-carboxi-4',5'-dicloro-2',7'-dimetoxifluoresceína), Cy5TM, TAMRA, MAX, TETTM, ROX (carboxi-X-rodamina), TYETM 563, Hexaclorofluoresceína, TEX 615, TYETM 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, ATTOTM 488, ATTOTM 532, ATTOTM 550, ATTOTM 565, ATTOTM Rho101, ATTOTM 590, ATTOTM 633, ATTOTM 647N, Rhodamine GreenTM-X, Rhodamine RedTM-X, 5-TAMRATM, Texas Red®-X, Lightcycler® 640 y Dy 750.

Un par de bases no natural, en algunas realizaciones, permite la inclusión específica del sitio de diferentes funcionalidades en el ADN para la Evolución sistemática de ligandos mediante aplicaciones de enriquecimiento exponencial (SELEX), que incluye la generación de aptámeros de ADN y/o ARN. Los aptámeros de ADN y ARN tienen una variedad de objetivos, que incluyen ácidos nucleicos, moléculas pequeñas, péptidos, carbohidratos, y células. SELEX incluye la creación de una biblioteca de moléculas de ácido nucleico, que contacta la biblioteca con moléculas objetivo para seleccionar moléculas de ácido nucleico que se unen a las moléculas objetivo, y amplifican los miembros de la biblioteca que se unen a las moléculas objetivo. Se continúan rondas adicionales de selección y amplificación hasta que se recuperan suficientes aptámeros. Un aptámero, en un aspecto, incluye cualquier base no natural divulgada en la presente memoria. En algunas realizaciones, un experimento SELEX, en el que los componentes de la biblioteca comprenden nucleobases no naturales, genera una afinidad de aptámero contra una molécula objetivo en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o menos rondas de selección que una biblioteca que no comprende nucleobases no naturales. En algunas realizaciones, un aptámero que comprende una o más nucleobases no naturales tiene una mayor afinidad por una molécula objetivo que un aptámero que contiene solo nucleobases naturales. La adición de una o más nucleobases no naturales en una biblioteca SELEX aumenta la diversidad química y estructural de los aptámeros de ADN o ARN resultantes. En algunas realizaciones, un aptámero no natural tiene al menos una afinidad nanomolar contra su molécula objetivo. En algunas realizaciones, un aptámero no natural tiene al menos una afinidad picomolar contra su molécula objetivo. Por ejemplo, un aptámero no natural tiene una afinidad por su molécula objetivo que está entre 1 y 1.000 pM. En algunas realizaciones, un aptámero no natural tiene al menos una afinidad femtomolar por su molécula

65

- objetivo. Por ejemplo, un aptámero no natural tiene una afinidad por su molécula objetivo que está entre 1 y 1.000 fM. Un aptámero no natural seleccionado mediante el uso de SELEX puede comprender 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 o más nucleobases no naturales. En algunas realizaciones, un aptámero no natural comprende dTPT3 o un derivado o análogo del mismo. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14a$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14b$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14c$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14d$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14e$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\alpha 14f$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\beta 8a$ o un derivado o análogo de la misma. En algunas realizaciones, un aptámero no natural comprende una nucleobase que tiene la fórmula $\beta 8b$ o un derivado o análogo de la misma.
- Se pueden proporcionar varias combinaciones de los componentes expuestos anteriormente con respecto a mezclas de reacción ejemplares y procedimientos de reacción en forma de kit. Dicho kit puede incluir componentes individuales que están separados entre sí, por ejemplo, que se transportan en recipientes o paquetes separados. Un kit puede incluir una o más subcombinaciones de los componentes expuestos en la presente memoria, mediante la separación de una o más subcombinaciones de otros componentes del kit. Las subcombinaciones pueden combinarse para crear una mezcla de reacción expuesta en la presente memoria (o se combina para realizar una reacción expuesta en la presente memoria). En realizaciones particulares, una subcombinación de componentes que está presente en un recipiente o paquete individual es insuficiente para realizar una reacción expuesta en la presente memoria.
- Sin embargo, el kit en su conjunto puede incluir una colección de recipientes o paquetes cuyo contenido puede combinarse para realizar una reacción expuesta en la presente memoria.
- Un kit puede incluir un material de embalaje adecuado para albergar el contenido del kit. El material de embalaje se construye mediante procedimientos bien conocidos, preferentemente para proporcionar un ambiente estéril, libre de contaminantes. Los materiales de embalaje empleados en la presente memoria pueden incluir, por ejemplo, aquellos utilizados habitualmente en kits comerciales vendidos para su uso con sistemas de secuenciación de ácido nucleico. Los materiales de embalaje ejemplares incluyen, sin limitación, vidrio, plástico, papel, papel de aluminio y similares, capaces de contener dentro de límites fijos un componente expuesto en la presente memoria.
- El material de embalaje puede incluir una etiqueta que indique un uso particular para los componentes. El uso para el kit que se indica en la etiqueta puede ser uno o más de los procedimientos expuestos en la presente memoria según sea apropiado para la combinación particular de componentes presentes en el kit. Por ejemplo, una etiqueta puede indicar que el kit es útil para un procedimiento de conjugación de una molécula de carga con una fracción enlazadora de una nucleobase no natural en un oligonucleótido.
- Las instrucciones para el uso de los reactivos o componentes empaquetados también se pueden incluir en un kit. Las instrucciones típicamente incluirán una expresión tangible que describa los parámetros de reacción, tales como las cantidades relativas de los componentes del kit y la muestra a mezclar, los períodos de tiempo de mantenimiento para los reactivos/aditivos de muestra, la temperatura, las condiciones del tampón y similares.
- Se entenderá que no todos los componentes necesarios para una reacción particular necesitan estar presentes en un kit particular. En su lugar, se pueden proporcionar uno o más componentes adicionales de otras fuentes. Las instrucciones proporcionadas con un kit pueden identificar el(los) componente(s) adicional(es) que se deben proporcionar y dónde se pueden obtener.
- En una realización, un kit proporciona una o más nucleobases no naturales o derivados de las mismas y reactivos configurados para realizar una funcionalización específica del sitio mediante el uso de una o más nucleobases o derivados no naturales de las mismas.

Ejemplos

- Actualmente, los nucleósidos y fosforamiditas libres de d5SICS y dNaM están disponibles comercialmente en Berry and Associates (Dexter, MI).

Ejemplo 1. Cribado basado en PCR para identificar pares de bases no naturales.

- Los trifosfatos del grupo $\alpha 6$ se prepararon a partir de los nucleósidos informados previamente (Kubelka, T., Slavetinska, L., Eigner, V. y Hocek, 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) de acuerdo a Ludwig, J. y Eckstein, F. Rapid and efficient synthesis of nucleoside 5'-O-(1-thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.*, 54, 631-635. La pureza de todos los demás trifosfatos se confirmó mediante MALDI-TOF y UV-VIS. Las ADN polimerasas Taq y OneTaq se adquirieron de New England Biolabs (Ipswich, MA). Se adquirió una mezcla de dNTP de Fermentas (Glen Burnie, MD).

La tinción de gel de ácido nucleico SYBR Green I (10.000 ×) se adquirió de Life Technologies (Carlsbad, CA). La síntesis de las plantillas de ADN, D8 (Malyshev, D.A., Dhami, K., Quach, H.T., Lavergne, T., Ordoukhanian, P., Torkamani, A. y 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 (30), 2012, 12005-12010), usada para para las rondas de cribado 1-5, y D6 (Malyshev, D.A., Seo, Y.J., Ordoukhanian, P. y Romesberg, F.E. PCR with an expanded genetic alphabet. J. Am. Chem Soc., 131, 14620-14621), usada para todas las demás amplificaciones, se describió anteriormente. La secuenciación de Sanger se realizó como se describió anteriormente (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). Las trazas de secuenciación de Sanger sin procesar se usaron para determinar el por ciento de retención de los pares de bases no naturales, que se convirtió en fidelidad por duplicación, como se describe (Malyshev, D.A., Dhami, K., Quach, H.T., Lavergne, T., Ordoukhanian, P., Torkamani, A. y 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. y Romesberg, F.E. PCR with an expanded genetic alphabet. J. Am. Chem. Soc., 131, 14620-14621).

Todas las amplificaciones por PCR se realizaron en un sistema de detección de PCR en tiempo real CFX Connect (Bio-Rad), en un volumen total de 25 µl mediante el uso de las siguientes condiciones: Tampón de reacción OneTaq 1 ×, Sybr Green I 0,5 ×, MgSO4 ajustado a 4,0 mM, 0,2 mM de cada dNTP, 50 µM de cada trifosfato no natural, 1 mM de Primer1 y Primer2 (Véase la Tabla 2), y 0,02 U/µl de ADN polimerasa. Otras condiciones específicas para cada ronda de cribado se describen en la Tabla 3. Los productos amplificados se purificaron mediante el uso de columnas de centrifugación DNA Clean y Concentrator-5 de Zymo Research (Irvine, CA). Después de la purificación, los productos de PCR se secuenciaron en un analizador de ADN 3730 (Applied Biosystems) para determinar la retención del par de bases no natural como se describe a continuación. La fidelidad se caracterizó por la retención de pares de bases no naturales (UBP) según se determinó mediante secuenciación con Primer1 en un analizador de ADN 3730 (Applied Biosystems).

Tabla 2. Secuencias de ADN.

Nombre	Secuencia (5' a 3')	Anotaciones
Cebador1	<u>CACACAGGAAACAGCTATGAC</u>	Cebadores para PCR
Cebador2	<u>GAAATTAATACGACTCACTATAGG</u>	
Cebador1-poli-dT	TT TTTTTTCACACAGGAAACAGCTATGAC	Cebadores para la secuenciación de Sanger
Cebador2-poli-dT	TT TTTTTTTTGAAATTAATACGACTCACTATAGG	
D8	<u>CACACAGGAAACAGCTATGACCCGGGTTATTACATGCGCTAGCACTT</u> GGAATTCACAATACT NaM TCTTTAAGGAAACCATAGTAAATCTCCTT CTTAAAGTTAAGCTTAACCCCTATAGTGAGTCGTATTAATTC	
D6	<u>CACACAGGAAACAGCTATGACCCGGGTTATTACATGCGCTAGCACTT</u> GGAATTCACCAGACGNNN NaM NNNCGGGACCCATAGTAAATCTCCTT TCTTAAAGTTAAGCTTAACCCCTATAGTGAGTCGTATTAATTC	N = nucleótido natural aleatorizado

Tabla 3. Condiciones de PCR para cada ronda consecutiva del cribado por PCR.

Componentes de la reacción	Rondas 1-4	Ronda 5	Ronda 6	Caracterización final de PCR	
Tampón	OneTaq 1 ×	OneTaq 1 ×	OneTaq 1 ×	OneTaq 1 ×	OneTaq 1 ×
Enzima	OneTaq	Taq	OneTaq o Taq	Taq	OneTaq
Plantilla	D8 ^a (0,1 ng)	D8 ^a (0,1 ng)	D6 ^a (0,01 ng)	D6 ^a (0,01 ng)	D6 ^a (0,01 ng)
Condiciones térmicas					
Desnaturalización inicial	96 °C, 60 s	96 °C, 60 s	96 °C, 60 s	96 °C, 60 s	96 °C, 60 s
Desnaturalización	96 °C, 10 s	96 °C, 10 s	96 °C, 5 s	96 °C, 5 s	96 °C, 10 s
Hibridación	60 °C, 15 s	60 °C, 15 s	60 °C, 5 s	60 °C, 5 s	60 °C, 5 s
Extensión	68 °C, 60 s	68 °C, 60 s	68 °C, 10 s	68 °C, 10 s	68 °C, 30 s
número de ciclos	16	16	24	16 +16 +20 ^b	16 +16 +20 ^b
^a Véase la Tabla 2 para las secuencias de las plantillas y los cebadores. ^b La cantidad inicial de plantilla fue de 0,01 ng. La mezcla de PCR se amplificó durante 16 ciclos, se diluyó 40.000 veces y se amplificó durante otros 16 ciclos. La etapa de dilución/amplificación se repitió y dió como resultado 52 ciclos totales de amplificación.					

5 Condiciones específicas del ensayo de PCR. La PCR con los UBP más prometedores se realizó con las condiciones que se describen en la Tabla 3. Los productos de PCR se purificaron adicionalmente en geles de agarosa al 2 %, seguido de una escisión de banda única y posterior limpieza mediante el uso del kit de recuperación de ADN ZymoClean Gel de Zymo Research. Después de la elución con 20 µl de agua, se midió la concentración de ADN mediante el uso de un colorante fluorescente (kit de ensayo Quant-iT dsDNA HS, Life Technologies), y los amplicones purificados se secuenciaron por triplicado con Primer1 y Primer2 para determinar la retención de UBP y, por lo tanto, la fidelidad de amplificación. La amplificación de ADN que contiene los pares que involucran análogos del grupo α6 se realizó con la polimerasa OneTaq en las siguientes condiciones de ciclo térmico: desnaturalización inicial a 96 °C durante 1 min; 16 ciclos de 96 °C durante 10 s, 60 °C durante 15 s, 68 °C durante 1 min. La fidelidad se determinó mediante secuenciación de amplicones en la dirección de Primer1 por triplicado. La amplificación del ADN que contiene los UBP formados entre dTPT3 y d2MN o dDM2 se realizó mediante el uso de polimerasas OneTaq o Taq durante 16 ciclos en las siguientes condiciones de ciclo térmico: 1) OneTaq: desnaturalización inicial a 96 °C durante 1 min, 96 °C durante 10 s, 60 °C durante 15 s, 68 °C durante 1 min; o 2) Taq: desnaturalización inicial a 96 °C durante 1 min, 96 °C durante 5 s, 60 °C durante 5 s, 68 °C durante 10 s. La fidelidad se determinó mediante la secuenciación de los amplicones en la dirección de Primer1 por triplicado.

10 Resultados. Para cribar los UBP bien replicados, los trifosfatos de desoxinucleósidos no naturales se agruparon para su análisis en análogos de dMMO2/dNaM- o d5SICS/dTPT3, aunque la distinción no está completamente clara en todos los casos. En total, 80 análogos dMMO2/dNaM se agruparon en doce "grupos α" (α1 - α12; Figura 8), y 31 análogos d5SICS/dTPT3 se agruparon en seis "grupos β" (β1 - β6; Figura 9). Tenga en cuenta que las designaciones de grupo usados aquí no deben confundirse con la designación de anómero (todos los análogos examinados son glucósidos β). Además, para aumentar el contenido de la relación estructura-actividad (SAR) del cribado, se fosforilaron e incluyeron como grupo α6 siete análogos de nucleósidos informados previamente (dTOK576-dTOK588) con nucleobases de piridilo sustituido (Kubelka, T., Slavetinska, L., Eigner, V. y Hocekm 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.). Para el cribado, se amplificó por PCR una plantilla de ADN monocatenario de 134 mer que contiene un dNaM ubicada centralmente (se denomina D8) en presencia de trifosfatos naturales (200 µM cada uno), todas las combinaciones de un grupo α y β de un trifosfato se muestra en las Figuras 8 y 9 (50 µM cada una) y 0,02 U/µl de ADN polimerasa. Durante la primera ronda de PCR, las plantillas dNaM incorporan un análogo α y después se reemplaza por un análogo β cuando la cadena original se copia en la segunda ronda, con la amplificación de los UBP resultantes en rondas posteriores. El producto de amplificación de cada reacción se analizó mediante secuenciación de Sanger. La presencia de un nucleótido no natural da como resultado la terminación abrupta del cromatograma de secuenciación, lo que permite cuantificar el nivel de retención de UBP por la cantidad de lectura. El porcentaje de UBP retenido en el ADN después de la amplificación durante cada ronda de cribado se muestra en la Figura 14.

15 La primera ronda de cribado empleó 0,1 ng de plantilla y 16 ciclos de amplificación en condiciones relativamente permisivas que incluyeron la polimerasa OneTaq y un tiempo de extensión de 1 minuto. Para este ejemplo, OneTaq se considera permisivo porque es una mezcla de Taq (una polimerasa de la familia A) y Deep Vent (una polimerasa de la familia B), y esta última posee corrección de lectura exonucleotídica que permite la escisión de un trifosfato incorporado incorrectamente. En estas condiciones, solo los pares que involucran al grupo β5 o β6 mostraron alta retención.

20 Las combinaciones de β5 o β6 y los grupos α que mostraron la retención más alta progresaron a una segunda ronda de cribado, en la que se dividieron en grupos más pequeños (se indican por a, b o c; Figuras 8 y 9). Se observó una alta retención (≥97 %) con β5a y α2c, α9a, α9c, α10a, α10c, α12b o α12c; con β5b y α9a, α9b, α10c o α12b; y con β6b y

α 10c. Se observó retención moderada (84 - 96 %) con β 5a y α 1a, α 1b, α 6a, α 9b, α 10b o α 12a; β 5b y α 1a, α 1b, α 2c, α 6a, α 9c, α 10a, α 12a o α 12c; β 6a y α 1b o α 10c; y β 6b y α 1a, α 6a, α 9a-c, α 10a, α 10b o α 12a-c.

Para una tercera ronda de cribado, los análogos α se analizaron en grupos de solo uno a tres compuestos, y el grupo β 6a se subdividió en sus dos trifosfatos constituyentes, dTPT1TP y dFPT1TP. La retención más alta (≥ 90 %) se observó con β 5a y α 1a, α 2cII, α 9a-c, α 10aI, α 10aII, α 10c, α 12b o dTfMOTP; β 5b y α 9a, α 9c, o α 10c; dFPT1TP y α 10aI; y β 6b y α 1a, α 9a-c, α 10aI, α 10aII, α 10c, α 12b, dNMOTP, dTfMOTP, o dCNMOTP. Solo se observó ligeramente menos retención (80-89 %) con β 5a y α 2cI, α 12a, dNMOTP, dQMOTP, o dTOK587TP; β 5b y α 1a, α 2cII, α 10aI, α 10aII, α 12b o dTOK587TP; dFPT1TP y α 10c; y β 6b y α 12a, dQMOTP, dFuMO1TP o dTOK587TP.

Para una cuarta ronda de cribado, todos los derivados α de la Figura 9 se analizaron como trifosfatos individuales, con la excepción de α 9b y α 9c, que permanecieron agrupados. La retención más alta (≥ 91 %) se observó con β 5a y α 9b, α 9c, dFIMOTP, dIMOTP, dFEMOTP, dMMO2TP, d2OMeTP, dDMOTP, d5FMTP, dNaMTP, dVMOTP, dZMOTP, dC1MOTP, dTfMOTP, dQMOTP, d2MNTP, dDM2TP, o dTOK587TP, β 5b y α 9b, α 9c, dFIMOTP, dIMOTP, dFEMOTP, dNaMTP, dZMOTP, dC1MOTP, dQMOTP, dMM1TP, dDM2TP, o dTOK587TP; análogo β 6 de dFPT1TP y análogos α de d2OMeTP o dNaMTP; y β 6b y α 9b, α 9c, dFIMOTP, dIMOTP, dFEMOTP, dMMO2TP, dDMOTP, dTMOTP, dNMOTP, d5FMTP, dNaMTP, dVMOTP, dZMOTP, dC1MOTP, dTfMOTP, dQMOTP, dCNMOTP, d2MNTP, dTOK587TP, o dFuMO2TP.

Para aumentar la rigurosidad del cribado, se realizó una quinta ronda con polimerasa Taq en lugar de OneTaq, ya que esta carece de actividad de corrección de lectura de exonucleasa y, por lo tanto, aumenta la sensibilidad a la síntesis de pares erróneos. Esta ronda también separó todos los grupos α y β restantes en trifosfatos individuales. La retención más alta (≥ 90 %) se observó con dSICSTP y dNaMTP; dSNICSTP y dNaMTP; dTPT2TP y dFDMOTP; dTPT3TP y dFIMOTP, dIMOTP o dNaMTP; y dFTPT3TP y dFIMOTP, dIMOTP, dFEMOTP, dNMOTP, dNaMTP, dC1MOTP, dTfMOTP o dCNMOTP.

Para diferenciar mejor entre los UBP, avanzamos los sesenta y dos UBP candidatos más prometedores a una sexta ronda de cribado en la que la concentración de la plantilla se redujo 10 veces (a 10 pg) para permitir una mayor amplificación y, por lo tanto, permitir una mayor discriminación, y la plantilla se cambió a 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), donde los tres nucleótidos flanqueantes a cada lado del nucleótido no natural se aleatorizan entre los nucleótidos naturales. Además, los pasos de desnaturalización e hibridación se redujeron a 5 s cada uno, y el tiempo de extensión se redujo a 10 s. En estas condiciones, exploramos la amplificación con OneTaq o con Taq solo. Los resultados con OneTaq mostraron la mayor retención (> 95 %) con dSICSTP y dNaMTP; dSNICSTP y dFEMOTP; dTPT3TP y dFIMOTP, dIMOTP, dFEMOTP, dZMOTP, o dNaMTP; y dFTPT3TP y dIMOTP o dFEMOTP. Se observó retención moderada (86 % - 94 %) con dSICSTP y dFEMOTP o dDM2TP; d5SICSTP y dNaMTP; dSNICSTP o dIMOTP; dTPT2TP y dNaMTP; dTPT3TP y dNMOTP, dC1MOTP, dQMOTP, dCNMOTP o d2MNTP; y dFTPT3TP y dFIMOTP, dNaMTP, dZMOTP, dC1MOTP, dTfMOTP o dCNMOTP. Si bien la retención durante la amplificación mediada por Taq se redujo en general en relación con la de OneTaq, las tendencias generales fueron similares. La retención más alta (> 96 %) se observó con dTPT3TP y dFIMOTP o dIMOTP, y con dFTPT3TP y dFIMOTP. Solo se observó una retención ligeramente inferior (89 %-94 %) con dTPT3TP y dFEMOTP, dNaMTP o dCNMOTP; y dFTPT3TP y dIMOTP, dFEMOTP, dNaMTP, dC1MOTP, dCNMOTP o d2MNTP.

La amplificación con las combinaciones más prometedoras de trifosfatos, dTPT3TP o dFTPT3TP y dFIMOTP, dIMOTP, dFEMOTP o dNaMTP, se realizó después durante 52 ciclos con Taq y un tiempo de extensión de 10 s, para explorar condiciones particularmente estrictas, o con OneTaq y 30 s tiempo de extensión, para explorar condiciones más prácticas (Tabla 4). Las dos cadenas amplificadas se secuenciaron por triplicado para determinar la retención de UBP con alta precisión. Con Taq, dTPT3-dNaM, dTPT3-dFIMO, dFTPT3-dNaM, y dFTPT3-dFIMO mostraron la mayor retención, mientras que los pares que implican dIMO y dFEMO mostraron algo menos de retención. dTPT3-dNaM y dFTPT3-dNaM mostraron la mayor retención con OneTaq, seguidos de cerca por dFTPT3-dFIMO y dTPT3-dFIMO.

Tabla 4.

dβTP	dαTP	Amplificación, 10 ¹² x	Retención, %	Fidelidad por duplicación, %
<i>Taq</i> 10 s de Extensión				
TPT3	FIMO	8,5	84 ± 3	99,60 ± 0,09
	IMO	6,3	81 ± 5	99,50 ± 0,15
	FEMO	5,0	79 ± 3	99,44 ± 0,09
	NaM	5,8	86,5 ± 0,5	99,66 ± 0,01
FTPT3	FIMO	4,8	84 ± 3	99,60 ± 0,09
	IMO	5,6	82 ± 5	99,54 ± 0,13
	FEMO	5,7	81 ± 4	99,51 ± 0,11
	NaM	3,7	91 ± 6	99,76 ± 0,15
5SICS	NaM	9,3	< 50 ^b	< 85 ^b
<i>OneTaq</i> , 1 min de Extensión				
TPT3	FIMO	8,7	84,7 ± 1,1	99,61 ± 0,03
	IMO	9,4	82,9 ± 1,7	99,56 ± 0,05
	FEMO	10,4	82,2 ± 1,0	99,55 ± 0,03
	NaM	8,3	91,2 ± 1,3	99,79 ± 0,03
FTPT3	FIMO	8,2	86 ± 3	99,65 ± 0,08
	IMO	7,1	76,8 ± 1,6	99,38 ± 0,05
	FEMO	6,3	72,4 ± 1,4	99,24 ± 0,04
	NaM	7,0	90 ± 2	99,76 ± 0,06
5SICS	NaM	8,1	77,1 ± 0,7	99,00 ± 0,02

^aLa retención y la fidelidad determinadas como se describe en el Ejemplo 1. ^bLa retención de UBP por debajo del 50 % y, por lo tanto, se estima que la fidelidad sea < 85 %.

Los datos del cribado sugieren que varios pares formados entre dTPT3 y los derivados de α6 en base a piridina que no se examinaron previamente se replicaron razonablemente bien. Por lo tanto, examinamos por triplicado la amplificación del ADN que contiene estos UBP mediante el uso de OneTaq y 16 ciclos de amplificación con tiempos de extensión de 1 min (Tabla 5). Los pares formados entre dTPT3 y dTOK580, dTOK582 o dTOK586 se replicaron mal. Sin embargo, los pares formados entre dTPT3 y dTOK588, dTOK581, dTOK576 y dTOK587 se amplificaron con una retención del 62 %, 65 %, 85 % y 94 %, respectivamente.

Tabla 5. Los datos de amplificación y fidelidad de dTPT3 contra derivados basados en piridina del grupo $\alpha 6$; DM5, MMO2, DMO y NaM también se caracterizaron para la comparación de andamios.

d β TP	d α TP	Amplificación	Retención, %	Fidelidad, %
TPT3	TOK576	780	85,20 \pm 1,12	98,35 \pm 0,14
TPT3	TOK580	1056	< 50 ^b	< 85 ^b
TPT3	TOK581	1034	65,07 \pm 0,15	95,80 \pm 0,02
TPT3	TOK582	1240	< 50 ^b	< 85 ^b
TPT3	TOK586	948	< 50 ^b	< 85 ^b
TPT3	TOK587	818	93,81 \pm 1,35	99,34 \pm 0,15
TPT3	TOK588	666	61,98 \pm 7,09	94,99 \pm 1,13
TPT3	DM5	- ^a	- ^a	- ^a
TPT3	MMO2	1096	90,95 \pm 3,63	99,06 \pm 0,40
TPT3	DMO	864	84,02 \pm 1,92	98,23 \pm 0,23
TPT3	NaM	1004	99,23 \pm 1,12	99,92 \pm 0,11

^a No se detectó amplificación para esta muestra. ^b La retención de pares de bases no naturales fue inferior al 50 % y, por lo tanto, la fidelidad se estima inferior al 85 %.

5

10

Finalmente, los datos del cribado sugieren que los pares formados entre dTPT3 y d2MN o dDM2 se replican razonablemente bien, a pesar de que ni d2MN ni dDM2 poseen un aceptor de enlace orto H supuestamente esencial. Por lo tanto, estos pares se examinaron adicionalmente a través de 16 ciclos de amplificación con OneTaq o Taq solo, y con tiempos de extensión de 1 min o 10 s (Tabla 6). Con Taq solo, se observó solamente una pobre retención. Sin embargo, con OneTaq, la retención fue mejor para ambos pares. La retención del par dTPT3-dDM2 es 58 % y 69 % con tiempos de extensión de 1 min y 10 s, respectivamente. Sorprendentemente, dTPT3-d2MN se amplifica con retenciones de 96 % y 94 % con tiempos de extensión de 1 min y 10 s, respectivamente.

15

Tabla 6. Amplificación y datos de fidelidad de dTPT3 contra d2MN o dDM2, análogos α sin un aceptor de enlace H orto.

Enzima	d β TP	d α TP	Tiempo extensión de	Amplificación	Retención, %	Fidelidad duplicación, % por
OneTaq	TPT3	2MN	1 min	880	95,54 \pm 1,55	99,53 \pm 0,17
			10 s	224	93,53 \pm 1,42	99,15 \pm 0,20
	TPT3	DM2	1 min	1420	57,92 \pm 6,02	94,89 \pm 0,94
			10 s	376	68,46 \pm 4,34	95,65 \pm 0,72
Taq	TPT3	2MN	1 min	- ^a	-	-
			10 s	- ^a	-	-
	TPT3	DM2	1 min	334	< 50 ^b	< 85 ^b
			10 s	266	< 50 ^b	< 85 ^b

^a No se detectó amplificación para estas muestras. ^b La retención de pares de bases no naturales fue inferior al 50 % y, por lo tanto, la fidelidad se estima inferior al 85 %.

20

Discusión. En esta memoria se realizó un cribado basado en PCR para identificar los UBP más prometedores. Para aumentar el contenido de SAR del cribado, se incluyeron siete derivados α novedosos que se basan en un andamio de piridilo con diferentes sustituyentes en las posiciones orto y para al enlace glucosídico.

25

Datos de relación estructura-actividad. Incluso en condiciones permisivas, donde estaba presente la actividad de corrección de lectura exonucleotídica y los tiempos de extensión fueron de 1 minuto, solo los grupos mixtos de análogos α con análogos β mostraron niveles significativos de retención, lo que demuestra que la replicación eficiente requiere el emparejamiento de un andamio α con un andamio β . Sin embargo, los únicos grupos d β que mostraron alta retención fueron $\beta 5$ y $\beta 6$. Esto revela el estatus privilegiado del andamio tipo d5 SIC S/dTPT3 en relación con todos los demás examinados. La contribución dominante a la alta retención con el grupo $\beta 5$ resultó no ser el resultado de pares que involucran d5SICS, sino más bien de pares que involucran dSICS y, en menor medida, dSNICS. Por ejemplo, en todas

5 las condiciones, dSICS-dNaM se replicó mejor que d5SICS-dNaM. d5SICS resultó de la optimización de dSICS para emparejarse con dMMO2; aparentemente, el aumento del volumen de dNaM hace que el grupo metilo añadido sea perjudicial. Además, dSNICS-dNaM se replica casi tan bien (con OneTaq) o mejor (con Taq) que d5SICS-dNaM, lo que sugiere que un sustituyente 6-aza optimiza la síntesis de UBP tras facilitar la inserción del trifosfato no natural opuesto a dNaM o tras aumentar la eficiencia con que las plantillas de nucleótidos no naturales realizan la inserción de dNaMTP. Finalmente, dSNICS-dFEMO también se replica mejor que d5SICS-dNaM, pero solo en presencia de corrección de lectura, lo que sugiere que si bien la inserción de trifosfato puede ser menos eficiente, una mayor eficiencia de la extensión da como resultado un aumento general de la fidelidad. La contribución dominante a la retención de alta fidelidad con el grupo β 6 la proporcionaron dTPT3 y dFTPT3. En general, ambos se emparejaron bien con dNaM, dFEMO, dFIMO o dIMO. dTPT3 se apareó especialmente bien con dFIMO y dIMO, lo que sugiere que el sustituyente para-yodo media interacciones favorables, y este también se apareó bien con dFEMO y especialmente dNaM cuando la actividad de exonucleasa estaba presente. dFTPT3 se apareó bien con dIMO o dFEMO en presencia de actividad de exonucleasa, así como con dFIMO y dNaM en su ausencia.

15 Mientras que el sustituyente de nitrógeno de los análogos α en base a piridina (grupo α 6) fue en general perjudicial para la replicación, un análisis más detallado de los UBP formados con dTPT3 reveló varias tendencias. Un sustituyente metilo, cloro o amino en la posición orto al enlace C-glucosídico dio como resultado pares replicados pobremente, lo que presumiblemente se debe a una extensión pobre después de la incorporación del trifosfato no natural. El sustituyente orto-metoxi de dTOK581 dio como resultado una mejor replicación, que presumiblemente se debe a su capacidad de empaquetarse hidrofóbicamente con la plantilla durante la síntesis de UBP y aceptar un enlace H con un donante de enlace H en base a polimerasa durante la extensión. Los datos también revelaron que el sustituyente orto de metilsulfanilo de dTOK588, dTOK576 y especialmente dTOK587 da como resultado una mejor replicación. Es probable que esta mejora se deba a un compromiso más optimizado entre la capacidad de empaquetarse hidrofóbicamente y la capacidad de aceptar un enlace H de la polimerasa en el terminal del cebador. Además, el sustituyente para en esta serie de derivados puede contribuir a una replicación eficiente, un sustituyente bromo es el mejor, seguido de un segundo grupo metilsulfanilo y, finalmente, un grupo metilo simple. Cuando dTOK587, con su combinación de los sustituyentes orto-metilsulfanilo y para-bromo, se combinó con dTPT3, el UBP resultante se replicó con OneTaq y 1 min de tiempo de extensión con una fidelidad (calculado a partir del nivel de retención) del 99,3 %, que es ligeramente mejor que d5SICS-dMMO2 en condiciones similares. Claramente, los sustituyentes similares orto-metilsulfanilo y para-bromo deben examinarse con los andamios similares a α replicados de manera más eficiente, tales como dFIMO y dNaM.

También merece discusión la replicación de los pares formados entre dTPT3 y d2MN o dDM2. El ADN que contiene estos pares no se amplifica solo con Taq, sino que también se amplifica bien con OneTaq. Este resultado fue inesperado porque ni d2MN ni dDM2 poseen el aceptor de enlace H orto que se ha postulado que es esencial para la extensión del terminal cebador naciente (natural o no natural). Específicamente, cuando se coloca un nucleótido en el terminal del cebador en crecimiento, el aceptor de enlace H se dispone en el surco menor en desarrollo donde este acepta un enlace H de la polimerasa, y se cree que este enlace H es necesario para la alineación correcta del terminal. Cuando se amplifica con OneTaq y un tiempo de extensión de 1 min, dTPT3-d2MN se replica con una fidelidad del 99,5 %, que solo cae al 99,1 % cuando el tiempo de extensión se reduce a 10 s. La ausencia de amplificación en ausencia de corrección de lectura, junto con la única pequeña disminución observada en presencia de corrección de lectura cuando se reducen los tiempos de extensión, implica que la amplificación del ADN de fidelidad sorprendentemente alta que contiene dTPT3-d2MN da como resultado la extensión selectiva del UBP en relación con emparejamientos erróneos. Esto sugiere que la ausencia de un aceptor de enlaces H orto es más perjudicial para la extensión de un emparejamiento erróneo que para la extensión de los UBP.

45 Esfuerzos hacia la expansión del alfabeto genético. En general, los datos confirman que dTPT3-dNaM es el UBP más prometedor de los analizados. Sin embargo, también son prometedores, los pares formados entre dTPT3 y dFEMO, dFIMO, o dIMO, o entre dFTPT3 y dNaM, dFEMO, dFIMO, o dIMO. Además de los UBP más prometedores mencionados anteriormente, es digno de mención que un número notable de pares nuevos adicionales se replican con solo una fidelidad moderadamente reducida, o se replican con una alta fidelidad cuando la amplificación se realiza en condiciones menos estrictas (Tabla 7). Junto con los UBP replicados más eficientemente, estos pares proporcionan una amplia gama de andamios con diversas propiedades fisicoquímicas para esfuerzos de optimización adicionales, donde se espera que diferentes propiedades fisicoquímicas otorguen diferentes propiedades de tipo farmacocinético a los nucleótidos constituyentes.

55

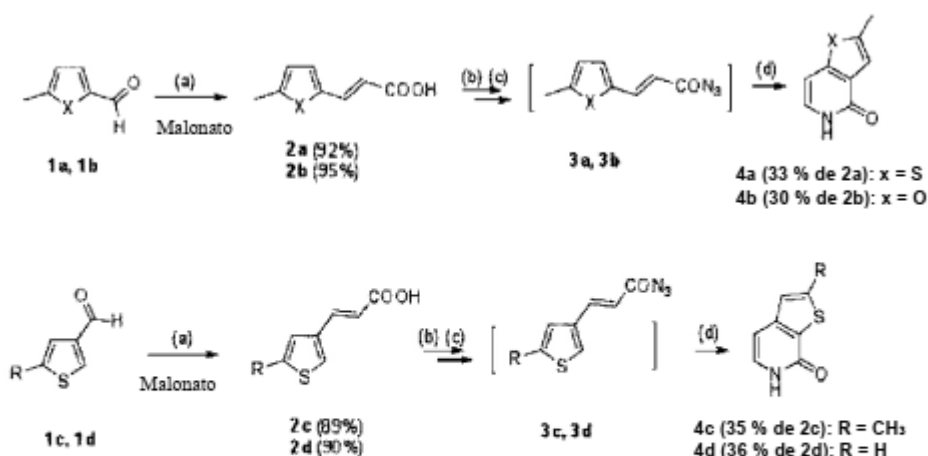
Tabla 7.

dβTP	dαTP	Retención %
SICS	NaM	99 ^a
SICS	FEMO	92 ^b
SNICS	NaM	90 ^a
SNICS	FEMO	95 ^b
SNICS	IMO	88 ^b
TPT3	NMO	89 ^b
TPT3	ZMO	86 ^b
TPT3	CIMO	90 ^b
TPT3	QMO	90 ^b
TPT3	CNMO	91 ^b
FTPT3	NMO	94 ^a
FTPT3	ZMO	88 ^a
FTPT3	CIMO	97 ^a
FTPT3	QMO	87 ^a
FTPT3	CNMO	94 ^a

^aCondiciones de PCR: Plantilla de 100 pg de D8 amplificada durante 16 ciclos con polimerasa Taq en condiciones de termociclado: desnaturalización inicial a 96 °C durante 1 min, 96 °C durante 30 s, 60 °C durante 15 s, 68 °C durante 60 s. ^bCondiciones de PCR: Plantilla de 10 pg de D6 amplificada durante 24 ciclos con polimerasa OneTaq en condiciones de termociclado: desnaturalización inicial a 96 °C durante 1 min, 96 °C durante 5 s, 60 °C durante 5 s, 68 °C durante 10 s.

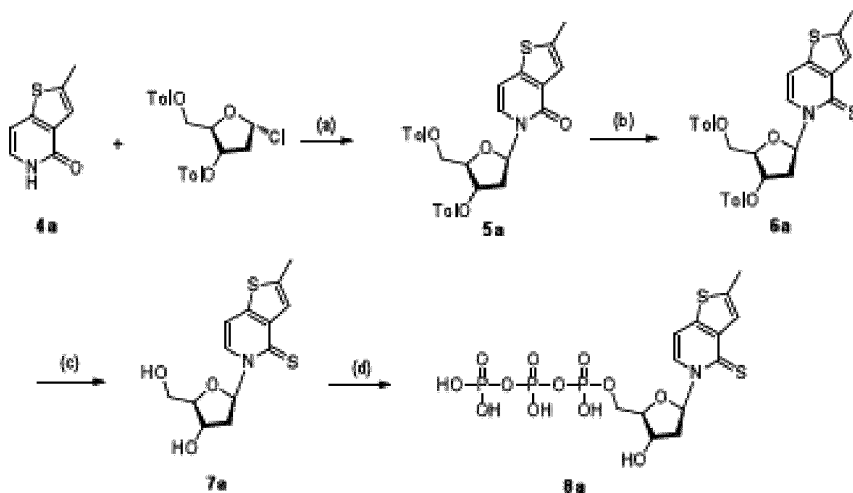
5 Ejemplo 2. Procedimiento general de síntesis de trifosfato.

Se disolvieron la esponja de protones (1,3 equiv) y el derivado de nucleósido libre (1,0 equiv) en trimetil fosfato seco (40 equiv) y se enfriaron a -15 °C en atmósfera de nitrógeno. Se añadió POC13 recién destilado (1,3 equiv) gota a gota y la mezcla resultante se agitó a -10 °C durante 2 h. Se añadieron tributilamina (6,0 equiv) y una solución de pirofosfato de tributilamonio (5,0 eq) en dimetilformamida (0,5 M). Durante 30 minutos, la reacción se dejó calentar lentamente a 0 °C y después se inactivó mediante la adición de Et3NH2CO3 acuoso 0,5 M (TEAB) pH 7,5 (2 vol-equiv.). La mezcla se diluyó dos veces con H2O y se aisló el producto en una columna DEAE Sephadex (GE Healthcare) con un gradiente de elución de TEAB de 0 a 1,2 M, se evaporó y se destiló junto con H2O (3×). Se realizó una purificación adicional por HPLC de fase inversa (C18) (0 - 35 % de CH3CN en TEAB 0,1 M, pH 7,5), (10 % - 31 % de rendimiento).



Esquema S1. (a) Piperidina, Py, 100 °C, 12 h, después reflujo por 1 h; (b) SOC12, DMF, C reflujo, 3 h, (c) NaN3, 1,4 dioxano, H2O, 5 °C, 0,5 h; (d) difenil éter, 230 °C, 1 h.

Los análogos de nucleobase 4a, 4b, 4c y 4d se sintetizaron en base a los procedimientos de la literatura¹. 2 como se muestra en el Esquema S1. Brevemente, la condensación del aldehído (1a-d) con ácido malónico a 100 °C en piridina como disolvente y piperidina como catalizador durante 12 h, seguido de un reflujo durante 1 h, produjo los intermedios de ácido acrílico correspondientes (2a-d). La cloración de estos ácidos con cloruro de tionilo en cloroformo en presencia de DMF proporcionó los cloruros de acilo, los cuales no se purificaron pero podrían usarse directamente en la preparación de las azidas (3a-d). Los compuestos 3a-d se prepararon en una mezcla bifásica de 1,4-dioxano y agua a 5 °C con azida de sodio. Las mezclas crudas de 3a-d en soluciones de CHCl₃ se agregaron en porciones al éter difenílico y se calentaron a 230 °C para dar los isocianatos que experimentaron la posterior ciclación intramolecular a los sistemas bicíclicos 6-5 fusionados 4a-d.

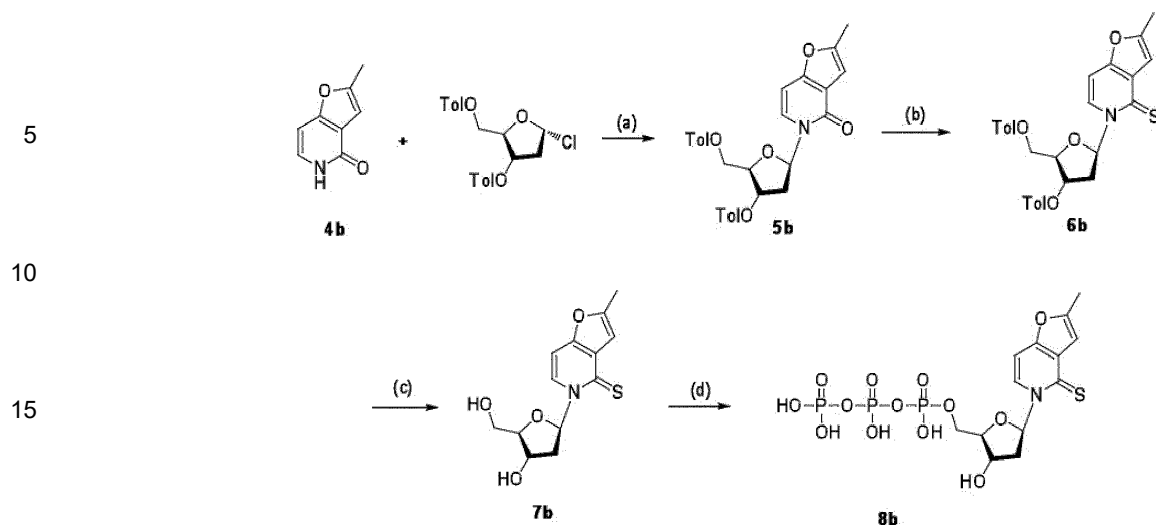


Compuesto 5a. Se añadió bis (trimetilsilil) acetamida (83 mg, 0,39 mmol) a una solución de 4a (54 mg, 0,33 mmol) en CH₂Cl₂ (8 ml) a temperatura ambiente en atmósfera de nitrógeno. Después de agitar durante 40 minutos, se añadió cloruro de 3,5-bis(toluoil)-2-desoxirribosilo (196 mg, 0,39 mmol). La mezcla de reacción se enfrió a 0 °C y se añadió SnCl₄ gota a gota (1,0 M en CH₂Cl₂, 160 µl, 0,16 mmol). La solución se agitó por 2 h a temperatura ambiente. La mezcla de reacción se diluyó con EtOAc, se inactivó con NaHCO₃ acuoso saturado, se extrajo con EtOAc, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 5a como espuma blanca (77 mg, 0,15 mmol, 45 %). ¹H NMR (500 MHz, CDC13) δ 7,97-6,82 (m, 11H, Ar-H), 6,44 (d, J = 7,5 Hz, 1H, H-1'), 5,63 (d, J = 6,5 Hz, 1H, H-3'), 4,76-4,68 (m, 2H, H-5'a, 5'b), 4,59 (d, J = 2,5 Hz, H-4'), 2,89 (dd, J = 1,5, 0,5 Hz, 1H, H-2'a), 2,59 (s, 3H, Ar-CH₃), 2,43 (s, 3H, Ar-CH₃), 2,43 (s, 3H, Ar-CH₃), 2,36-2,30 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13, δ) 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 se calculó para C₂₉H₂₈NO₆S (M+H⁺) 518,1632, se encontró 518,1621.

Compuesto 6a. El compuesto 5a (27 mg, 0,052 mmol) se secó por 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1 ml). Se añadió el reactivo de Lawesson (41,5 mg, 0,10 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración sobre algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 6a como una espuma amarilla (16 mg, 0,03 mmol, 58 %). ¹H NMR (500 MHz, CDC13) δ 8,00-7,89 (m, 4H, Ar-H), 7,71 (m, 1H, Ar-H), 7,49-7,48 (m, 1H, H-1'), 7,29-7,21 (m, 4H, Ar-H), 7,65-7,62 (m, 1H, Ar-H), 6,90 (d, J = 7,5 Hz, 1H, Ar-H), 5,64-5,62 (m, 1H, H-4'), 4,85-4,74 (m, 2H, H-5'a), 4,68-4,67 (m, 1H, H-5'b), 3,38-3,34 (m, 1H, H-3'), 2,26 (s, 3H), 2,44 (s, 3H), 2,41 (s, 3H), 2,28-2,22 (m, 1H). ¹³C NMR (125 MHz, CDC13) δ 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 se calculó para C₂₉H₂₈NO₅S₂ (M+H⁺) 534,1403, se encontró 534,1404.

Compuesto 7a. Se añadió NaOMe al 30 % (8,66 mg, 0,16 mmol) gota a gota a una solución de 6a (20 mg, 0,037 mmol) en metanol (1,0 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. Después, se concentró la mezcla de reacción y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 7a como una espuma amarilla (9,2 mg, 0,031 mmol, 85 %). ¹H NMR (500 MHz, CD₃OD) δ 8,36 (d, J = 4 Hz, 1H, Ar-H), 7,58 (d, J = 1Hz, 1H, Ar-H), 7,35 (t, J = 4 Hz, 1H, H-1'), 7,22 (d, J = 8 Hz, 1H, Ar-H), 4,07-4,06 (m, 1H, H-4'), 4,07 (d, J = 4 Hz, 1H, H-3'), 3,80 (dd, J = 2,4, 4 Hz, 2H, H-5'a, b), 2,79-2,76 (m, 1H, H-2'a), 2,13-2,08 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) δ 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 se calculó para C₁₃H₁₆NO₃S₂ (M+H⁺) 298,0566, se encontró 298,0569.

Compuesto 8a (dTPT1TP). Se sintetizó el compuesto 8a (11,2 mg, 20,8 µmol, 31 %) mediante el uso del Procedimiento general para la síntesis de trifosfato descrita anteriormente a partir de 7a (20 mg, 67,3 µmol). ³¹P NMR (162 MHz, D₂O) δ -10,3 (d, J = 19,8 Hz, yP), -10,9 (d, J = 20,1 Hz, aP), -22,8 (t, J = 19,4 Hz, f3-P). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]⁻ se calculó para C₁₃H₁₇NO₁₂P₃S₂, 536,3, se encontró, 536,7.

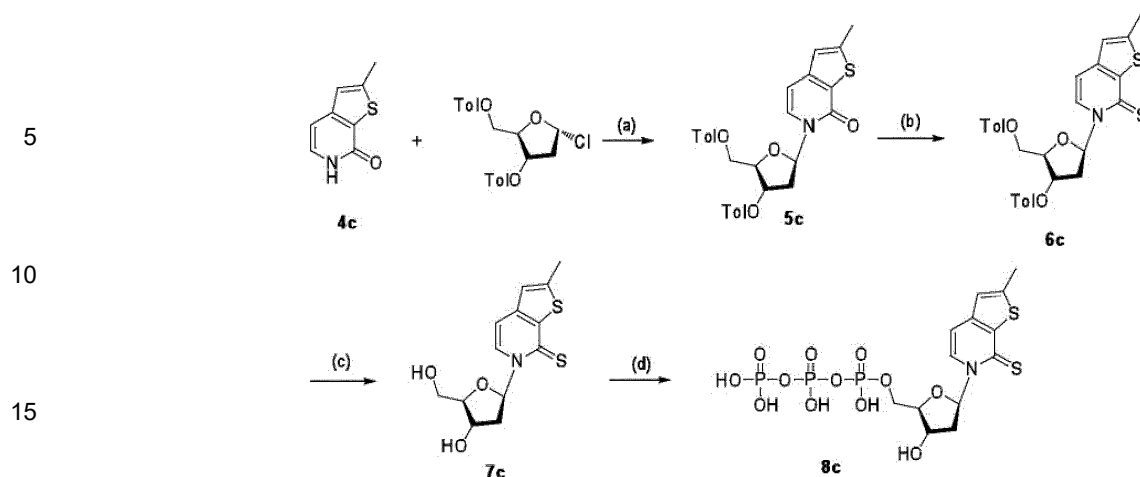


Compuesto 5b. Se añadió bis (trimetilsilil)acetamida (165 mg, 0,81 mmol) a una solución de 4b (100 mg, 0,67 mmol) en CH₂Cl₂ (8 ml) a temperatura ambiente en atmósfera de nitrógeno. Después de agitar durante 40 minutos, se añadió cloruro de 3,5-bis (toluolil)-2-desoxirribosilo (292 mg, 0,81 mmol). La mezcla de reacción se enfrió a 0 °C y se añadió SnCl₄ gota a gota (1,0 M en CH₂Cl₂, 200 µl, 0,2 mmol). La solución se agitó por 2 h a temperatura ambiente. La mezcla de reacción se diluyó con EtOAc, se inactivó con NaHCO₃ acuoso saturado, se extrajo con EtOAc, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 5b como espuma blanca (137 mg, 0,27 mmol, 41 %). ¹H NMR (500 MHz, CDC13) δ 7,99 (d, J = 8,1 Hz, 2H, Ar-H), 7,93 (d, J = 8,1 Hz, 2H, Ar-H), 7,55 (d, J = 7,7 Hz, 1H, Ar-H), 7,33 - 7,28 (m, 2H, Ar-H), 7,27 - 7,20 (m, 2H, Ar-H), 6,82 (dd, J = 8,3, 5,6 Hz, 1H, Ar-H), 6,57 (d, J = 0,9 Hz, 1H, Ar-H), 6,41 (d, J = 7,7 Hz, 1H, H-1'), 5,68 - 5,61 (m, 1H, H-4'), 4,75 (dd, J = 12,1, 3,4 Hz, 2H, H-5'a, b), 4,62 (q, J = 3,1 Hz, 1H, H-3'), 2,94 (ddd, J = 14,3, 5,6, 1,7 Hz, 1H, H-2'a), 2,48 - 2,39 (s, 3x3H, Ar-CH₃), 2,36 - 2,26 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) δ 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 se calculó para C₂₉H₂₈NO₇ (M+H⁺) 502,1860, se encontró 502,1885.

Compuesto 6b. El compuesto 5b (29 mg, 0,056 mmol) se secó por 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1 ml). Se añadió el reactivo de Lawesson (41,5 mg, 0,10 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración sobre algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 6b como una espuma amarilla (15 mg, 0,029 mmol, 52 %). ¹H NMR (500 MHz, CDC13) δ 8,10-7,89 (m, 5H, Ar-H), 7,52-7,48 (m, 1H, H-1'), 7,29-7,22 (m, 4H, Ar-H), 6,8 (d, J = 1Hz, 1H, Ar-H), 6,73 (d, J = 7,5 Hz, 1H, Ar-H), 5,65-5,62 (m, 1H, H-4'), 4,84-4,74 (m, 2H, H-5'a, b), 4,67-4,65 (m, 1H, H-3'), 3,36-3,32 (m, 1H, H-2'a), 2,44 (s, 3H, Ar-CH₃), 2,43 (s, 3H, Ar-CH₃), 2,41 (s, 3H, Ar-CH₃), 2,27-2,21 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) δ 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 se calculó para C₂₉H₂₈NO₆S (M+H⁺) 518,1632, se encontró 518,1638.

Compuesto 7b. Se añadió 30 % NaOMe (8,66 mg, 0,16 mmol) a una solución de 6b (20 mg, 0,039 mmol) en metanol (1,0 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. La mezcla de reacción se concentró después y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 7b como espuma amarilla (9,3 mg, 0,033 mmol, 85 %). ¹H NMR (500 MHz, CD₃OD) δ 8,57 (d, J = 5 Hz, 1H, Ar-H), 7,42 (t, J = 4 Hz, 1H, H-1'), 7,13 (d, J = 7,5 Hz, 1H, Ar-H), 6,80 (s, 1H, Ar-H), 4,50-4,47 (m, 1H, H-4'), 4,12 (d, J = 3,5 Hz, 1H, H-3'), 3,95 (dd, J = 30, 3 Hz, 2H, H-5'a, b), 2,81-2,77 (m, 1H, H-2'a), 2,50 (s, 3H, Ar-CH₃), 2,18-2,14 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) 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+) m/z se calculó para C₁₃H₁₆NO₄S (M+H⁺) 282,0795, se encontró 282,0790.

Compuesto 8b (dFPT1TP). Se sintetizó el compuesto 8b (3,7 mg, 7,1 µmol, 10 %) mediante el uso del Procedimiento general de síntesis de trifosfato descrita anteriormente a partir de 7b (20 mg, 71,2 µmol). ³¹P NMR (162 MHz, D₂O) δ -10,4 (d, J = 20,0 Hz, yP), -10,9 (d, J = 19,4 Hz, aP), -22,8 (t, J = 20,0 Hz, f3-P). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]⁻ se calculó para C₁₃H₁₇NO₁₃P₃S, 520,3, se encontró, 520,1.

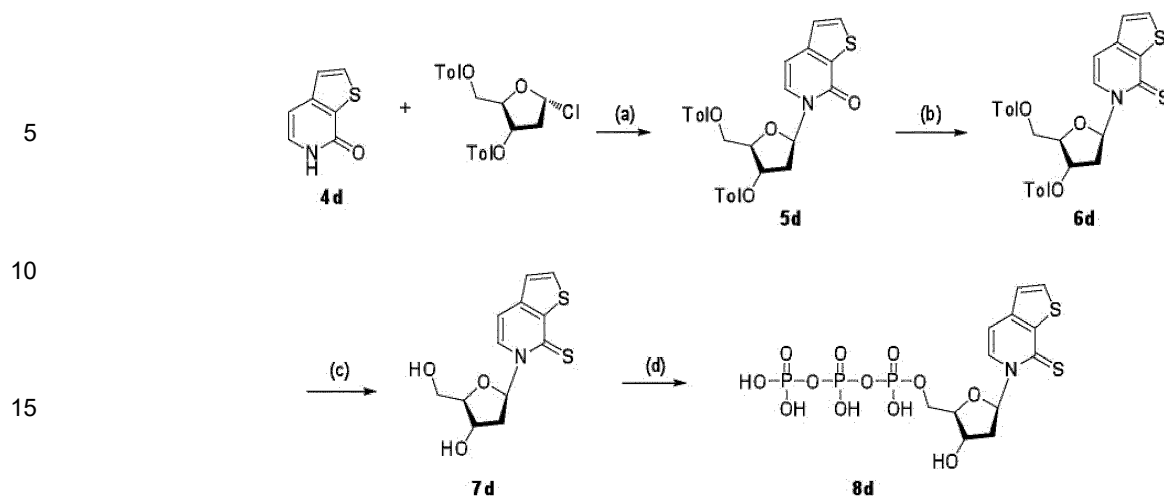


Compuesto 5c. Se añadió bis (trimetilsilil) acetamida (66 mg, 0,33 mmol) a la solución de 4c (46 mg, 0,28 mmol) en CH₂Cl₂ (8 ml) a temperatura ambiente en atmósfera de nitrógeno. Después de agitar durante 40 minutos, se añadió cloruro de 3,5-bis (toluoil)-2-desoxirribosilo (120 mg, 0,33 mmol). La mezcla de reacción se enfrió a 0 °C y se añadió SnCl₄ gota a gota (1,0 M en CH₂Cl₂, 140 µl, 0,14 mmol). La solución se agitó por 2 h a temperatura ambiente. La mezcla de reacción se diluyó con EtOAc, se inactivó con NaHCO₃ acuoso saturado, se extrajo con EtOAc, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 5c como espuma blanca (58 mg, 0,11 mmol, 40 %). ¹H NMR (500 MHz, CDC₁₃) δ 7,98-7,90 (m, 4H, Ar-H), 7,53 (d, J = 7,4 Hz, 1H, Ar-H), 7,27- 7,21 (m, 4H, Ar-H), 6,83-6,82 (m, 2H, Ar-H), 6,44 (d, J = 7,5 Hz, 1H, H-1'), 5,63 (d, J = 6,5 Hz, 1H, H-4'), 4,76 - 4,60 (m, 2H, H-5'a, b), 4,59 (d, J = 2,5 Hz, 1H, H-3'), 2,89 (dd, J = 13, 5,5 Hz, H-2'a), 2,59 (s, 3H, Ar-CH₃), 2,43 (s, 3H, Ar-CH₃), 2,40 (s, 3H, Ar-CH₃), 2,37-2,30 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC₁₃) 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⁺) m/z se calculó para C₂₉H₂₈NO₆S (M+H⁺) 518,1632, se encontró 518,1631.

Compuesto 6c. El compuesto 5c (50 mg, 0,097 mmol) se secó por 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1,5 ml). Se añadió el reactivo de Lawesson (83 mg, 0,20 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración sobre algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 6c como una espuma amarilla (16 mg, 0,03 mmol, 31 %). ¹H NMR (500 MHz, CDC₁₃) δ 8,13-7,97 (m, 5H, Ar-H), 7,52-7,49 (m, 1H, H-1'), 7,37-7,29 (m, 4H, Ar-H), 6,99 (d, J = 1Hz, 1H, Ar-H), 6,91 (d, J = 7,5 Hz, 1H, Ar-H), 5,73-5,71 (m, 1H, H-4'), 4,91-4,82 (m, 2H, H-5'a, b), 4,76-4,74 (m, 1H, H-3'), 3,41-3,37 (m, 1H, H-2'a), 2,68 (s, 3H, Ar-CH₃), 2,52 (s, 3H, Ar-CH₃), 2,49 (s, 3H, Ar-CH₃), 2,39-2,34 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC₁₃) 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⁺) m/z se calculó para C₂₉H₂₈NO₅S₂ (M+H⁺) 534,1403, se encontró 534,1406.

Compuesto 7c. Se añadió 30 % de NaOMe (8,66 mg, 0,16 mmol) gota a gota a una solución de 6c (20 mg, 0,037 mmol) en metanol (1,0 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. Después, se concentró la mezcla de reacción y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 7c como espuma amarilla (8,9 mg, 0,03 mmol, 81 %). ¹H NMR (500 MHz, CD₃OD) δ 8,48, (d, J = 7.5 Hz, 1H, Ar-H), 7,42 (t, J = 5 Hz, 1H, H-1'), 7,20 (d, J = 5 Hz, 1 H, Ar-H), 7,12 (s, 1H, Ar-H), 4,51-4,48 (m, 1H, H-4'), 4,13 (d, J = 5 Hz, 1H, H-3'), 3,95 (dd, J = 30, 5 Hz, 2H, H-5'a, b), 2,81-2,78 (m, 1H, H-2'a), 2,67 (s, 3H, Ar-CH₃), 2,21-2,16 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) 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⁺) m/z se calculó para C₁₃H₁₆NO₃S₂ (M+H⁺) 298,0566, se encontró 298.0566.

Compuesto 8c. El compuesto 7c (10,8 mg, 20,2 µmol, 30 %) se sintetizó mediante el uso del Procedimiento general de síntesis de trifosfato descrita anteriormente a partir de 7c (20 mg, 67,3 µmol). ³¹P NMR (162 MHz, D₂O) δ -10,8 (d, J = 19,8 Hz, yP), -11,5 (d, J = 20,1 Hz, aP), - 23,3 (t, J = 20,1 Hz, f3-P). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]⁻ se calculó para C₁₃H₁₇NO₁₂P₃S₂, 536,3, se encontró, 536,1

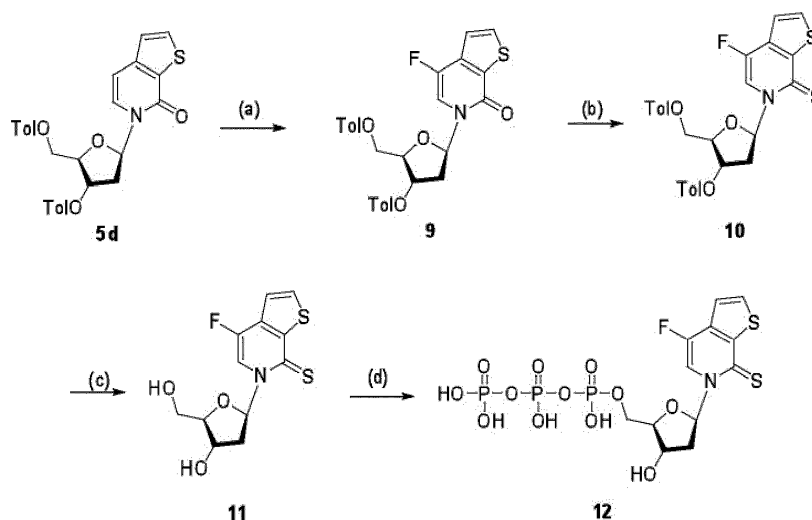


Compuesto 5d. Se añadió bis (trimetilsilil) acetamida (298 mg, 1,46 mmol) a una solución de 4d (200 mg, 1,32 mmol) en CH₂Cl₂ (8 ml) a temperatura ambiente en atmósfera de nitrógeno. Después de agitar durante 40 minutos, se añadió cloruro de 3,5-bis(toluoil)-2-desoxirribosilo (563 mg, 1,46 mmol). La mezcla de reacción se enfrió a 0 °C y se añadió SnC14 gota a gota (1,0 M en CH₂Cl₂, 660 µl, 0,66 mmol). La solución se agitó por 2 h a temperatura ambiente. La mezcla de reacción se diluyó con EtOAc, se inactivó con NaHCO₃ acuoso saturado, se extrajo con EtOAc, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 5d como una espuma blanca (260 mg, 0,52 mmol, 39 %). ¹H NMR (500 MHz, CDC13) δ 7,98-7,90 (m, 4H, Ar-H), 7,70 (d, J = 6 Hz, 1H, Ar-H), 7,55 (d, J = 9,5 Hz, 1H, Ar-H), 7,28-7,16 (m, 5H, Ar-H), 6,84-6,85 (m, 1H, Ar-H), 6,57 (d, J = 9,5 Hz, H-1'), 5,66-5,64 (m, 1H, H-4'), 4,75-4,72 (m, 2H, H-5'a, b), 4,61 (m, 1H, H-3'), 2,95-2,90 (m, 1H, H-2'a), 2,43 (s, 3H, Ar-CH₃), 2,40 (s, 3H, Ar-CH₃), 2,39-2,31(m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) 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+) m/z se calculó para C₂₀H₂₀Cl₂N₂O₅S (M+H+) 504,1475, se encontró 504,1480.

Compuesto 6d. El compuesto 5d (50 mg, 0,1 mmol) se secó por 3 co-evaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1 ml). Se añadió el reactivo de Lawesson (48 mg, 0,12 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración en algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 6d como una espuma amarilla (17 mg, 0,033 mmol, 33 %). ¹H NMR (500 MHz, CDC13) δ 8,14-7,82 (m, 7H, Ar-H), 7,51 (dd, J = 7,5, 6,0 Hz, 1H, H-1'), 7,32-7,23 (m, 5H, Ar-H), 6,99 (d, J = 7,2 Hz, 1H, Ar-H), 7,4-5,73 (m, 1H, H-4'), 4,92-4,83 (m, 2H, H-5'a, b), 4,78-4,77 (m, 1H, H-3'), 3,43-3,40 (m, 1H, H-2'a), 2,51 (s, 3H, Ar-CH₃), 2,48 (s, 3H, Ar-CH₃), 2,39-2,36 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) 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+) m/z se calculó para C₂₈H₂₆NO₅S₂ (M+H+) 520,1247, se encontró 520,1241.

Compuesto 7d. Se añadió 30 % de NaOMe (8,66 mg, 0,16 mmol) gota a gota a una solución de 6d (20 mg, 0,039 mmol) en metanol (1,0 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. Después, la mezcla de reacción se concentró y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 7d como espuma amarilla (9,0 mg, 0,032 mmol, 82 %). ¹H NMR (500 MHz, CD₃OD) δ 8,48 (d, J = 5 Hz, 1H, Ar-H), 8,01 (d, J = 5 Hz, 1H, Ar-H), 7,40-7,38 (m, 2 H, Ar-H), 7,29 (d, J = 10 Hz, 1H, H-1'), 4,47-4,46 (m, 1H, H-4'), 4,10 (m, 1H, H-3'), 3,94-3,88 (m, 2H, H-5'a, b), 2,77-2,76 (m, 1H, H-2'a), 2,19-2,14 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) 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 se calculó para C₁₂H₁₄NO₃S₂ (M+H+) 284,0411, se encontró 284,0410.

Compuesto 8d (dTPT3TP). El compuesto 8d (5,7 mg, 10,9 µmol, 31 %) se sintetizó mediante el uso del Procedimiento general de síntesis de trifosfato descrita anteriormente a partir de 7d (10 mg, 35,3 µmol). ³¹P NMR (162 MHz, D₂O) δ -9,3 (d, J = 19,5 Hz, yP), -10,8 (d, J = 19,8 Hz, aP), -22,4 (t, J = 20,0 Hz, f3-P). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]-se calculó para C₁₂H₁₅NO₁₂P₃S₂-, 521,9, se encontró, 521,9.



20 Compuesto 9. El compuesto 5d (55 mg, 0,11 mmol) se disolvió en 1,0 ml de MeOH-CH₃CN (1: 1 v/v), se añadió selectfluor (42 mg, 0,12 mmol) y se calentó la mezcla a reflujo durante 3 h, después se evaporó el disolvente, se disolvió el residuo en EtOAc (20 ml), se lavó la fase orgánica con agua tres veces. Después, se evaporó el disolvente orgánico, y se secó el residuo sólido por 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en 1 ml de TfOH-CH₂C12 (1: 1 v/v) y la mezcla se agitó a temperatura ambiente durante 1 h, después se concentró la mezcla y el producto crudo se sometió a una cromatografía en columna de gel de sílice (hexano/EtOAc) para proporcionar el compuesto 9 como un sólido blanco (49 mg, 0,093 mmol, 85 %). ¹H NMR (500 MHz, CDC13) δ 7,98-7,92 (m, 4H, Ar-H), 7,75 (d, J = 5 Hz, 1H, Ar-H), 7,52 (d, J = 7,5 Hz, 1H, Ar-H), 7,32-7,21 (m, 5H, Ar-H), 6,82-6,78 (m, 1H, H-1'), 5,64-5,61 (m, 1H, H-4'), 4,80-4,59 (m, 2H, H-5'a, b), 4,62-4,59 (m, 1H, H-3'), 2,93- 2,87 (m, 1H, H-2'a), 2,43 (s, 3H, Ar-CH₃), 2,39 (s, 3H, Ar-CH₃), 2,34-2,27 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) 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. ¹⁹F NMR (376 MHz, CDC13) δ -151,5. HRMS (ESI+) m/z se calculó C₂₈H₂₅FNO₆S (M+H⁺) 522,1381, se encontró con 522.1380.

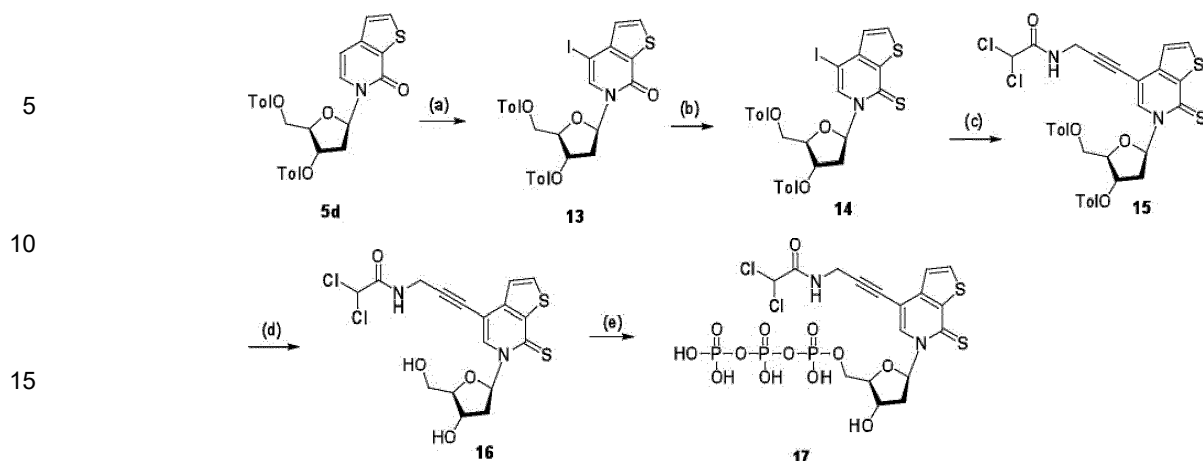
35 Compuesto 10. El compuesto 9 (20 mg, 0,038 mmol) se secó mediante 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1 ml). Se añadió el reactivo de Lawesson (18,5 mg, 0,046 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración sobre algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 10 como una espuma amarilla (6,5 mg, 0,012 mmol, 32 %). ¹H NMR (500 MHz, CDC13) δ 8,11-7,85 (m, 6H, Ar-H), 7,40-7,39 (m, 2H, Ar-H, H-1'), 7,28-7,21 (m, 4H, Ar- H), 5,64-5,63 (m, 1H, H-4'), 4,83 (m, 2H, H-5'a, b), 4,69 (m, 1H, H-3'), 3,34-3,29 (m, 1H, H-2'a), 2,44 (s, 1H, Ar-CH₃), 2,40 (s, 3H, Ar-CH₃), 2,30-2,26 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC13) 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. ¹⁹F NMR (376 MHz, CDC13) δ -142,9. HRMS (ESI+) m/z se calculó para C₂₈H₂₅FNO₅S₂ (M+H⁺) 538,1153, se encontró 538,1155.

45 Compuesto 11. Se añadió 30 % de NaOMe (4,33 mg, 0,08 mmol) gota a gota a una solución de 10 (10 mg, 0,019 mmol) en metanol (1,5 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. Después, se concentró la mezcla de reacción y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂C12) para proporcionar el compuesto 11 como espuma amarilla (4,9 mg, 0,016 mmol, 85 %). ¹H NMR (500 MHz, CD₃OD) δ 8,68 (d, J = 5 Hz, 1H, Ar-H), 8,12 (d, J = 5 Hz, 1H, Ar-H), 7,52 (d, J = 5 Hz, 1H, Ar-H), 7,28 (t, J = 6,5 Hz, 1H, H-1'), 4,48 (m, 1H, H-4'), 4,10 (m, 1H, H-3'), 3,94 (dd, J = 35, 3 Hz, 2H, H-5'a, b), 2,78-2,75 (m, 1H, H-2'a), 2,24-2,19 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) δ 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. ¹⁹F NMR (376 MHz, CD₃OD) δ -145,3. HRMS (ESI+) m/z se calculó para C₁₂H₁₃FNO₃S₂ (M+H⁺) 302,0315, se encontró 302,0314.

55 Compuesto 12 (dFTPT3TP). El compuesto 12 (2,0 mg, 3,7 μ mol, 22 %) se sintetizó mediante el uso del Procedimiento general de síntesis de trifosfato descrita anteriormente a partir de 11 (5 mg, 16,6 μ mol). ³¹P NMR (162 MHz, D₂O) δ -10,9 (d, J = 20,0 Hz, yP), -11,6 (d, J = 21,1 Hz, aP), -23,3 (t, J = 23,1 Hz, f3-P). ¹⁹F NMR (376 MHz, D₂O) δ -138,5 (s). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]⁻ se calculó para C₁₂H₁₄FNO₁₂P₃S₂-, 539,9, se encontró, 540,1.

60

65



Compuesto 13. Se añadió monoyoduro de yodo (1,0 M en CH₂Cl₂, 0,15 ml, 0,15 mmol) gota a gota a una solución de 5d (73 mg, 0,145 mmol) en CH₂Cl₂ (1 ml) a 0 °C en atmósfera de nitrógeno. La mezcla resultante se agitó a temperatura ambiente durante una noche. La mezcla de reacción se inactivó con NaHCO₃ acuoso saturado y Na₂S₂O₃ acuoso saturado, se extrajo con CH₂Cl₂, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 13 como una espuma blanca (57 mg, 0,091 mmol, 63 %). ¹H NMR (500 MHz, CDC₁₃) δ 7,98-7,93 (m, 4H, Ar-H), 7,83 (s, 1H, Ar-H), 7,72 (d, J = 5 Hz, 1H, Ar-H), 7,28-7,17 (m, 5H, Ar-H), 6,78-6,75 (m, 1H, H-1'), 5,65-5,63 (m, 1H, H-4'), 4,76 (m, 2H, H-5'a, b), 4,63-4,62 (m, 1H, H-3'), 2,95-2,91 (m, 1H, H-2'a), 2,43 (s, 3H, Ar-CH₃), 2,35 (s, 3H, Ar-CH₃), 2,34-2,29 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC₁₃) δ 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 se calculó para C₂₈H₂₅INO₅S (M+H⁺) 630,0442, se encontró 630,0440.

Compuesto 14. El compuesto 13 (30 mg, 0,048 mmol) se secó mediante 3 coevaporaciones con tolueno anhidro. El residuo se disolvió en tolueno anhidro (1 ml), se añadió el reactivo de Lawesson (23 mg, 0,057 mmol) y se calentó la mezcla durante una noche a reflujo. Después de la filtración sobre algodón, se concentró el filtrado y el producto crudo se sometió a una cromatografía en columna de gel de sílice (Hexano/EtOAc) para proporcionar el compuesto 14 como una espuma amarilla (8,4 mg, 0,013 mmol, 27 %). ¹H NMR (500 MHz, CDC₁₃) δ 8,31 (s, 1H, Ar-H), 7,99-7,82 (m, 5H, Ar-H), 7,39-7,36 (m, 1H, H-1'), 7,29-7,20 (m, 5H, Ar-H), 5,65-5,64 (m, 1H, H-4'), 4,83-4,81 (m, 2H, H-5'a, b), 4,71-4,70 (m, 1H, H-3'), 3,35 (dd, J = 15, 5,5 Hz, 1H, H-2'a), 2,44 (s, 3H, Ar-CH₃), 2,39 (s, 3H, Ar-CH₃), 2,27-2,21 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC₁₃) δ 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 se calculó para C₂₈H₂₅INO₅S₂ (M+H⁺) 646,0213, se encontró 646,0219.

Compuesto 15. Se añadió (PPh₃)₄Pd (1,7 mg, 0,0015 mmol), CuI (0,57 mg, 0,011 mmol) y Et₃N (5 µl, 0,030 mmol) a una solución de 14 (10 mg, 0,015 mmol) en DMF (2 ml) en atmósfera de nitrógeno. La mezcla de reacción se desgasificó y se añadió una solución de C₁₂CHCONHCH₂CCH (3,8 mg, 0,0225 mmol) en DMF (0,5 ml). La mezcla de reacción se agitó una noche a temperatura ambiente y se controló mediante TLC. La mezcla de reacción se diluyó con EtOAc, se inactivó con NaHCO₃ acuoso saturado, se extrajo con EtOAc, se secó, se filtró y se evaporó. El producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 15 como espuma amarilla (9,2 mg, 0,0135 mmol, 91 %). ¹H NMR (500 MHz, CDC₁₃) δ 8,26 (s, 1H, Ar-H), 7,99-7,82 (m, 5H, Ar-H), 7,40-7,37 (m, 2H, Ar-H, H-1'), 7,29-7,21 (m, 4 H, Ar-H), 6,71 (br, 1H, NH), 6,95 (s, 1H, CHC₁₂), 5,65-5,64 (m, 1H, H-4'), 4,85-4,79 (m, 2H, H-5'a, b), 4,73 (m, 1H, H-3'), 4,26-4,11 (m, 2H, NHCH₂), 3,38-3,34 (m, 1H, H-2'a), 2,44 (s, 3H, Ar-CH₃), 2,40 (s, 3H, Ar-CH₃), 2,31-2,25 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CDC₁₃) δ 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 se calculó para C₃₃H₂₉Cl₂N₂O₆S₂ (M+H⁺) 683,0839, se encontró 683,0854.

Compuesto 16. Se añadió NaOMe al 30 % (2,92 mg, 0,32 mmol) gota a gota a una solución de 15 (9,2 mg, 0,0135 mmol) en metanol (1,0 ml). La mezcla de reacción se agitó por 1 h a temperatura ambiente y se controló mediante TLC. La mezcla de reacción se concentró y el producto crudo se sometió a cromatografía en columna de gel de sílice (MeOH/CH₂Cl₂) para proporcionar el compuesto 16 como espuma amarilla (4,5 mg, 0,01 mmol, 74 %). ¹H NMR (500 MHz, CD₃OD) δ 8,69 (s, 1H, Ar-H), 8,06 (d, J = 5 Hz, 1H, Ar-H), 7,53 (d, J = 5 Hz, 1H, Ar-H), 7,30 (t, J = 5 Hz, 1H, H-1'), 6,33 (s, 1H, CHC₁₂), 4,47-4,46 (m, 1H, H-4'), 4,36 (s, 2H, NHCH₂), 4,11-4,08 (m, 1H, H-3'), 3,97 (dd, J = 12, 3 Hz, 2H, H-5'a, b), 2,79-2,74 (m, 1H, H-2'a), 2,21-2,16 (m, 1H, H-2'b). ¹³C NMR (125 MHz, CD₃OD) δ 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 se calculó para C₁₇H₁₇Cl₂N₂O₄S₂ (M+H⁺) 447,0001, se encontró 447,0020.

Compuesto 17 (dTPT³ATP). El compuesto 17 (2,2 mg, 3,1 µmol, 28 %) se sintetizó mediante el uso del Procedimiento general de síntesis de trifosfato descrita anteriormente a partir de 16 (5 mg, 11,2 µmol). ³¹P NMR (162 MHz, D₂O) δ -

10,85 (d, J = 19,9 Hz, y-P), -11,63 (d, J = 20,0 Hz, a-P), -23,07 (s), -23,26 (t, J = 19,7 Hz, f3-P). MS (MALDI-TOF-, matriz: 9-aminoacridina) (m/z): [M-H]- se calculó para C₁₇H₁₈C₁₂N₂O₁₃P₃S₂-, 684,9, se encontró, 685,0.

Ejemplo 3. Procedimiento general para el ensayo de amplificación por PCR para determinar la fidelidad.

5

Materiales. Las ADN polimerasas Taq y OneTaq se adquirieron de New England Biolabs (Ipswich, MA). Se adquirió una mezcla de dNTP de Fermentas (Glen Burnie, MD). La tinción de gel de ácido nucleico SYBR Green I (10.000 ×) se adquirió de Life Technologies (Carlsbad, CA).

10

DNA oligonucleotides. En la Tabla 8 se divulgan las secuencias de oligonucleótidos completas. Los cebadores completamente naturales se adquirieron de Intergrated DNA Technologies (Coralville, Iowa). Los reactivos y solventes para la síntesis de cebadores no naturales 1-3 se obtuvieron de Glen Research (Sterling, VA) y/o Applied Biosystems (Foster City, CA). Los oligonucleótidos se prepararon mediante el uso de síntesis de ADN automatizada estándar con fosforamidas naturales ultra suaves (Glen Research) y fosforamida dNaM (Berry & Associates, Inc., Dexter, MI) en soportes de vidrio de poro controlado (0,20 μmol, 1.000 Å, Glen Research) y un sintetizador ABI Expedite 8905. Después de la síntesis automatizada, los oligonucleótidos se escindieron del soporte, se desprotegieron por incubación en conc. de amoníaco acuoso durante una noche a temperatura ambiente, se purificaron mediante purificación DMT (cartucho glen-pak™, Glen Research) y se desalaron sobre Sephadex G-25 (columnas NAP-25, GE Healthcare). La concentración de oligonucleótidos monocatenarios se determinó mediante absorción UV a 260 nm.

20

Ensayo PCR. Las amplificaciones por PCR se realizaron en un volumen total de 25 μl y con condiciones específicas para cada ensayo como se describe en la Tabla 9. Después de la amplificación, se analizó una alícuota de 5 μl en un gel PAGE no desnaturalizante al 6 % junto con una escalera de 50 pb (Life Technologies) para confirmar el tamaño del amplicón. La solución restante se purificó por columna de centrifugación (DNA Clean y Concentrator-5; Zymo Research, Irvine, CA), seguido de gel de agarosa al 4 %, recuperado con el Kit de recuperación de ADN de gel Zymoclean (Zymo Research), se cuantificó por unión de tinte fluorescente (kit de ensayo Quant-iT dsDNA HS, Life Technologies), y se secuenció en un analizador de ADN 3730 (Applied Biosystems). La fidelidad se determinó como el % de retención promedio del par de bases no natural mediante la duplicación como se describe a continuación.

25

30

Determinación de la fidelidad. El por ciento de retención de un par de bases no natural (F) se midió mediante el uso de secuenciación de datos sin procesar y se normalizó a fidelidades por duplicación. Brevemente, la presencia de un nucleótido no natural conduce a una terminación brusca del perfil de secuenciación, mientras que la mutación a un nucleótido natural da como resultado una "lectura completa". Por lo tanto, la extensión de la "lectura completa", se correlaciona inversamente con la retención del par de bases no natural. Para usar los datos de secuenciación como una medida cuantitativa de la fidelidad de la PCR, realizamos experimentos de calibración en el rango de 50-100 % de retención del par de bases no natural. Por lo tanto, la retención baja (< 50 %) y la "lectura" alta hacen que la cuantificación sea inexacta.

35

40

Se realizó la cuantificación de la retención alta (> 50 %) mediante el ajuste de los puntos de inicio y parada para el programa de computación Sequencing Analysis (Applied Biosystems) y después se determinó la intensidad de señal promedio individualmente para cada canal (A, C, G y T) para los picos dentro de aquellos puntos (35-45 nucleótidos de longitud) definidos antes (sección L) y después de S36 (sección R) del nucleótido no natural. La relación R/L se normalizó mediante el uso de gráficos de calibración de la secuenciación para tener en cuenta tanto el ruido en los cromatogramas de secuenciación como la lectura en las muestras de control. La relación R/L de después de la normalización (R/L_{norm}) se corresponde al porcentaje de las secuencias naturales en el grupo. Finalmente, se calculó F como la norma 1 - (R/L) y se calculó la retención del par de bases no natural por duplicación (fidelidad, f) como $1/(F^{\log_2 A})$, donde A es una amplificación y $\log_2 A$ es el número de duplicaciones. Se secuenció cada muestra anterior y la amplificación por PCR por triplicado en cada dirección para minimizar el error de secuenciación. Los datos correspondientes se divulgan en la Tabla 10. En condiciones de PCR estándar, OneTaq amplificó el ADN que contiene dTPT3-dNaM con una eficiencia que es solo 4 veces menor que la del ADN que contiene solo los pares de bases naturales, y con una fidelidad superior al 99,98 %. Esta fidelidad corresponde a una tasa de error de 10^{-4} por nucleótido, que se superpone con la tasa de error de ADN completamente natural de 10^{-4} a 10^{-7} con sistemas de PCR de uso común. Con la polimerasa Taq, la eficiencia es solo 2,5 veces menor que la de un par de bases natural, y la fidelidad es del 99,7 %. Esta fidelidad corresponde a una tasa de error de 10^{-3} , que es similar a la observada con la amplificación del ADN natural mediada por Taq.

50

55

ES 2 773 547 T3

Tabla 8. Secuencias de ADN.

Nombre	Secuencia (5' a 3')	Anotaciones
	Regiones del cebador subrayadas	
Fend1	<u>CACACAGGAAACAGCTATGAC</u>	Cebadores para PCR (plantillas D6 y 134mer)
Fend2	<u>GAAATTAATACGACTCACTATAGG</u>	
Fend1-poly- dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTT <u>CACACAG</u> <u>GAAACAGCTATGAC</u>	Cebadores para la secuenciación Sanger (plantillas D6 y 134mer)
Fend2-poly- dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTT <u>GAAAT</u> <u>TAATACGACTCACTATAGG</u>	
D6	<u>CACACAGGAAACAGCTATGACCCGGGTTATTAC</u> <u>ATGCGCTAGCACTTGGAATTCACCAGACGNNN</u> <u>NaM NNNCGGGACCCATAGTAAATCTCCTTCTT</u> <u>AAAGTTAAGCTTAACCCTATAGTGAGTCGTATT</u> <u>AATTTC</u>	N = nucleótido natural aleatorizado
134mer	<u>CACACAGGAAACAGCTATGACCCGGGTTATTAC</u> <u>ATGCGCTAGCACTTGGAATTCACAATACT NaM</u> <u>TCTTTAAGGAAACCATAGTAAATCTCCTTCTT</u> <u>AAAGTTAAGCTTAACCCTATAGTGAGTCGTATT</u> <u>AATTTC</u>	
Cebador 1	<u>NaM CCTGCGTCAATGTAATGTTT</u>	Cebadores para PCR con Temp1-3
Cebador2	<u>TTCACGGT NaM AGCACGCATAGG</u>	
Cebador3	<u>CCAATGTACC NaM TGCGTATGTTT</u>	
Cebador-rev	<u>CCCTGCGTTTATCTGCTCTC</u>	
Temp1	<u>CCCTGCGTTTATCTGCTCTCTCGGTCGTTTCGGC</u> <u>TGCGGCGGAAACATTACATTGACGCAGG</u>	Los nucleótidos mostrados en negrita forman un par erróneo con dNaM en la primera ronda de PCR
Temp2	<u>CCCTGCGTTTATCTGCTCTCTCGGTCGTTTCGGC</u> <u>TGCGGCGCCTATGCGTGCTTACCGTGAA</u>	
Temp3	<u>CCCTGCGTTTATCTGCTCTCTCGGTCGTTTCGGC</u> <u>TGCCGGAACATACGCATGGTACATTGG</u>	

Tabla 9. Condiciones de la PCR.

	OneTaq	Taq	PCR para el marcaje con biotina (par de bases no natural ubicada centralmente)	PCR para el marcaje con biotina (par de bases no natural ubicada en las posiciones 1, 9, 11)
Tampón	OneTaq 1 x	Taq 1 x	OneTaq 1 x	OneTaq 1 x
Enzima, U/ μ l	OneTaq, 0,02	Taq, 0,02	OneTaq, 0,02	OneTaq, 0,02
Plantilla	D6 (0,01 ng)	D6 (0,01 ng)	134mer (0,5 ng)	60mer (0,5 ng)
dNTPs, μ M	200	200	200	200
dNaMTP, μ M	100	100	100	100
dXTP, μ M	100	100	100 de d5SICS ^{PA} TP o dTPT3 ^{PA} TP	100 de d5SICS ^{PA} TP o dTPT3 ^{PA} TP
Mg ²⁺ , mM	3	3	3	3
Cebadores, μ M	1	1	1	1
SYBR Green I	0.5 x	0.5 x	0.5 x	0.5 x
Condiciones térmicas				
Desnaturalización inicial	-	-	96 °C, 1 min	96 °C, 1 min
Desnaturalización	96 °C, 10 s	96 °C, 10 s	96 °C, 15 s	96 °C, 15 s
Hibridación	60 °C, 15 s	60 °C, 15 s	60 °C, 30 s	64 °C, 30 s
Extensión	68 °C, 60 s	68 °C, 15 s	68 °C, 2 min	68 °C, 2 min
# de ciclos	16+16+16	20	12	12

Tabla 10.

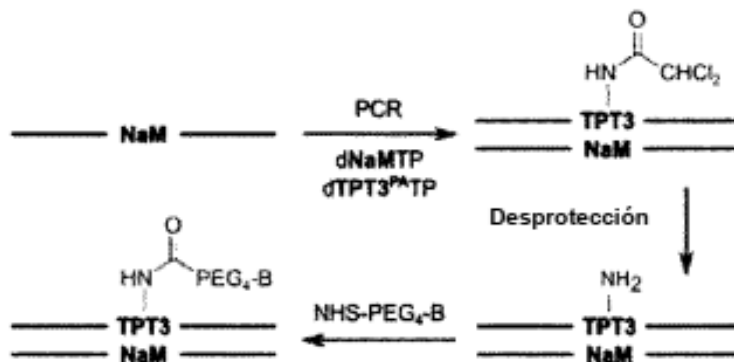
	OneTaq PCR (48 ciclos)			Taq PCR (20 ciclos)		
	amplificación x 10 ¹²	Retención, %	fidelidad, %	amplificación x 10 ¹³	Retención, %	fidelidad, %
5SICS	9,4	96,3 \pm 1,7	99,91 \pm 0,04	7,7	86,7 \pm 1,0	98,90 \pm 0,01
TPT3	12,9	> 99	> 99,98	11,7	95,6 \pm 1,7	99,66 \pm 0,13
TPT3 ^{PA}	4,7	98,6 \pm 1,2	99,97 \pm 0,03	3,5	85 \pm 4	98,7 \pm 0,4
5SICS ^{PA}	9,2	45 \pm 2	98,16 \pm 0,12	6,4	^a	^a

^aPar de bases no naturales perdida durante la amplificación

Ejemplo 4: Marcaje específico del sitio de TPT3: análisis mediante el ensayo de cambio de gel de estreptavidina.

Se sintetizó un ADN de 134 mer que comprende dTPT3 en posición central^{PA}-dNaM o d5SICS^{PA}-dNaM. Las plantillas de ADN se amplificaron por PCR en las condiciones descritas en la Tabla 9. Tras finalizar, se añadió NaOH (1 M, 12,5 μ l) directamente a las muestras de PCR a una concentración final de 0,2 M y se incubó durante 5 horas a temperatura ambiente. Después de la adición de NaOAc (3 M, pH 5,5, 7,5 μ l) y 200 μ l de etanol frío, las muestras se mezclaron, se incubaron en hielo durante una noche, y el ADN se precipitó por centrifugación a 10,000 rfu durante 30 minutos a 4 °C. Se extrajo el sobrenadante y se lavaron los gránulos cuidadosamente con etanol al 80 %. Las muestras se resuspendieron en 50 μ l del tampón de hibridación (fosfato de sodio 50 mM, pH 7,5, NaCl 100 mM, EDTA 1 mM), se calentaron a 95 °C y se enfriaron a temperatura ambiente durante 30 minutos. Se mezcló la solución de NHS-PEG₄-biotina (Thermo Scientific) en el tampón de hibridación (40 mM, 50 μ l) con las muestras de ADN y se incubó durante una noche a temperatura ambiente. Las muestras se purificaron mediante columna de centrifugación (DNA Clean and Concentrator-5, Zymo Research) y se eluyeron en 10 μ l del tampón de elución. Se mezcló la mitad de la muestra (5 μ l)

con 1 μg de estreptavidina (Promega) en tampón de hibridación, se incubó durante 30 minutos a 37 °C, se mezcló con tampón de carga no desnaturante 5 x (Qiagen) y se cargó en PAGE 6 % sin desnaturar. La mitad restante se mezcló con tampón de carga no desnaturante 5 x y se cargó directamente en el gel como un control. Después de correr el gel a 110 V durante 30 minutos, el gel se sumergió en Sybr Gold Nucleic Acid Stain (Life Technologies) 1 x durante 30 minutos y se visualizó mediante el uso de un Molecular Imager Gel Doc XR+ equipado con filtro 520DF30 (Bio-Rad). A continuación se muestra un esquema de la estrategia de marcaje descrita.



Ejemplo 5. Procedimiento general de transcripción de un par de bases no natural.

Para caracterizar la transcripción de los pares de bases no naturales formados por dTPT3 y dNaM, o análogos o derivados de los mismos (en los que los derivados incluyen fracciones enlazadoras), se sintetizan los ribonucleótidos y los desoxinucleótidos y se convierten en los trifosfatos o desoxifosforamiditos correspondientes, y los desoxifosforamiditos se incorporan a las plantillas de ADN mediante el uso de la síntesis automatizada de ADN. Los experimentos de transcripción se llevan a cabo con sustrato de ADN 100 nM, tampón Takara 1 x (Tris-HCl 40 mM, pH 8,0, MgCl₂ 8 mM, espermidina 2 mM), agua estéril tratado con DEPC y libre de nucleasas (Fisher), polimerasa T7 (50 unidades), NTP natural 20 μM cada una, α -³²P-ATP (2,5 μCi , MP Biomedicals) y TPT3TP 5 μM o NamTP 5 μM . Después de la incubación durante 2 horas a 37 °C, se inactiva la reacción mediante la adición de 10 μl de solución de carga de gel (urea 10 M, azul de bromofenol al 0,05 %), y la mezcla de reacción se carga en un gel de poliacrilamida al 20 % con urea 7 M, se somete a electroforesis, y se analiza mediante fosforimagen. La eficiencia de la transcripción se examina tras medir (con un porcentaje de conversión bajo) la cantidad de producto completo formado en función del tiempo.

Ejemplo 6. Procedimiento general para el análisis termodinámico de un ADN dúplex que comprende un par de bases no natural.

Los experimentos de fusión UV se llevan a cabo mediante el uso de un espectrofotómetro Cary 300 Bio UV-visible. La absorbancia de una muestra (3 μl de oligonucleótido que comprende un par de bases no natural, tampón PIPES 10 mM, pH 7,0, NaCl 100 mM, MgCl₂ 10 mM) se controla a 260 nm de 21 °C a 80 °C a una velocidad de calentamiento de 0,5 °C por minuto. Las temperaturas de fusión se determinan mediante el procedimiento derivativo mediante el uso del programa de aplicación térmica Cary Win UV.

Los parámetros termodinámicos se determinan por el análisis de van't Hoff $T_m^{-1} = R[\text{En}([C_T]/4)]\Delta H + \Delta S^\circ/\Delta H^\circ$, donde ΔH° y ΔS° son los cambios estándar de entalpía y entropía determinados a partir de experimentos UV, respectivamente, R es la constante de gas universal y $[C_T]$ es la concentración total de la cadena oligonucleotídica. Los cambios en el número de moléculas de agua asociadas con el proceso de fusión, Δn_w , se obtienen de la dependencia de T_m sobre la actividad del agua (a_w) de acuerdo con la ecuación $\Delta n_w = (-\Delta H/R)[\delta(T_m^{-1})/\delta(\text{En } a_w)]$. La pendiente del gráfico de temperatura de fusión recíproca (K^{-1}) contra el logaritmo de la actividad del agua a diferentes concentraciones (0, 2, 5, 7, 10, 12 y 15 % en peso) de etilenglicol se toma como el valor de $\delta(T_m^{-1})/\delta(\text{En } a_w)$.

Los experimentos de CD se llevan a cabo con un espectropolarímetro Aviv modelo 61 DS equipado con una unidad de control de temperatura termoeléctrica Peltier (concentración de oligonucleótidos 3 μM , tampón PIPES 10 mM, pH 7,0, NaCl 100 mM, MgCl₂ 10 mM). Los datos se recopilan mediante el uso de una cubeta de cuarzo de 1 cm de longitud de recorrido con escaneo de 360 a 220 nm, una constante de tiempo de 3 segundos y un tamaño de paso de longitud de onda de 0,5 nm a 25 °C.

Ejemplo 7. Selección in vitro con nucleobases no naturales.

Se genera una biblioteca de oligonucleótidos que comprende ácidos nucleicos no naturales. Una muestra de la biblioteca se somete a unión secuencial y elución de una molécula objetivo, por ejemplo, una proteína. El conjunto de ácidos nucleicos de unión se amplifica por PCR y se somete a otra ronda de selección para unirse a la molécula objetivo. Este proceso de selección se repite varias veces. Para aumentar la presión de selección, en las últimas rondas

de selección, se reduce la concentración de la molécula objetivo y/o el tiempo de incubación. Los ácidos nucleicos sobrevivientes se secuencian como aptámeros potenciales. Las afinidades de unión de los aptámeros potenciales se determinan mediante el uso de citometría de flujo.

5 **Ejemplo 8. Procedimiento general de reacción de clic de ADN.**

10 Se añadió 1 µl de azida-PEG(3+3)-SS-Biotina (20 mM en H₂O), a una solución de ADN (0,2 pmol) en 14 µl de DMSO, seguido de 2 µl de ligando (BimC₄A)₃(4 mM en H₂O), 1 µl de ascorbato de sodio (100 mM en H₂O), y 1 µl de tampón PBS (5 ×), la mezcla se agitó en un vórtex y como último componente, se añadió una solución de 1 µl de CuSO₄ recién preparada (4 mM en H₂O). Se agitó la solución durante 2 a 37 °C, y después se purificó el ADN del producto resultante (kit DNA Clean & Concentrator-5, Zymo Research Corp.). Las muestras purificadas se usaron directamente para ensayos de movilidad en gel (véase más abajo).

15 **Ejemplo 9. Procedimientos generales para el marcaje de ADN post-amplificación (De Seo y otros, JACS 2011,133,19878).**

20 Para el marcaje de la síntesis posenzimática, se incubó ADNd con un grupo amino libre con EZ-Link sulfo-NHS-SS-biotina 10 mM o EZ-Link NHS-PEG₄-biotina (Thermo Scientific) durante 1 hora a temperatura ambiente en tampón marcado con fosfato (fosfato de sodio 50 mM, pH 7,5, NaCl 150 mM, EDTA 1 mM), y después se purificó mediante el uso del kit de purificación Qiagen PCR. Con derivados de amina protegidos con dicloroacetilo como dTPT3^{PA} o d5SICS^{PA}, la amina requirió primero la desprotección, que se realizó mediante incubación durante una noche en una solución acuosa concentrada de amoníaco a temperatura ambiente. El amoníaco se elimina mediante un concentrador SpeedVac (aspirador de agua seguido de una bomba de vacío de aceite). Para escindir los enlaces que contienen disulfuro (es decir SS-biotina o SS-PEG₄-biotina), el ADNd se trató con DTT (concentración final de 30 mM) durante 1 hora a 37 °C. Para el marcaje de la cadena principal, se incubó ADNd con un fosforotioato de cadena principal con EZ-Link yodoacetil-PEG 25 mM₂-biotina (Thermo Scientific) en tampón fosfato de marcaje durante una noche a 50 °C, y los productos se purificaron con Qiagen PCR Purification Kit. Todas las reacciones que manipulan fracciones de biotina unidas se cuantificaron mediante ensayos de cambio de gel de estreptavidina.

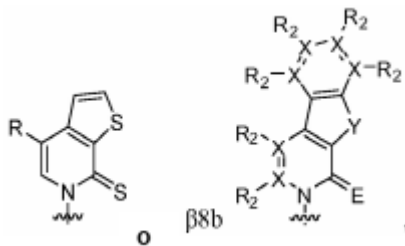
30 *Ensayos de movilidad en gel.* Las muestras de ADN (10-50 ng) se mezclaron con 1 µg de estreptavidina (Promega) en tampón fosfato de marcaje (fosfato de sodio 50 mM, pH 7,5, NaCl 150 mM, EDTA 1 mM), se incubaron durante 30 minutos a 37 °C, se mezclaron con tampón de carga no desnaturizante 5 × (Qiagen), y se cargado en PAGE no desnaturizante al 6 %. El gel se corrió a 150 V durante 25-40 min, después se tiñó con Sybr Gold Nucleic Acid Stain 1 × (Life Technologies) en TBE durante 30 min y se visualizó mediante el uso de un Molecular Imager Gel Doc XR+ equipado con un filtro 520DF30 (Bio-Rad). Se observaron bandas fuertes correspondientes a ADNd (en ~150 pb) y el complejo 1: 1 entre ADNd y estreptavidina (en ~400 pb). También se observaron bandas débiles correspondientes a complejos de ADN de alto orden (migración más lenta) y estreptavidina o de ADN monocatenario, no biotinilado resultante de la hibridación incompleta después de la PCR en algunos casos.

40 Son posibles diversas modificaciones equivalentes dentro del alcance de las realizaciones reivindicadas. Así, debe entenderse que aunque las presentes realizaciones se divulgan específicamente mediante realizaciones preferentes y características opcionales, los expertos en la materia pueden recurrir a la modificación y variación equivalente de los conceptos divulgados en la presente memoria, y que tales modificaciones y variaciones equivalentes se consideran dentro del alcance de esta divulgación como se define en las reivindicaciones adjuntas.

45

REIVINDICACIONES

1. Un compuesto que comprende un análogo de nucleobase de cualquiera de las fórmulas siguientes:



en las que R es hidrógeno o es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o R es un enlazador acoplado al cual se une una carga;

en las que cada X es carbono,

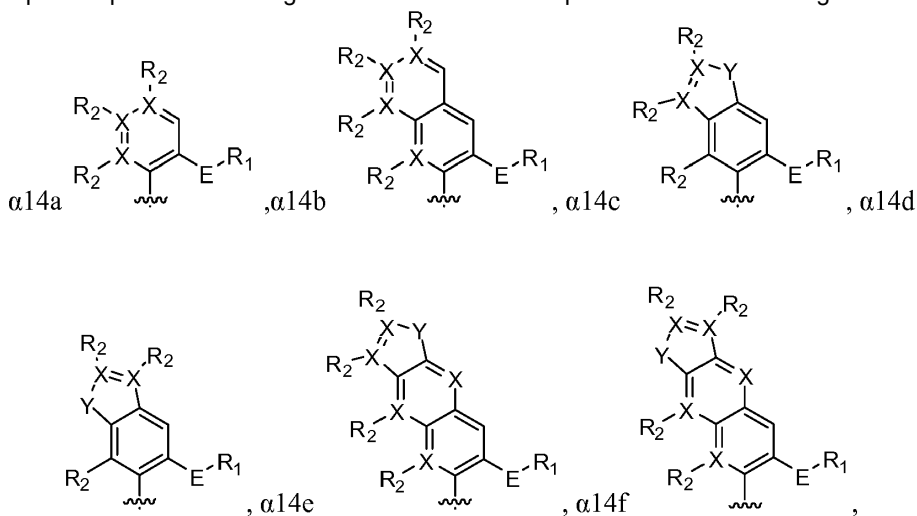
en las que cada R₂ es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoselenu, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga,

en las que cada Y es independientemente azufre, selenio o amina secundaria,

en las que cada E es azufre;

y la línea ondulada indica un punto de unión a una fracción ribosilo, desoxirribosilo o didesoxirribosilo o un análogo de las mismas, en la que la fracción ribosilo, desoxirribosilo o didesoxirribosilo o análogo de la misma está en forma libre, conectada a un grupo monofosfato, difosfato o trifosfato, que comprende opcionalmente un grupo α-tiotrifosfato, β-tiotrifosfato o γ-tiotrifosfato, o se incluye en un ARN o un ADN o en un análogo de ARN o un análogo de ADN.

2. Un compuesto que comprende un análogo de nucleobase de cualquiera de las fórmulas siguientes



en las que cada X es independientemente carbono o nitrógeno,

en las que cada R₁ es independientemente hidrógeno, grupo alquilo, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga,

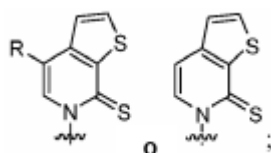
en las que cada R₂ está presente cuando X es carbono y es independientemente hidrógeno, grupo alquilo, alquenilo, alquinilo, metoxi, metanotiol, metanoselenu, halógeno, ciano, azida, un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o un enlazador acoplado al que se une una carga,

en las que cada Y es independientemente azufre, oxígeno, selenio o amina secundaria,

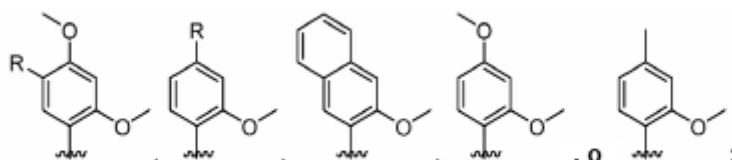
en las que cada E es azufre,

y la línea ondulada indica un punto de unión a una fracción ribosilo, desoxirribosilo o didesoxirribosilo o un análogo de los mismos, en el que la fracción ribosilo, desoxirribosilo o didesoxirribosilo o análogo de la misma está en forma libre, unida a un grupo monofosfato, difosfato o trifosfato, que comprende opcionalmente un grupo α-tiotrifosfato, β-tiotrifosfato o γ-tiotrifosfato, o se incluye en un ARN o un ADN o en un análogo de ARN o un análogo de ADN.

3. El compuesto de la reivindicación 1, en el que el centro reactivo comprende un grupo amino, un grupo acetilénico, un grupo tiol, un grupo aldehído o un grupo azida.
- 5 4. El compuesto de la reivindicación 2, en el que el centro reactivo comprende un grupo amino, un grupo acetilénico, un grupo tiol, un grupo aldehído o un grupo azida.
5. El compuesto de la reivindicación 1, en el que la fracción ribosilo o desoxirribosilo porta un grupo trifosfato o α -tiotrifosfato unido a un 5'-hidroxilo del mismo.
- 10 6. El compuesto de la reivindicación 2, en el que la fracción ribosilo o desoxirribosilo porta un grupo trifosfato o α -tiotrifosfato unido a un 5'-hidroxilo del mismo.
- 15 7. El compuesto de la reivindicación 1, en el que la fracción ribosilo o desoxirribosilo se incorpora en una cadena oligonucleotídica de ARN o ADN, respectivamente, o la fracción ribosilo o desoxirribosilo o análogo de las mismas se incorpora en un análogo de ARN o ADN.
- 20 8. El compuesto de la reivindicación 2, en el que la fracción ribosilo o desoxirribosilo se incorpora en una cadena oligonucleotídica de ARN o ADN, respectivamente, o la fracción ribosilo o desoxirribosilo o análogo de las mismas se incorpora en un análogo de ARN o ADN.
- 25 9. El compuesto de la reivindicación 1, en el que al menos un R_2 comprende independientemente: un grupo seleccionado del grupo que consiste en un grupo $-C\equiv C-CH_2NHR_3$, en la que R_3 es hidrógeno o es un grupo protector de amino; un grupo acetilénico adecuado para su uso en una reacción de clic con un reactivo de carga que comprende una carga y un grupo reactivo acetileno; un grupo tiol adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo tiol; un grupo aldehído adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo aldehído; un grupo azida adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo azida; y $-C\equiv C-(CH_2)_n-C=CH$, en la que n es 1, 2, 3, 4, 5 o 6; y $-C\equiv C-(CH_2)_{n1}-O-(CH_2)_{n2}-C=CH$, en la que n_1 y n_2 son cada uno independientemente 1, 2 o 3.
- 30 10. El compuesto de la reivindicación 2, en el que al menos un R_2 comprende independientemente: un grupo seleccionado del grupo que consiste en un grupo $-C\equiv C-CH_2NHR_3$, en la que R_3 es hidrógeno o es un grupo protector de amino; un grupo acetilénico adecuado para su uso en una reacción de clic con un reactivo de carga que comprende una carga y un grupo reactivo acetileno; un grupo tiol adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo tiol; un grupo aldehído adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo aldehído; un grupo azida adecuado para su uso en una reacción con un reactivo de carga que comprende una carga y un grupo reactivo azida; y $-C\equiv C-(CH_2)_n-C=CH$, en la que n es 1, 2, 3, 4, 5 o 6; y $-C\equiv C-(CH_2)_{n1}-O-(CH_2)_{n2}-C=CH$, en la que n_1 y n_2 son cada uno independientemente 1, 2 o 3.
- 35 40 11. Un par de nucleobases que comprende un primer análogo de nucleobase que tiene la fórmula



50 y un segundo análogo de nucleobase que tiene cualquiera de las fórmulas



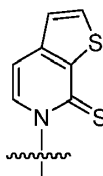
60 en las que R es independientemente hidrógeno o es un enlazador reactivo que comprende un centro reactivo adaptado para unirse a un reactivo de carga que comprende una carga y un grupo de reactividad complementaria al centro reactivo, o R es un enlazador acoplado al cual se une una carga;

65 y la línea ondulada indica un punto de unión a una fracción ribosilo, desoxirribosilo o didesoxirribosilo o un análogo de los mismos, en el que la fracción ribosilo, desoxirribosilo o didesoxirribosilo o análogo de la misma está en forma libre, unida a un grupo monofosfato, difosfato o trifosfato, que comprende opcionalmente un grupo α -tiotrifosfato, β -tiotrifosfato o γ -tiotrifosfato, o se incluye en un ARN o un ADN o en un análogo de ARN o un análogo de ADN.

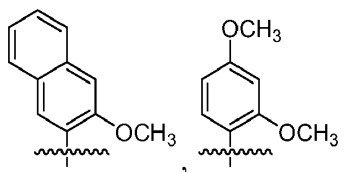
12. Un oligonucleótido bicatenario dúplex en el que una primera cadena oligonucleotídica comprende un compuesto de la reivindicación 1, y una segunda cadena oligonucleotídica complementaria comprende una nucleobase complementaria de emparejamiento de bases en un sitio complementario de emparejamiento de bases de la misma.

13. Un oligonucleótido bicatenario dúplex en el que una primera cadena oligonucleotídica comprende un compuesto de la reivindicación 2, y una segunda cadena oligonucleotídica complementaria comprende una nucleobase complementaria de emparejamiento de bases en un sitio complementario de emparejamiento de bases de la misma.

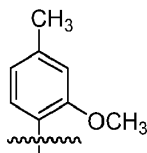
14. El oligonucleótido bicatenario duplex de la reivindicación 12, en el que la primera cadena comprende una nucleobase que tiene la fórmula



y la segunda cadena comprende una nucleobase de emparejamiento de bases complementaria seleccionada de



o

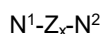


en un sitio complementario de emparejamiento de bases de la misma.

15. Un procedimiento para llevar a cabo una funcionalización específica del sitio de un oligonucleótido bicatenario dúplex, que comprende: incorporar una nucleobase no natural que comprende un enlazador reactivo que comprende un centro reactivo, se selecciona la nucleobase de un compuesto de la reivindicación 1, en una primera cadena oligonucleotídica; después, se sintetiza una segunda cadena oligonucleotídica complementaria a la primera cadena oligonucleotídica, la segunda cadena oligonucleotídica comprende una nucleobase complementaria a la nucleobase no natural en una posición complementaria específica del sitio, en condiciones tales que la primera cadena oligonucleotídica y la segunda cadena oligonucleotídica forman un oligonucleótido bicatenario dúplex; después, mediante el contacto del oligonucleótido bicatenario dúplex que incorpora la nucleobase no natural que comprende la fracción del enlazador reactivo con un reactivo de carga que comprende una carga y un grupo de reactividad complementaria, en condiciones adecuadas para que se produzca la reacción del enlazador reactivo y el grupo de reactividad complementaria para producir un enlazador acoplado; para proporcionar el oligonucleótido bicatenario dúplex funcionalizado con la carga que se une al mismo a través de un enlazador acoplado.

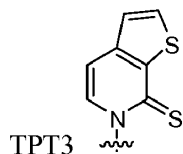
16. Un procedimiento para llevar a cabo una funcionalización específica del sitio de un oligonucleótido bicatenario dúplex, que comprende: incorporar una nucleobase no natural que comprende un enlazador reactivo que comprende un centro reactivo, se selecciona la nucleobase de un compuesto de la reivindicación 2, en una primera cadena oligonucleotídica; después, se sintetiza una segunda cadena oligonucleotídica complementaria a la primera cadena oligonucleotídica, la segunda cadena oligonucleotídica comprende una nucleobase complementaria a la nucleobase no natural en una posición complementaria específica del sitio, en condiciones tales que la primera cadena oligonucleotídica y la segunda cadena oligonucleotídica forman un oligonucleótido bicatenario dúplex; después, mediante el contacto del oligonucleótido bicatenario dúplex que incorpora la nucleobase no natural que comprende la fracción del enlazador reactivo con un reactivo de carga que comprende una carga y un grupo de reactividad complementaria, en condiciones adecuadas para que se produzca la reacción del enlazador reactivo y el grupo de reactividad complementaria para producir un enlazador acoplado; para proporcionar el oligonucleótido bicatenario dúplex funcionalizado con la carga que se une al mismo a través de un enlazador acoplado.

17. Una estructura que comprende la fórmula:

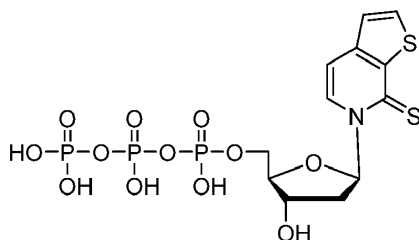
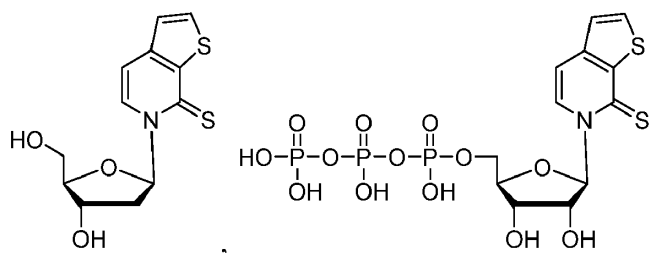
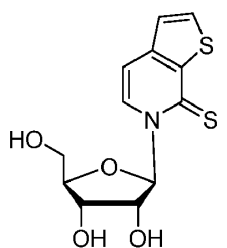


en la que N¹ es un nucleótido o análogo del mismo, o un grupo fosfato terminal; en la que N² es un nucleótido o análogo del mismo, o un grupo hidroxilo terminal; en la que Z es un compuesto de la reivindicación 1 o la reivindicación 2, y en la que x es un número entero de 1 a 20.

5 18. El compuesto de la reivindicación 1, que tiene una fórmula seleccionada de:

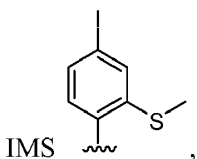
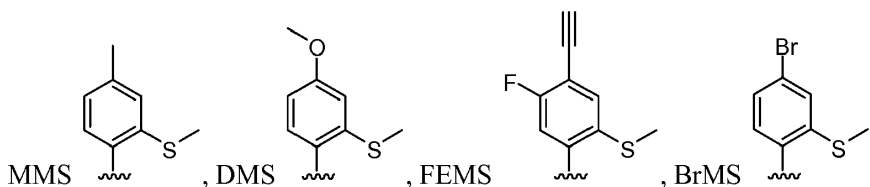


y derivados o análogos del mismo,



30 o

35 19. El compuesto de la reivindicación 2 seleccionado del grupo que consiste en:



50 y derivados o análogos del mismo.

Figura 1

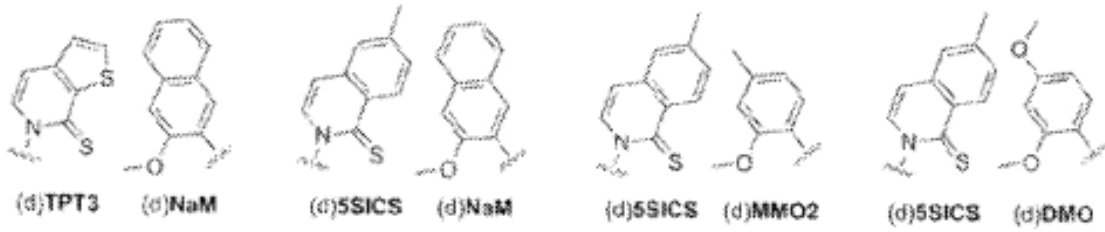


Figura 2

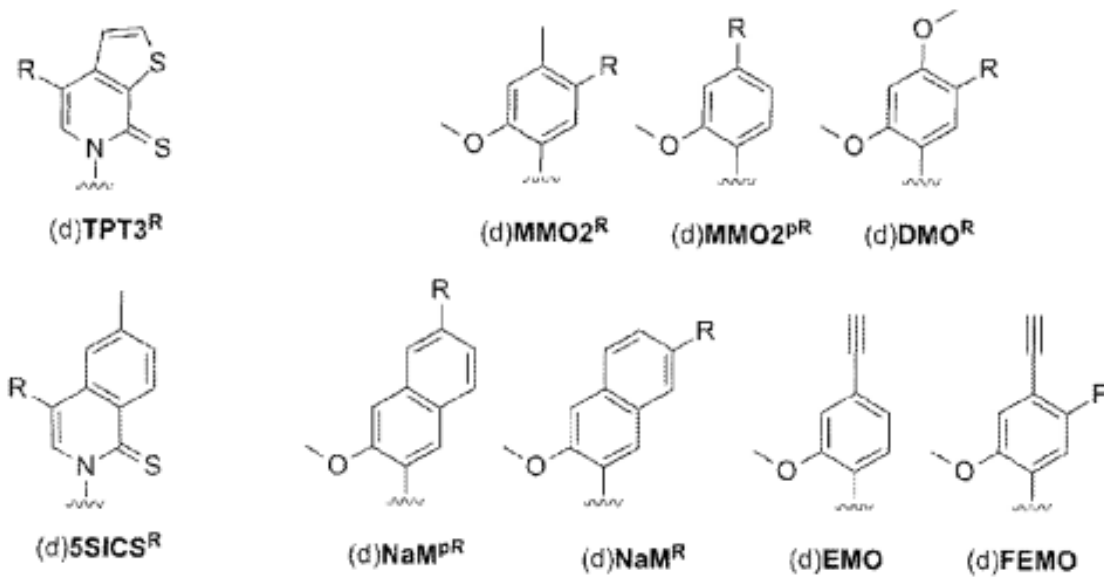


Figura 3

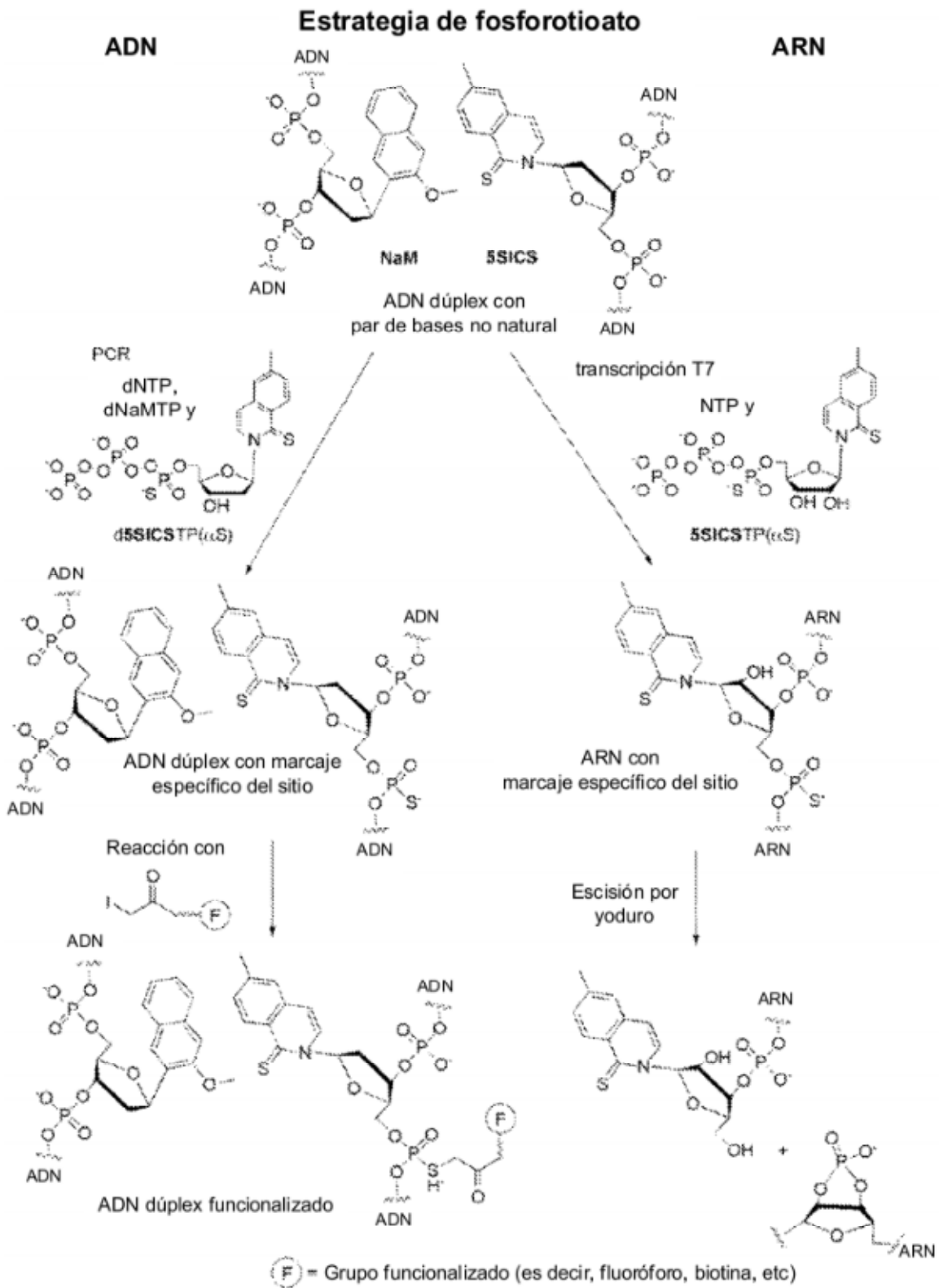


Figura 4

Estrategia amino

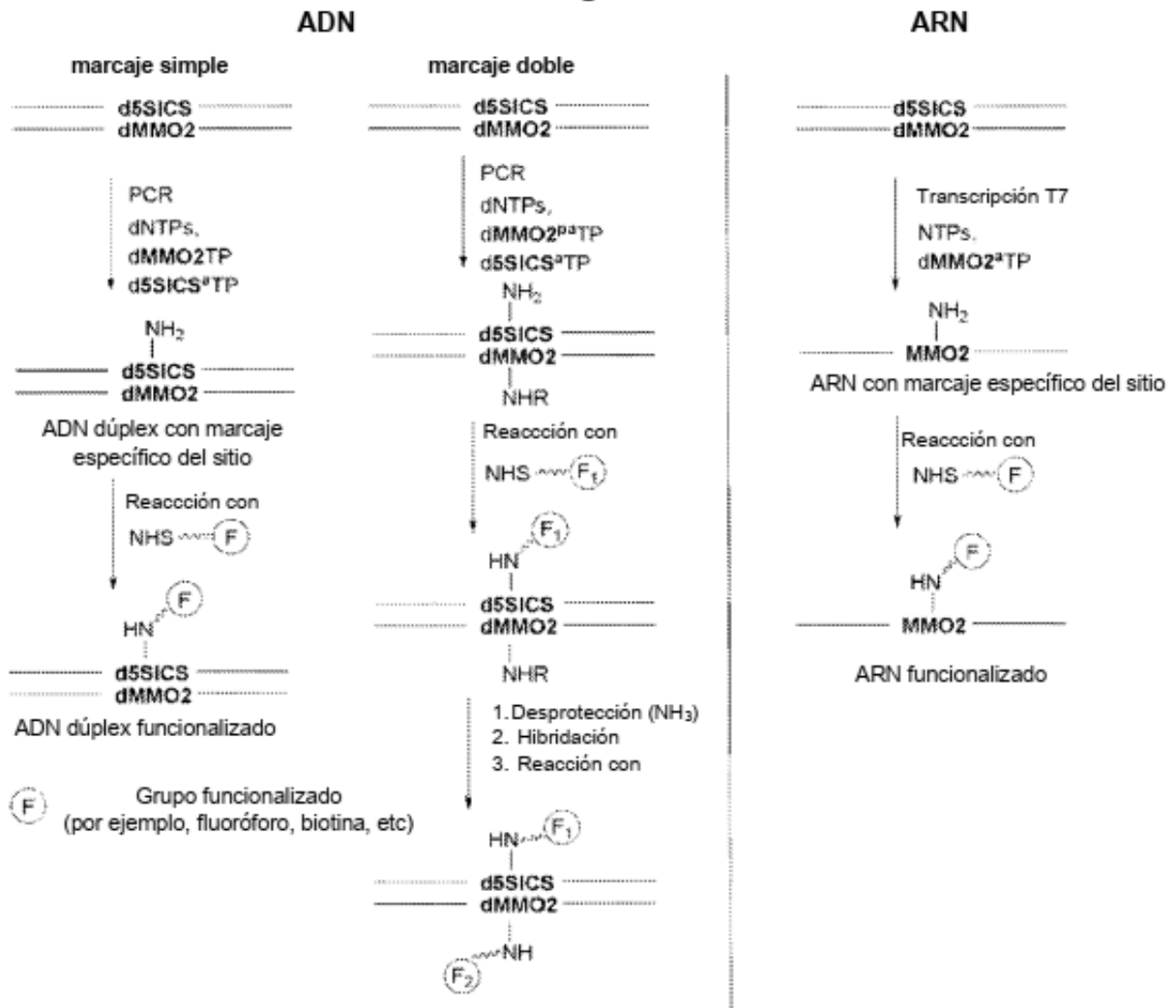


Figura 5

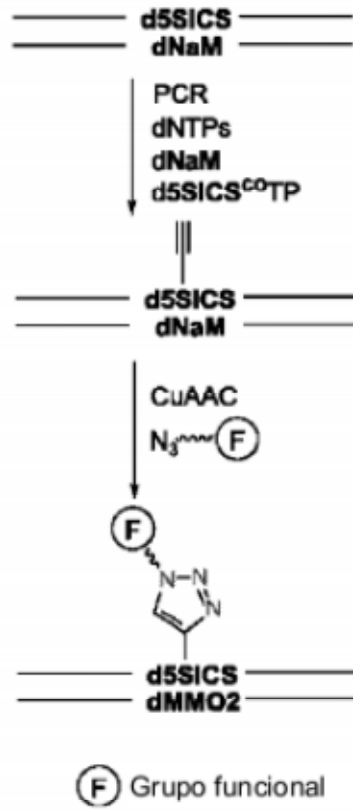


Figura 6

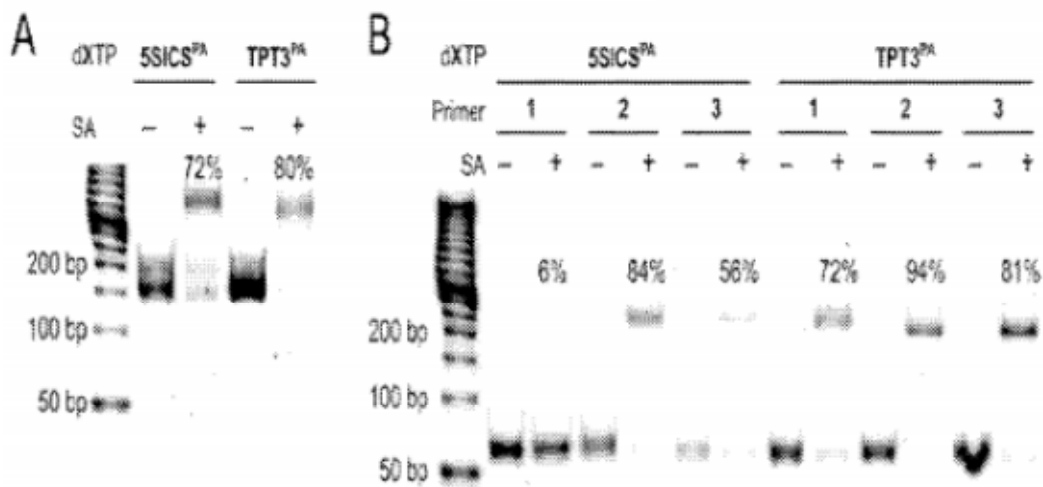


Figura 7

A

5'-d(ATAATACGACTCACTATAGGG)

3'-d(TATTATGCTGAGTGATATCCCTTAGGGCTC~~X~~TCAC)

↓
T7 RNA polymerase
NTPs, [α - 32 P] ATP
YTP

5'-GGGAAUCCCGAGYAGUG

B

X:

NaM

5SICS

Y:

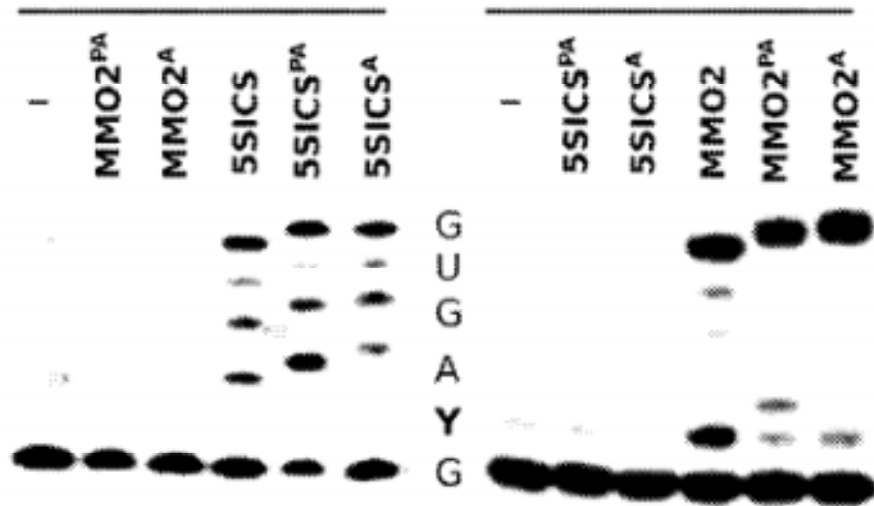


Figure 8

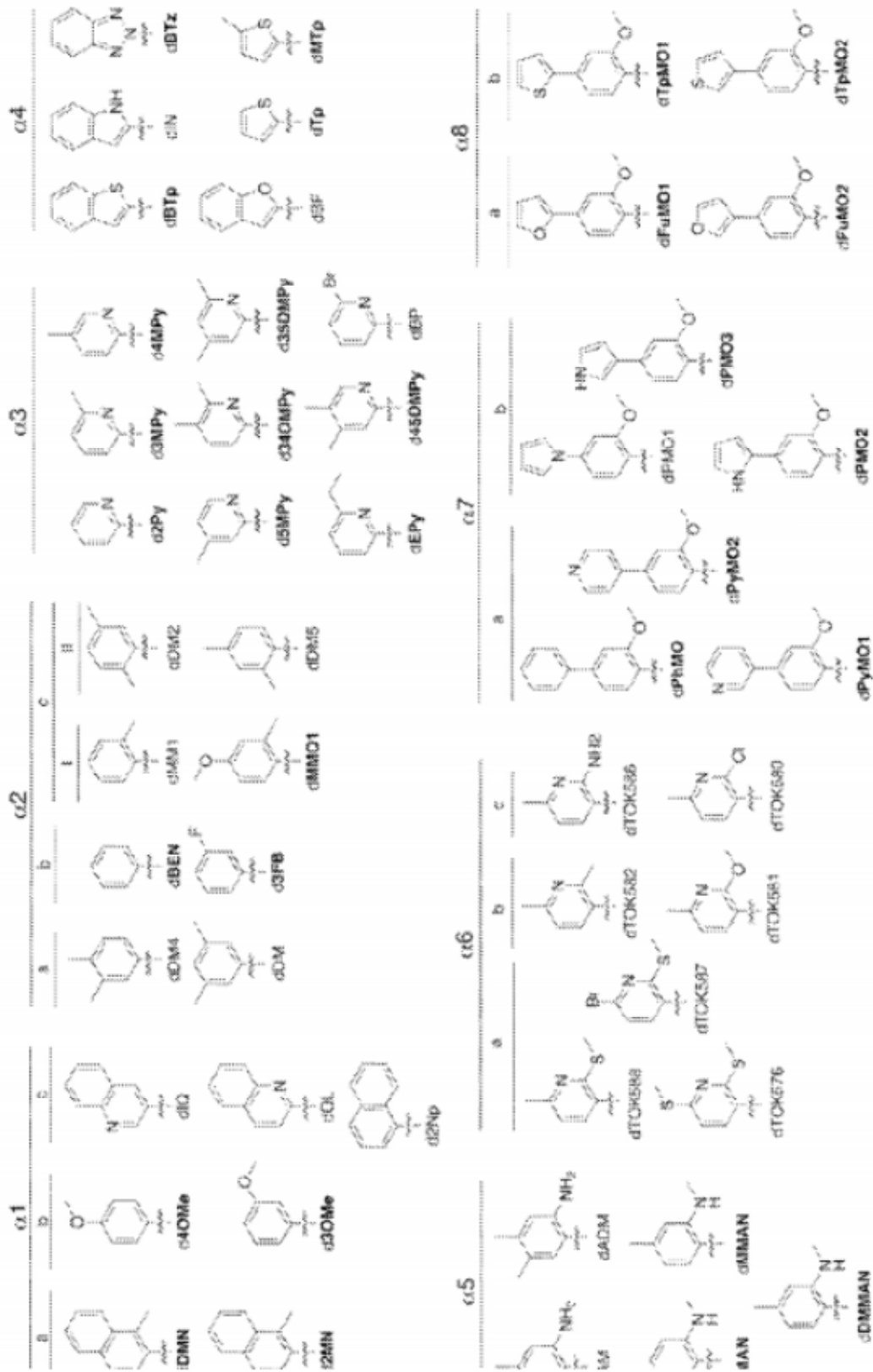


Figura 8
Continuación

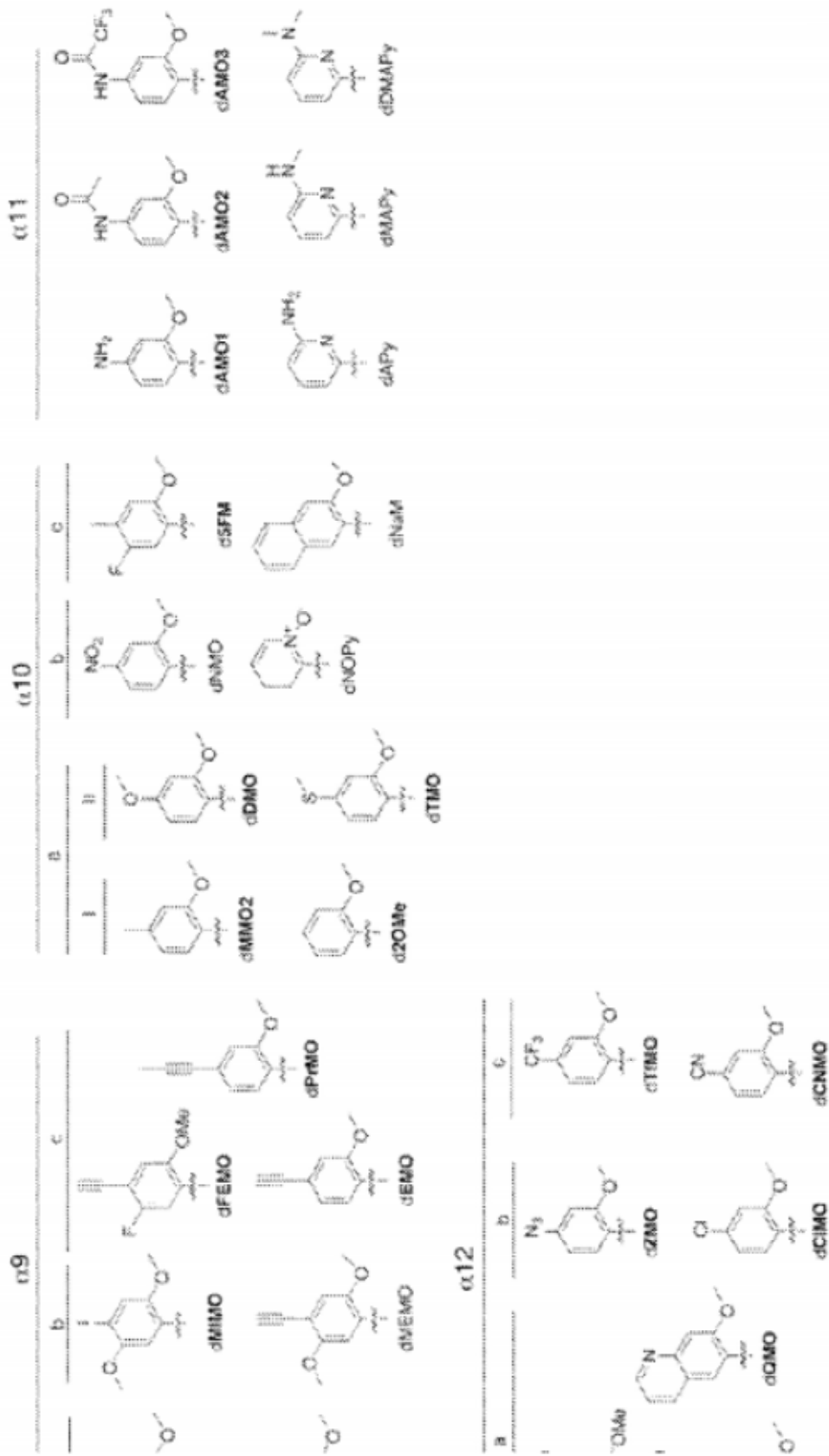


Figura 9

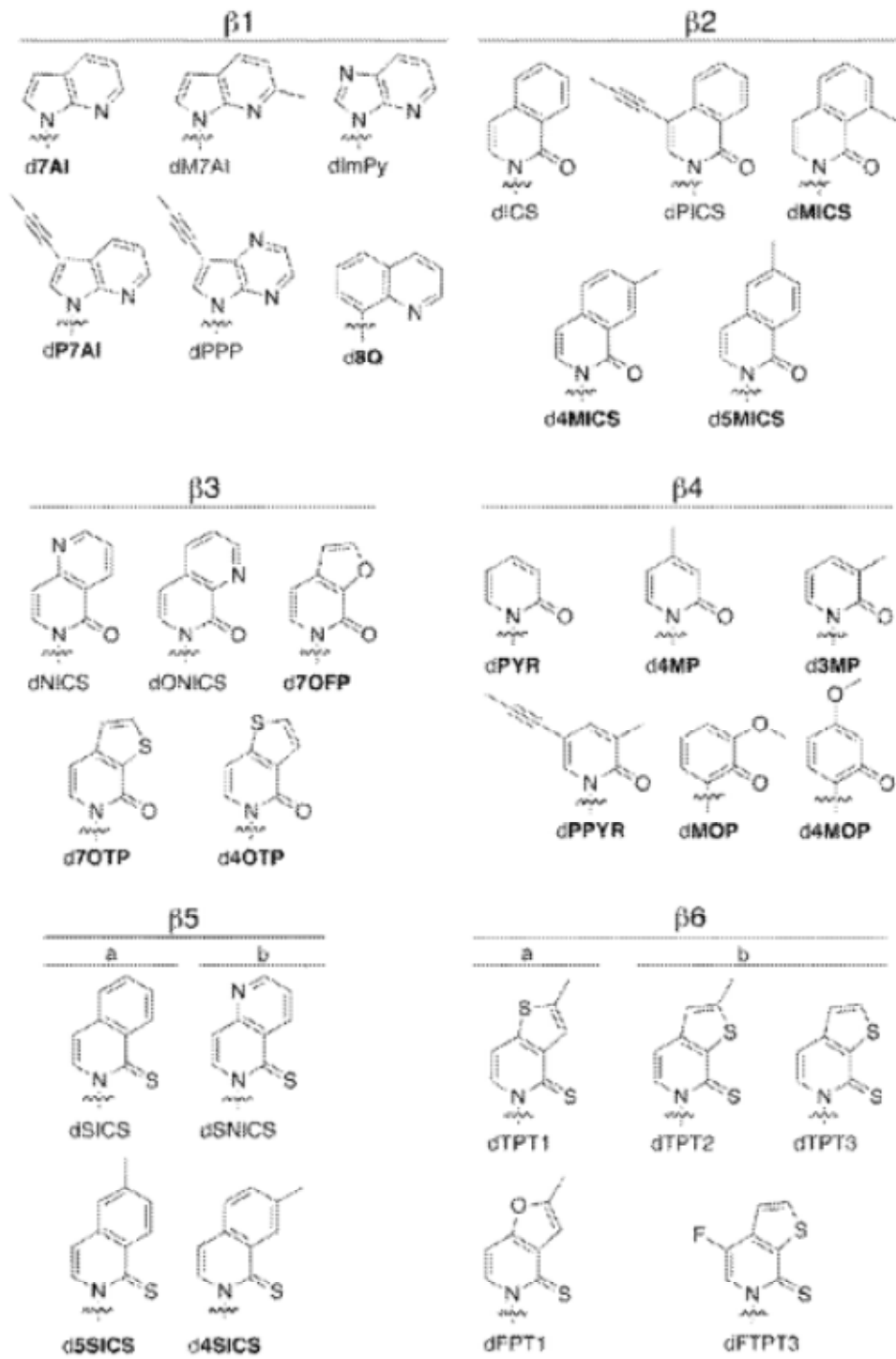


Figura 10

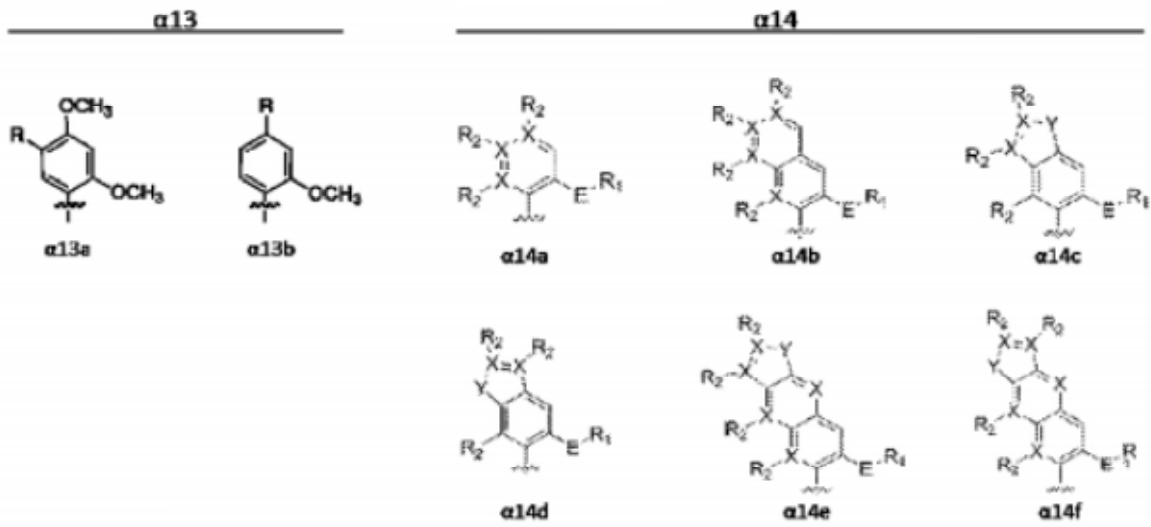


Figura 11

$\alpha 14a$

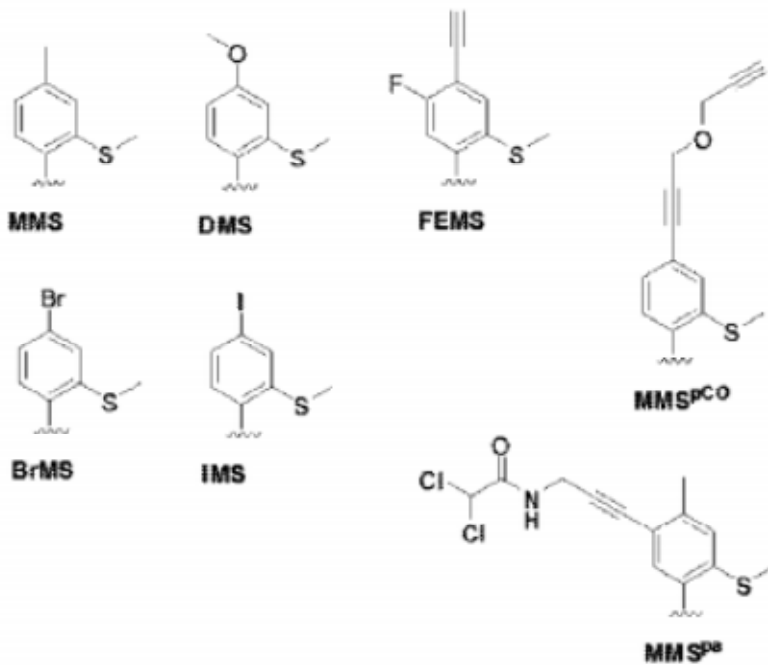


Figura 12

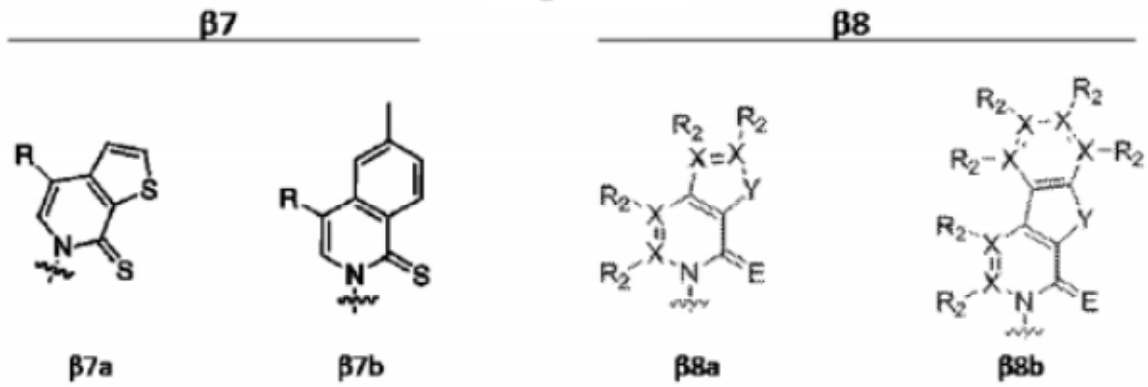


Figura 13

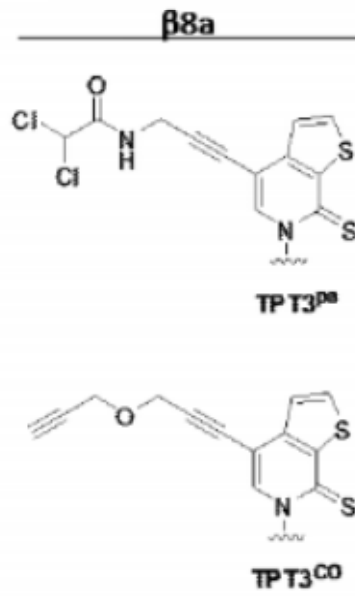


Figura 14

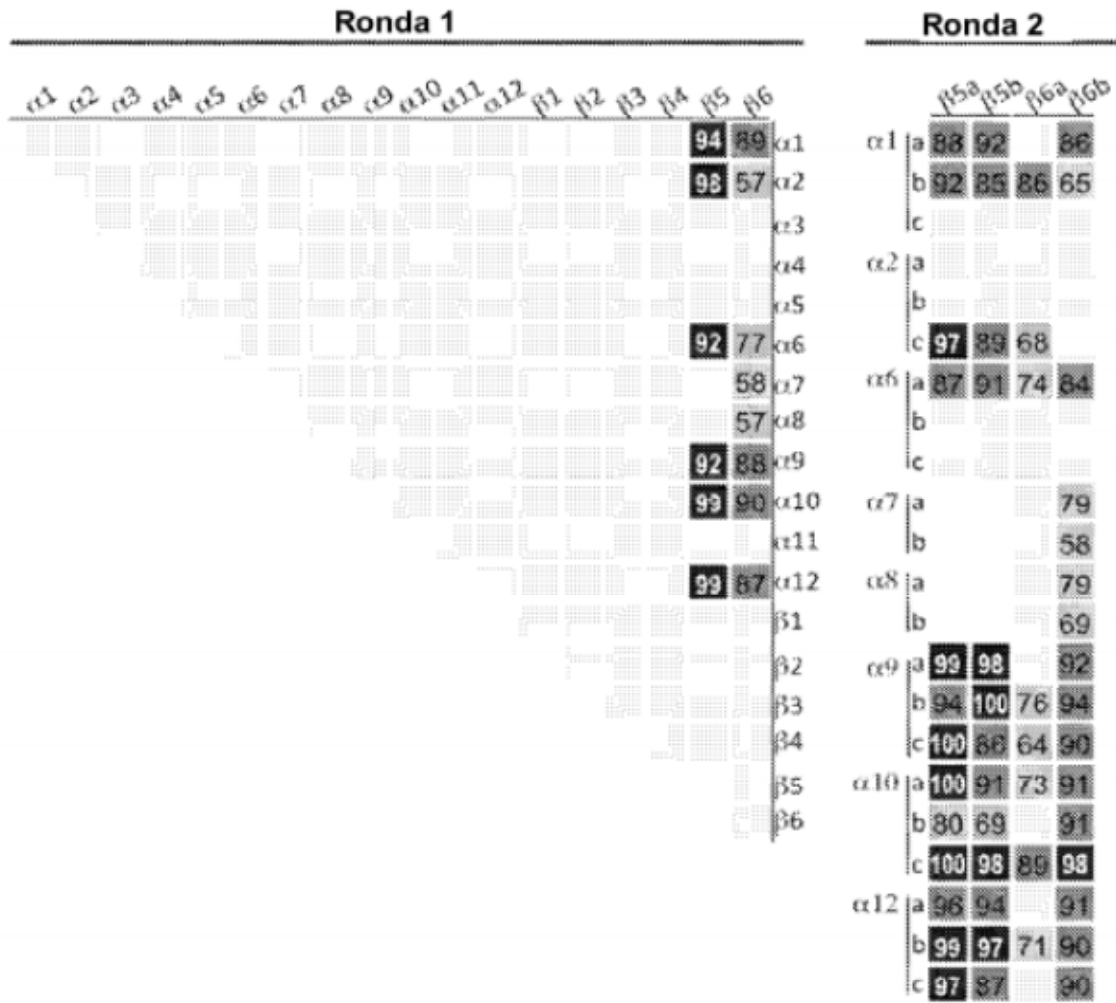


Figura 14
Continuación

Ronda 3					Ronda 4							
		β 5a	β 5b	TPT1	FPT1	β 6b			β 5a	β 5b	FPT1	β 6b
α 1	a	91	84			90	α 1	DMN				
	b							2MN	93	76		91
α 2	ci	83	71		63	55	α 2	MMO1				
	cl	97	89		69	78		MM1	78	96		65
α 6	TOK587	80	87			84		DM5				
	TOK588							DM2	100	91		76
	TOK576	53	73			54	α 6	TOK587	93	92		93
α 7	a					78		TOK576		73		
	PhMO					52	α 7	PyMO1				77
	b					79		PyMO2				80
	PMO3					61		PMO1				67
α 8	FuMO1					85		PMO2				76
	FuMO2					79	α 8	FuMO1				93
	TpMO1					77		FuMO2				91
	TpMO2					62		TpMO1				84
α 9	a	98	93			96	α 9	FIMO	99	98		100
	b	92	78			91		IMO	99	97		100
	c	96	94		73	99		FEMO	100	100		100
α 10	al	99	88		90	94		b	100	96		97
	all	97	88		56	94		c	100	99		100
α 12	NOPy						α 10	MMO2	99	89	70	91
	NMO	80	66			96		2OMe	100	84	93	81
	c	99	92	71	89	95		DMO	98	88		94
α 12	a	85	77			89		TMO	88	78		95
	QMO	88	76			86		NOPy				
	b	100	86		71	99		NMO	84	78		100
	TfMO	98	77		54	97		5FM	100	88	74	94
	CNMO	78	74			97		NaM	99	96	100	100
α 12							α 12	FDMO	83			63
								VMO	99	76		100
								ZMO	100	95		98
								CIMO	100	95		98
								TfMO	98	76		100
								QMO	93	93		94
								CNMO	73	63		98

Figura 14
Continuación

		Ronda 5							Ronda 6									
		SICS	5SICS	4SICS	3SICS	FP	TPT1	TPT2	TPT3	FTPT3								
α1	2MN				71			83	77									
	α2	MM1																
	DM2																	
α6	TOK587	77			81			66	76									
	TOK576				68													
α7	PyMO2																	
α8	FuMO1																	
	FuMO2																	
	TPMO1																	
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																	
α10	MMO2									76								
	ZOMe																	
	DMO																	
	TMO								67									
	NOPy																	
	NMO								86	94								
	5FM																	
	NaM	99	70		90			70	97	99								
	α12	FDMO							95	85	88							
		VMO									82							
ZMO		55			62			74	92	88								
CIMO		61			69			61	84	97								
TfMO									92	92								
QMO					70				71	87								
CNMO									84	94								

		OneTaq							Taq									
		SICS	5SICS	4SICS	3SICS	FP	TPT1	TPT2	TPT3	FTPT3								
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							

Figura 15

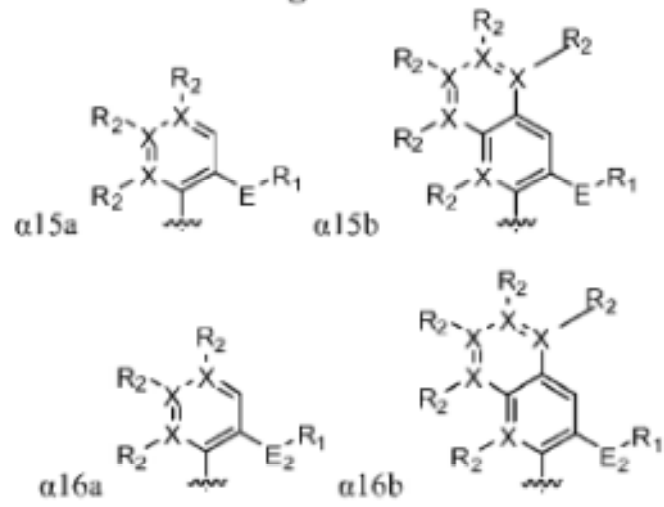


Figura 16

