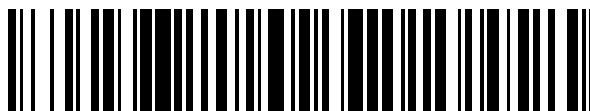


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 685 549**

51 Int. Cl.:

C07H 19/073 (2006.01)

C07H 19/10 (2006.01)

C07H 19/14 (2006.01)

C07H 19/173 (2006.01)

C07H 19/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.03.2013 PCT/EP2013/055466**

87 Fecha y número de publicación internacional: **18.09.2014 WO14139596**

96 Fecha de presentación y número de la solicitud europea: **15.03.2013 E 13711333 (8)**

97 Fecha y número de publicación de la concesión europea: **30.05.2018 EP 2970356**

54 Título: **Nucleósidos o nucleótidos modificados**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.10.2018

73 Titular/es:
**ILLUMINA CAMBRIDGE LIMITED (100.0%)
Chesterford Research Park
Little Chesterford Nr Saffron Walden Essex CB10
1XL, GB**

72 Inventor/es:
**LIU, XIAOHAI;
WU, XIAOLIN y
SMITH, GEOFFREY PAUL**

74 Agente/Representante:
ELZABURU, S.L.P

ES 2 685 549 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

Nucleósidos o nucleótidos modificados

Antecedentes**Campo de la invención**

- 5 Algunas realizaciones que se describen en la presente memoria se refieren a nucleótidos o nucleósidos modificados que comprenden grupos protectores 3'-hidroxi y a su uso en métodos de secuenciación de polinucleótidos. Algunas realizaciones que se describen en la presente memoria se refieren a un método de preparación de los nucleótidos o nucleósidos 3'-hidroxi protegidos.

Descripción de la técnica relacionada

- 10 Los avances en el estudio de moléculas han sido conducidos, en parte, por la mejora en las tecnologías usadas para caracterizar las moléculas o sus reacciones biológicas. En particular, el estudio de los ácidos nucleicos ADN y ARN se ha beneficiado del desarrollo de las tecnologías usadas para el análisis de secuencias y el estudio de sucesos de hibridación.

- 15 Un ejemplo de las tecnologías que han mejorado el estudio de los ácidos nucleicos es el desarrollo de agrupaciones fabricadas de ácidos nucleicos inmovilizados. Estas agrupaciones consisten por lo general en matrices de alta densidad de polinucleótidos inmovilizados sobre un material de soporte sólido. Véase, por ejemplo, Fodor *et al.*, *Trends Biotech.* 12: 19-26, 1994, que describe las formas de ensamblar los ácidos nucleicos usando una superficie de vidrio sensibilizada químicamente protegida con una máscara, pero expuesta en áreas definidas para permitir la unión de fosforamiditos de nucleótido modificados adecuadamente. Las agrupaciones fabricadas también se pueden fabricar mediante la técnica de "formación de puntos" de polinucleótidos conocidos sobre un soporte sólido en posiciones predeterminadas (por ejemplo, Stimpson *et al.*, *Proc. Natl. Acad. Sci.* 92: 6379-6383, 1995).

- 20 Una forma de determinar la secuencia de nucleótidos de un ácido nucleico unido a una agrupación se denomina "secuenciación por síntesis" o "SBS". Esta técnica para determinar la secuencia de ADN requiere de forma ideal la incorporación controlada (es decir, uno en cada ocasión) del nucleótido complementario correcto opuesto del ácido nucleico que se secuencia. Esto permite una secuenciación precisa mediante la adición de nucleótidos en múltiples ciclos ya que cada resto de nucleótido se secuencia uno en cada ocasión, previniendo de ese modo que se produzca una serie incontrolada de incorporaciones. El nucleótido incorporado se lee usando una etiqueta apropiada unida al mismo antes de la retirada del resto de la etiqueta y la nueva ronda de secuenciación posterior.

- 30 Con el fin de asegurar que solo se produzca una única incorporación, se añade una modificación estructural ("grupo protector") a cada nucleótido etiquetado que se añade a la cadena creciente para asegurar que solo se incorpora un único nucleótido. Después de que se haya añadido el nucleótido con el grupo protector, el grupo protector se retira a continuación, en condiciones de reacción que no interfieren con la integridad del ADN que se secuencia. El ciclo de secuenciación puede continuar a continuación con la incorporación del siguiente nucleótido protegido etiquetado.

- 35 Para que sean útiles en la secuenciación de ADN, los nucleótidos, y más habitualmente los trifosfatos de nucleótido, requieren generalmente un grupo protector de 3'-hidroxi de modo que se prevenga que la polimerasa usada para incorporarlos en la cadena de polinucleótidos continúe la replicación una vez se ha añadido la base en el nucleótido. Existen numerosas limitaciones en los tipos de grupos que se pueden añadir a un nucleótido y aún son adecuados. El grupo protector debería prevenir que las moléculas de nucleótido adicionales se añadan a la cadena de polinucleótido mientras que, simultáneamente, deberían ser fáciles de retirar del resto de azúcar sin causar daño a la cadena de polinucleótido. Además, el nucleótido modificado necesita ser tolerado por la polimerasa u otra enzima apropiada usada para incorporarlo en la cadena de polinucleótido. Por lo tanto, el grupo protector ideal exhibe estabilidad a largo plazo, se incorpora de forma eficaz mediante la enzima polimerasa, causa el bloqueo de la incorporación de nucleótidos secundarios o adicionales y tiene la capacidad de retirarse en condiciones suaves que no causan daño a la estructura del polinucleótido, preferentemente en condiciones acuosas.

- 45 Se han descrito anteriormente grupos protectores reversibles. Por ejemplo, Metzker *et al.*, (*Nucleic Acids Research*, 22 (20): 4259-4267, 1994) describe la síntesis y el uso de ocho 5'-trifosfatos de 2-desoxirribonucleósido 3'-modificados (dNTP 3'-modificados) y se someten a ensayo en dos ensayos de plantilla de ADN para la actividad de incorporación. El documento de Patente WO 2002/029003 describe un método de secuenciación que puede incluir el uso de un grupo protector alilo para taponar el grupo 3'-OH en una cadena en crecimiento de ADN en una reacción de polimerasa.

- 50 Los documentos de Patente WO 2012/162429 y WO 2012/083249 describen nucleósidos y nucleótidos modificados que comprenden una base de purina o pirimidina y un resto de azúcar de ribosa o desoxirribosa que tienen un grupo protector de 3'-hidroxi retirable.

- 55 Además, los presentes inventores han informado previamente del desarrollo de una diversidad de grupos protectores reversibles y métodos de desprotección de los mismos en condiciones compatibles con el ADN en el documento de Publicación de Solicitud Internacional n.º WO 2004/018497.

Sumario

- Algunas realizaciones que se describen en la presente memoria se refieren a una molécula de nucleótido o nucleósido modificada que comprende una base de purina o pirimidina y un resto de azúcar de ribosa o desoxirribosa que tiene un grupo protector de 3'-hidroxi retirable que forma una estructura $-O-C(R)_2N_3$ unida covalentemente al átomo de carbono 3', en donde
- 5 R se selecciona entre el grupo que consiste en hidrógeno, $-C(R^1)_m(R^2)_n-C(=O)OR^3$, $-C(=O)NR^4R^5$, $-C(R^6)_2O(CH_2)_pNR^7R^8$ y $-C(R^9)_2O-Ph-C(=O)NR^{10}R^{11}$;
- cada R^1 y R^2 se selecciona independientemente entre hidrógeno o halógeno;
- R^3 se selecciona entre hidrógeno o alquilo opcionalmente sustituido;
- 10 cada R^4 y R^5 se selecciona independientemente entre hidrógeno, alquilo opcionalmente sustituido, arilo opcionalmente sustituido, heteroarilo opcionalmente sustituido, o aralquilo opcionalmente sustituido;
- cada R^6 y R^9 se selecciona entre hidrógeno, alquilo opcionalmente sustituido o halógeno;
- cada R^7 , R^8 , R^{10} y R^{11} se selecciona independientemente entre hidrógeno, alquilo opcionalmente sustituido, arilo opcionalmente sustituido, heteroarilo opcionalmente sustituido, o aralquilo opcionalmente sustituido;
- 15 m es un número entero de 0 a 3; y
- n es un número entero de 0 a 3; con la condición de que el total de $m + n$ sea igual a 3; y
- p es un número entero de 0 a 6; con la condición de que
- R^1 y R^2 no pueden ser halógeno; y
- al menos un R no es hidrógeno.
- 20 Algunas realizaciones que se describen en la presente memoria se refieren a un método para preparar un polinucleótido en crecimiento complementario a un polinucleótido monocatenario diana en una reacción de secuenciación, que comprende incorporar una molécula de nucleótido modificada que se describe en la presente memoria en el polinucleótido complementario en crecimiento, en donde la incorporación del nucleótido modificado previene la introducción de cualquier nucleótido posterior en el polinucleótido complementario en crecimiento.
- 25 Algunas realizaciones que se describen en la presente memoria se refieren a un método para determinar la secuencia de un polinucleótido monocatenario diana, que comprende monitorizar la incorporación secuencial de nucleótidos complementarios, en donde al menos un nucleótido complementario incorporado es una molécula de nucleótido modificada que se describe en la presente memoria; y detectar la identidad de la molécula de nucleótido modificada. En algunas realizaciones, la incorporación de la molécula de nucleótido modificada se consigue mediante una transferasa terminal, una polimerasa terminal o una transcriptasa inversa.
- 30 Algunas realizaciones que se describen en la presente memoria se refieren a un kit que comprende una pluralidad de moléculas de nucleótido o nucleósido modificadas que se describen en la presente memoria, y materiales de envase para los mismos. En algunas realizaciones, la identidad del nucleótido modificado se determina por detección de la etiqueta detectable unida a la base. En algunas de tales realizaciones, el grupo protector de 3'-hidroxi y la etiqueta detectable se retiran antes de introducir el siguiente nucleótido complementario. En algunas de tales realizaciones, el grupo protector de 3'-hidroxi y la etiqueta detectable se retiran en una etapa individual de reacción química.
- 35
- Breve descripción de los dibujos**
- La Figura 1A** ilustra una diversidad de grupos protectores de 3'-OH.
- 40 **La Figura 1B** ilustra la estabilidad térmica de diversos grupos protectores de 3'-OH.
- La Figura 2A** ilustra la curva de la velocidad de desprotección de tres grupos protectores de 3'-OH diferentes.
- La Figura 2B** muestra un gráfico del tiempo medio de desprotección de tres grupos protectores de 3'-OH diferentes.
- La Figura 3** muestra los valores de ajuste y preajuste de fase de diversos nucleótidos modificados con un grupo protector de 3'-OH térmicamente estable en comparación con el grupo protector estándar.
- 45 **La Figura 4A** muestra los datos de secuenciación de 2 x 400 pb de mono-F ffNs-A-isómero en la mezcla de incorporación (IMX).

La Figura 4B muestra los datos de secuenciación de 2 x 400 pb de mono-F ffNs-B-isómero en la mezcla de incorporación (IMX).

Descripción detallada

5 En la presente memoria se describe un nucleótido o nucleósido modificado que comprende un grupo protector de 3'-OH. En una variante, el grupo protector de 3'-OH es un grupo protector azidometilo sustituido con monofluorometilo. En otra variante, el grupo protector de 3'-OH es un grupo protector azidometilo C-amido sustituido. Otra variante más se refiere a nucleótidos modificados que tienen grupos protectores de 3'-OH azidometilo sustituido con difluorometilo.

Definiciones

10 A menos que se definan de otro modo, todos los términos y expresiones técnicos y científicos que se usan en la presente memoria tienen el mismo significado que se entiende habitualmente por el experto habitual en la materia. El uso de la expresión "que incluye" así como otras formas, tales como "incluir", "incluye", e "incluido" no es limitante. El uso de la expresión "que tiene" así como otras formas, tales como "tener", "tiene", y "tenía" no es limitante. Como se usa en la presente memoria descriptiva, ya sea en una expresión transicional o en el cuerpo de la reivindicación, se ha de interpretar que el término "comprende" y la expresión "que comprende" tienen un significado abierto. Es decir, el término y la expresión anteriores se han de interpretar de forma sinónima con las expresiones "que tiene al menos" o "que incluye al menos". Por ejemplo, cuando se usa en el contexto de un proceso, la expresión "que comprende" significa que el proceso incluye al menos las etapas recitadas, pero pueden incluir etapas adicionales. Cuando se usa en el contexto de un compuesto, composición, o dispositivo, la expresión "que comprende" significa que el compuesto, composición, o dispositivo incluye al menos las características o componentes recitados, pero también puede incluir características o componentes adicionales.

20

Como se usa en la presente memoria, las abreviaturas orgánicas comunes se definen como sigue a continuación:

	Ac	Acetilo
	Ac ₂ O	Anhídrido acético
25	ac.	Acuoso
	Bn	Bencilo
	Bz	Benzoílo
	BOC o Boc	terc-Butoxicarbonilo
	Bu	n-Butilo
30	cat.	Catalítico
	Cbz	Carbobenciloxi
	°C	Temperatura en grados centígrados
	dATP	Trifosfato de desoxiadenosina
	dCTP	Trifosfato de desoxicitidina
35	dGTP	Trifosfato de desoxiguanosina
	dTTP	Trifosfato de desoxitimidina
	ddNTP(s)	Didesoxinucleótido(s)
	DBU	1,8-Diazabicyclo[5.4.0]undec-7-eno
	DCA	Ácido dicloroacético
40	DCE	1,2-Dicloroetano
	DCM	Cloruro de metileno
	DIEA	Diisopropiletilamina
	DMA	Dimetilacetamida
	DME	Dimetoxietano

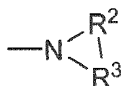
	DMF	N,N'-Dimetilformamida
	DMSO	Dimetilsulfóxido
	DPPA	Difenilfosforil azida
	Et	Etilo
5	EtOAc	Acetato de etilo
	ffN	Nucleótido completamente funcional
	g	Gramos
	GPC	Cromatografía de permeación en gel
	h o hr	Horas
10	iPr	Isopropilo
	KPi	Tampón de fosfato de potasio 10 mM a pH 7,0
	KPS	Persulfato de potasio
	IPA	Alcohol isopropílico
	IMX	Mezcla de incorporación
15	LCMS	Cromatografía líquida-espectrometría de masas
	LDA	Diisopropilamina de litio
	m o min	Minutos
	mCPBA	Ácido meta-cloroperoxibenzoico
	MeOH	Metanol
20	MeCN	Acetonitrilo
	Mono-F	-CH ₂ F
	Mono-F ffN	Nucleótidos modificados con -CH ₂ F sustituido en posición de metileno del grupo protector de 3'-OH azidometilo
	ml	Mililitros
25	MTBE	Metil terc-butil éter
	NaN ₃	Azida de sodio
	NHS	N-hidroxisuccinimida
	PG	Grupo protector
	Ph	Fenilo
30	Ppt	Precipitado
	ta	Temperatura ambiente
	SBS	Secuenciación por síntesis
	TEA	Trietilamina
	TEMPO	(2,2,6,6-Tetrametilpiperidin-1-il)oxilo
35	TCDI	1,1'-Tiocarbonil diimidazol
	Terc, t	Terciario
	TFA	Ácido trifluoroacético

THF	Tetrahidrofurano
TEMED	Tetrametiletilendiamina
μl	Microlitros

5 Como se usa en la presente memoria, el término "agrupación" se refiere a una población de moléculas de sonda diferentes que están unidas a uno o más sustratos de un modo tal que las diferentes moléculas de sonda se pueden diferenciar entre sí de acuerdo con la ubicación relativa. Una agrupación puede incluir diferentes moléculas de sonda que están cada una ubicadas en una ubicación direccionable diferente en un sustrato. Además o alternativamente, una agrupación puede incluir sustratos separados que portan cada uno una molécula de sonda diferente, en donde las moléculas de sonda diferentes se pueden identificar de acuerdo con las ubicaciones de los sustratos sobre una superficie a la que están unidos los sustratos o de acuerdo con las ubicaciones de los sustratos en un líquido. Algunas agrupaciones a modo de ejemplo en las que los sustratos separados están situados sobre una superficie incluyen, pero no se limitan a, las que incluyen perlas en las paredes como se describe, por ejemplo, en el documento de Patente de Estados Unidos n.º 6.355.431 B1, el documento de Patente US 2002/0102578 y el documento de Publicación PCT n.º WO 00/63437. Algunos formatos a modo de ejemplo que se pueden usar en la invención para distinguir perlas en una agrupación líquida, por ejemplo, usando un dispositivo microfluído, tal como un clasificador de células activado fluorescente (FACS), se describen, por ejemplo, en el documento de Patente de Estados Unidos n.º 6.524.793. Algunos ejemplos adicionales de agrupaciones que se pueden usar en la invención incluyen, pero no se limitan a, las que se describen en los documentos de Patente de Estados Unidos con números 5.429.807; 5.436.327; 5.561.071; 5.583.211; 5.658.734; 5.837.858; 5.874.219; 5.919.523; 6.136.269; 6.287.768; 20 6.287.776; 6.288.220; 6.297.006; 6.291.193; 6.346.413; 6.416.949; 6.482.591; 6.514.751 y 6.610.482; y los documentos de Patente WO 93/17126; WO 95/11995; WO 95/35505; EP 742 287; y EP 799 897.

25 Como se usa en la presente memoria, la expresión "unido covalentemente" o "enlazado covalentemente" se refiere a la formación de un enlace químico que está caracterizado por la compartición de pares de electrones entre átomos. Por ejemplo, un revestimiento de polímero unido covalentemente se refiere a un revestimiento de polímero que forma enlaces químicos con una superficie funcionalizada de un sustrato, en comparación con la unión a la superficie a través de otros medios, por ejemplo, adhesión o interacción electrostática. Se ha de entender que los polímeros que están unidos covalentemente a una superficie también pueden estar unidos a través de medios además de la unión covalente.

30 Como se usa en la presente memoria, cualquier grupo o grupos "R" tal como, sin limitación, R², R³, R⁴, R⁵, R⁶, R⁷, y R⁸ representa sustituyentes que pueden estar unidos al átomo indicado. Un grupo R puede estar sustituido o sin sustituir. Si se describe que dos grupos "R" se "toman juntos" los grupos R y los átomos a los que están unidos pueden formar un cicloalquilo, arilo, heteroarilo, o heterociclo. Por ejemplo, sin limitación, si R² y R³, o R², R³, o R⁴, y el átomo al que están unidos, se indica que se "toman juntos" o se "unen juntos" significa que están unidos covalentemente los unos a los otros para formar un anillo, un ejemplo del cual se expone a continuación:



35 Cada vez que se describe que un grupo está "opcionalmente sustituido" ese grupo puede estar sin sustituir o sustituido con uno o más de los sustituyentes indicados. De forma análoga, cuando se describe que un grupo está "sin sustituir o sustituido" si estuviera sustituido, el sustituyente se puede seleccionar entre uno o más de los sustituyentes indicados. Si no se indica ningún sustituyente, se pretende indicar que el grupo "opcionalmente sustituido" o "sustituido" indicado puede estar individual e independientemente sustituido con uno o más grupos individual e independientemente seleccionados entre un grupo de funcionalidades que incluyen, pero no se limitan a, alquilo, alqueno, alquino, cicloalquilo, cicloalqueno, cicloalquino, arilo, heteroarilo, heteroalíclico, aralquilo, heteroaralquilo, (heteroalíclico)alquilo, hidroxilo, hidroxilo protegido, alcoxi, arilo, acilo, mercapto, alquilo, arililo, ciano, halógeno, tiocarbonilo, O-carbamilo, N-carbamilo, O-tiocarbamilo, N-tiocarbamilo, C-amido, N-amido, S-sulfonamido, N-sulfonamido, C-carboxi, C-carboxi protegido, O-carboxi, isocianato, tiocianato, isotiocianato, nitro, sililo, sulfenilo, sulfinilo, sulfonilo, haloalquilo, haloalcoxi, trihalometanosulfonilo, trihalometanosulfonamido, amino, grupo amino monosustituido, grupo amino disustituido, y los derivados protegidos de los mismos.

50 Como se usa en la presente memoria, "alquilo" se refiere a una cadena de hidrocarburo lineal o ramificada que comprende un grupo hidrocarburo completamente saturado (sin dobles ni triples enlaces). En algunas realizaciones, el grupo alquilo puede tener de 1 a 20 átomos de carbono (cada vez que aparece en la presente memoria, un intervalo numérico tal como "1 a 20" se refiere a cada número entero en el intervalo dado inclusive los puntos de los extremos; por ejemplo, "de 1 a 20 átomos de carbono" significa que el grupo puede consistir en 1 átomo de carbono, 2 átomos de carbono, 3 átomos de carbono, etc., hasta e incluyendo 20 átomos de carbono, aunque la presente definición también incluye la aparición del término "alquilo" cuando no se indica ningún intervalo numérico). El grupo alquilo también puede ser un alquilo de tamaño medio que tiene de aproximadamente 7 a aproximadamente 10 átomos de carbono. El grupo alquilo también puede ser un grupo alquilo inferior que tiene de 1 a 6 átomos de carbono. El grupo alquilo de los compuestos se puede denominar "alquilo C₁-C₄" o denominaciones similares.

Únicamente a modo de ejemplo, "alquilo C₁-C₄" indica que hay de uno a cuatro átomos de carbono en la cadena de alquilo, es decir, la cadena de alquilo se selecciona metilo, etilo, propilo, iso-propilo, n-butilo, iso-butilo, sec-butilo, y t-butilo. Los grupos alquilo habituales incluyen, pero no se limitan en modo alguno a, metilo, etilo, propilo, isopropilo, butilo, isobutilo, terc-butilo, pentilo, y hexilos. El grupo alquilo puede estar sustituido o sin sustituir.

- 5 Como se usa en la presente memoria, "alquenilo" se refiere a un grupo alquilo que contiene en la cadena de hidrocarburo lineal o ramificada uno o más dobles enlaces. Un grupo alquenilo puede estar sin sustituir o sustituido.

Como se usa en la presente memoria, "alquinilo" se refiere a un grupo alquilo que contiene en la cadena de hidrocarburo lineal o ramificada uno o más triples enlaces. Un grupo alquinilo puede estar sin sustituir o sustituido.

- 10 Como se usa en la presente memoria, "cicloalquilo" se refiere a un sistema de anillos de hidrocarburo mono o multicíclico completamente saturado (sin dobles ni triples enlaces). Cuando está compuesto por dos o más anillos, los anillos se pueden unir juntos de una forma condensada. Los grupos cicloalquilo pueden contener de 3 a 10 átomos de carbono en el anillo o anillos. En algunas realizaciones, los grupos cicloalquilo pueden contener de 3 a ocho átomos en el anillo o anillos. Un grupo cicloalquilo puede estar sin sustituir o sustituido. Los grupos cicloalquilo habituales incluyen, pero no se limitan en modo alguno a, ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo, cicloheptilo, y ciclooctilo.

- 15 Como se usa en la presente memoria, "arilo" se refiere a un sistema de anillos aromático monocíclico o multicíclico carbocíclico (todo carbonos) (que incluye, por ejemplo, sistemas de anillos condensados, con puente, o espirocíclicos donde dos anillos de carbono comparten un enlace químico, por ejemplo, uno o más anillos de arilo con uno o más anillos de arilo o no arilo) que tienen un sistema de electrones pi completamente deslocalizado en al menos uno de los anillos. El número de átomos de carbono en un grupo arilo puede variar. Por ejemplo, en algunas realizaciones, el grupo arilo puede ser un grupo arilo C₆-C₁₄, un grupo arilo C₆-C₁₀, o un grupo arilo C₆. Algunos ejemplos de grupos arilo incluyen, pero no se limitan a, benceno, naftaleno, y azuleno. Un grupo arilo puede estar sustituido o sin sustituir.

- 20 Como se usa en la presente memoria, "heterociclilo" se refiere a sistemas de anillos que incluyen al menos un heteroátomo (por ejemplo, O, N, S). Tales sistemas pueden ser insaturados, pueden incluir alguna insaturación, o pueden contener alguna parte aromática, o pueden ser completamente aromáticos. Un grupo heterociclilo puede estar sustituido o sin sustituir.

- 25 Como se usa en la presente memoria, "heteroarilo" se refiere a un sistema de anillos aromático monocíclico o multicíclico (un sistema de anillos que tiene al menos un anillo con un sistema de electrones pi completamente deslocalizado) que contiene uno o más heteroátomos, es decir, un elemento distinto del carbono, que incluye, pero no se limita a, nitrógeno, oxígeno, y azufre, y al menos un anillo aromático. El número de átomos en el anillo o anillos del grupo heteroarilo puede variar. Por ejemplo, en algunas realizaciones, un grupo heteroarilo puede contener de 4 a 14 átomos en el anillo o anillos, de 5 a 10 átomos en el anillo o anillos o de 5 a 6 átomos en el anillo o anillos. Además, el término "heteroarilo" incluye sistemas de anillos condensados donde dos anillos, tales como al menos un anillo de arilo y al menos un anillo de heteroarilo, o al menos dos anillos de heteroarilo, comparten al menos un enlace químico. Algunos ejemplos de anillos de heteroarilo incluyen, pero no se limitan a, furano, furazano, tiofeno, benzotiofeno, ftalazina, pirrol, oxazol, benzoxazol, 1,2,3-oxadiazol, 1,2,4-oxadiazol, tiazol, 1,2,3-tiadiazol, 1,2,4-tiadiazol, benzotiazol, imidazol, benzoimidazol, indol, indazol, pirazol, benzopirazol, isoxazol, benzoisoxazol, isotiazol, triazol, benzotriazol, tiadiazol, tetrazol, piridina, piridazina, pirimidina, pirazina, purina, pteridina, quinolina, isoquinolina, quinazolina, quinoxalina, cinolina, y triazina. Un grupo heteroarilo puede estar sustituido o sin sustituir.

- 30 Como se usa en la presente memoria, "heteroalíclico" o "heteroalíclicilo" se refiere a un sistema de anillos monocíclico, bicíclico, y tricíclico de tres, cuatro, cinco, seis, siete, ocho, nueve, diez, hasta 18 miembros en donde los átomos de carbono junto con de 1 a 5 heteroátomos constituyen dicho sistema de anillos. Sin embargo, un heterociclo puede contener opcionalmente uno o más enlaces insaturados situados de un modo tal que no se produzca un sistema de electrones pi completamente deslocalizado en todos los anillos. Los heteroátomos se seleccionan independientemente entre oxígeno, azufre, nitrógeno. Un heterociclo puede contener además una o más funcionalidades carbonilo o tiocarbonilo, para hacer que la definición incluya sistemas oxo y sistemas tio tales como lactamas, lactonas, imidas cíclicas, tioimidas cíclicas, y carbamatos cíclicos. Cuando está compuesto por dos o más anillos, los anillos pueden estar unidos juntos en una forma condensada. Además, cualquier nitrógeno en un heteroalíclico puede estar cuaternarizado. Los grupos heteroalíclico o heteroalíclicilo pueden estar sin sustituir o sustituidos. Algunos ejemplos de tales grupos "heteroalíclico" o "heteroalíclicilo" incluyen, pero no se limitan a, 1,3-dioxina, 1,3-dioxano, 1,4-dioxano, 1,2-dioxolano, 1,3-dioxolano, 1,4-dioxolano, 1,3-oxatiano, 1,4-oxatiina, 1,3-oxatiolano, 1,3-ditiol, 1,3-ditiolano, 1,4-oxatiano, tetrahidro-1,4-tiazina, 2H-1,2-oxazina, maleimida, succinimida, ácido barbitúrico, ácido tiobarbitúrico, dioxopiperazina, hidantoína, dihidouracilo, trioxano, hexahidro-1,3,5-triazina, imidazolina, imidazolidina, isoxazolina, isoxazolidina, oxazolina, oxazolidina, oxazolidinona, tiazolina, tiazolidina, morfolina, oxirano, N-óxido de piperidina, piperidina, piperazina, pirrolidina, pirrolidona, pirrolidiona, 4-piperidona, pirazolina, pirazolidina, 2-oxopirrolidina, tetrahidropirano, 4H-pirano, tetrahidrotiopirano, tiamorfolina, sulfóxido de tiamorfolina, tiamorfolina sulfona, y sus análogos benzocondensados (por ejemplo, benzoimidazolidinona, tetrahidroquinolina, 3,4-metilendioxfenilo).

Como se usa en la presente memoria, "aralquilo" y "aril(alquilo)" se refieren a un grupo arilo conectado, a modo de sustituyente, a través de un grupo alquileo inferior. El grupo alquileo inferior y el grupo arilo de un aralquilo pueden estar sustituidos o sin sustituir. Algunos ejemplos incluyen, pero no se limitan a, bencilo, 2-fenilalquilo, 3-fenilalquilo, y naftilalquilo.

5 Como se usa en la presente memoria, "heteroaralquilo" y "heteroaril(alquilo)" se refiere a un grupo heteroarilo conectado, a modo de sustituyente, a través de un grupo alquileo inferior. El grupo alquileo inferior y el grupo heteroarilo del heteroaralquilo pueden estar sustituidos o sin sustituir. Algunos ejemplos incluyen, pero no se limitan a, 2-tienilalquilo, 3-tienilalquilo, furilalquilo, tienilalquilo, pirrolilalquilo, piridilalquilo, isoxazolilalquilo, e imidazolilalquilo, y sus análogos benzocondensados.

10 Como se usa en la presente memoria, "alcoxi" se refiere a la fórmula -OR en donde R es un alquilo, un alqueno, un alquino, un cicloalquilo, un cicloalqueno o un cicloalquino como se definió anteriormente. Una lista no limitante de alcoxis es metoxi, etoxi, n-propoxi, 1-metiletoxi (isopropoxi), n-butoxi, iso-butoxi, sec-butoxi, y terc-butoxi. Un alcoxi puede estar sustituido o sin sustituir.

15 Como se usa en la presente memoria, un grupo "C-amido" se refiere a un grupo " $-C(=O)N(R_aR_b)$ " en el que R_a y R_b pueden ser independientemente hidrógeno, alquilo, alqueno, alquino, cicloalquilo, cicloalqueno, cicloalquino, arilo, heteroarilo, heteroalíclico, aralquilo, o (heteroalíclico)alquilo. Un C-amido puede estar sustituido o sin sustituir.

Como se usa en la presente memoria, un grupo "N-amido" se refiere a un grupo " $RC(=O)N(R_a)$ " en el que R y R_a pueden ser independientemente hidrógeno, alquilo, alqueno, alquino, cicloalquilo, cicloalqueno, cicloalquino, arilo, heteroarilo, heteroalíclico, aralquilo, o (heteroalíclico)alquilo. Un N-amido puede estar sustituido o sin sustituir.

20 La expresión "átomo de halógeno", y los términos "halógeno" o "halo", como se usan en la presente memoria, significan uno cualquiera de los átomos radioestables de la columna 7 de la tabla periódica de los elementos, tal como flúor, cloro, bromo, y yodo.

25 El término "amina" como se usa en la presente memoria se refiere a un grupo $-NH_2$ en donde uno o más hidrógenos pueden estar opcionalmente sustituidos con un grupo R. R puede ser independientemente hidrógeno, alquilo, alqueno, alquino, cicloalquilo, cicloalqueno, cicloalquino, arilo, heteroarilo, heteroalíclico, aralquilo, o (heteroalíclico)alquilo.

El término "aldehído" como se usa en la presente memoria se refiere a un grupo $-R_c-C(O)H$, en donde R_c puede estar ausente o se puede seleccionar independientemente entre alquileo, alqueno, alquino, cicloalqueno, cicloalqueno, cicloalqueno, arileno, heteroarileno, heteroalíclico, aralqueno, o (heteroalíclico)alqueno.

30 El término "amino" como se usa en la presente memoria se refiere a un grupo $-NH_2$.

El término "hidroxi" como se usa en la presente memoria se refiere a un grupo $-OH$.

El término "ciano" como se usa en la presente memoria se refiere a un grupo $-CN$.

El término "azido" como se usa en la presente memoria se refiere a un grupo $-N_3$.

El término "tiol" como se usa en la presente memoria se refiere a un grupo $-SH$.

35 El término "ácido carboxílico" como se usa en la presente memoria se refiere a $-C(O)OH$.

El término "tiocianato" como se usa en la presente memoria se refiere a un grupo $-S-C=N$.

40 El término "oxo-amina" como se usa en la presente memoria se refiere a un grupo $-O-NH_2$, en donde uno o más hidrógenos del $-NH_2$ pueden estar opcionalmente sustituidos con un grupo R. R puede ser independientemente hidrógeno, alquilo, alqueno, alquino, cicloalquilo, cicloalqueno, cicloalquino, arilo, heteroarilo, heteroalíclico, aralquilo, o (heteroalíclico)alquilo.

45 Como se usa en la presente memoria, un "nucleótido" incluye una base heterocíclica que contiene nitrógeno, un azúcar, y uno o más grupos fosfato. Son unidades monoméricas de una secuencia de ácido nucleico. En el ARN, el azúcar es una ribosa, y en el ADN una desoxirribosa, es decir, un azúcar que carece de un grupo hidroxilo que está presente en la ribosa. La base heterocíclica que contiene nitrógeno puede ser una base de purina o pirimidina. Las bases de purina incluyen adenina (A) y guanina (G), y los derivados modificados o los análogos de las mismas. Las bases de pirimidina incluyen citosina (C), timina (T), y uracilo (U), y los derivados modificados o los análogos de las mismas. El átomo C-1 de la desoxirribosa está unido al átomo N-1 de una pirimidina o al átomo N-9 de una purina.

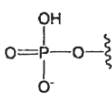
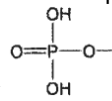
50 Como se usa en la presente memoria, un "nucleósido" es estructuralmente similar a un nucleótido, pero carece de los restos fosfato. Un ejemplo de un análogo de nucleósido sería uno en el que la etiqueta está unida a la base y no existe ningún grupo fosfato unido a la molécula de azúcar. El término "nucleósido" se usa en la presente memoria en su sentido habitual que se entiende por los expertos en la técnica. Algunos ejemplos incluyen, pero no se limitan a, un ribonucleósido que comprende un resto de ribosa y un desoxirribonucleósido que comprende un resto de

desoxirribosa. Un resto de pentosa modificado es un resto de pentosa en el que un átomo de oxígeno se ha reemplazado por un carbono y/o un carbono se ha reemplazado por un átomo de azufre o de oxígeno. Un "nucleósido" es un monómero que puede tener una base sustituida y/o un resto de azúcar. Además, un nucleósido se puede incorporar a polímeros y oligómeros de ADN y/o ARN de mayor tamaño.

- 5 La expresión "base de purina" se usa en la presente memoria en su sentido habitual que se entiende por los expertos en la técnica, e incluye sus tautómeros. De forma análoga, la expresión "base de pirimidina" se usa en la presente memoria en su sentido habitual que se entiende por los expertos en la técnica, e incluye sus tautómeros. Una lista no limitante de bases de purina opcionalmente sustituidas incluye purina, adenina, guanina, hipoxantina, xantina, aloxantina, 7-alquilguanina (por ejemplo, 7-metilguanina), teobromo, cafeína, ácido úrico e isoguanina.
- 10 Algunos ejemplos de bases de pirimidina incluyen, pero no se limitan a, citosina, timina, uracilo, 5,6-dihidrouracilo y 5-alquilitosina (por ejemplo, 5-metilcitosina).

Como se usa en la presente memoria, "derivado" o "análogo" significa un derivado de nucleótido o nucleósido sintético que tiene restos de base modificados y/o restos de azúcar modificados. Tales derivados y análogos se discuten, por ejemplo, en Scheit, *Nucleotide Analogs* (John Wiley & Son, 1980) y Uhlman *et al.*, *Chemical Reviews* 90:543-584, 1990. Los análogos de nucleótido también pueden comprender uniones fosfodiéster modificadas, incluyendo uniones fosforotioato, fosforoditioato, alquil-fosfonato, fosforanilidato y fosforamidato. "Derivado", "análogo" y "modificado", como se usan en la presente memoria, se pueden usar de forma intercambiable, y están incluidos por los términos "nucleótido" y "nucleósido" que se definen en la presente memoria.

Como se usa en la presente memoria, el término "fosfato" se usa en su sentido habitual que se entiende por los

- 20 expertos en la técnica, e incluye sus formas protonadas (por ejemplo  y ). Como se usa en la presente memoria, los términos "monofosfato", "difosfato", y "trifosfato" se usan en su sentido habitual que se entiende por los expertos en la técnica, e incluyen sus formas protonadas.

25 Las expresiones "grupo protector" y "grupos protectores" como se usan en la presente memoria se refieren a cualquier átomo o grupo de átomos que se añade a una molécula con el fin de prevenir que los grupos existentes en la molécula experimenten reacciones químicas no deseadas. En ocasiones, "grupo protector" y "grupo bloqueante" se pueden usar de forma intercambiable.

30 Como se usa en la presente memoria, los prefijos "foto" o "foto-" significan relacionado con la luz o la radiación electromagnética. El término puede incluir la totalidad o parte del espectro electromagnético que incluye, pero no se limita a, uno o más de los intervalos conocidos habitualmente como las partes de radio, microondas, infrarrojo, visible, ultravioleta, rayos X o rayos gamma del espectro. La parte del espectro puede ser una que se bloquea mediante una región de metal de una superficie tal como los metales que se exponen en la presente memoria. Además o alternativamente, la parte del espectro puede ser una que pasa a través de una región intersticial de una superficie tal como una región hecha de vidrio, plástico, sílice, u otro material que se expone en la presente memoria. En realizaciones particulares, se puede usar radiación que es capaz de pasar a través de un metal.

35 Además o alternativamente, se puede usar radiación que se enmascara por el vidrio, plástico, sílice, u otro material que se expone en la presente memoria.

40 Como se usa en la presente memoria, la expresión "ajuste de fase" se refiere a un fenómeno en SBS que está causado por la retirada incompleta de los terminadores y fluoróforos de 3', y el fallo en completar la incorporación de una parte de las cadenas de ADN en los conglomerados mediante las polimerasas en un ciclo de secuenciación dado. El preajuste de fase está causado por la incorporación de nucleótidos sin terminadores de 3' eficaces y el suceso de incorporación avanza 1 ciclo hacia delante. El ajuste y el preajuste de fase causan que las intensidades extraídas para un ciclo específico consistan en la señal del ciclo actual así como ruido de los ciclos precedente y siguiente. A medida que aumenta el número de ciclos, aumenta la fracción de secuencias por conglomerado afectada por ajuste de fase, dificultando la identificación de la base correcta. El preajuste de fase puede estar causado por la presencia de una cantidad traza de nucleótidos sin proteger o bloquear en 3'-OH durante la secuenciación por síntesis (SBS). Los nucleótidos sin proteger en 3'-OH se podrían generar durante el proceso de fabricación o posiblemente durante el almacenamiento y los procesos de manipulación de los reactivos. Por consiguiente, el descubrimiento de análogos de nucleótido que disminuyen la incidencia del preajuste de fase es sorprendente y proporciona una gran ventaja en las aplicaciones de SBS con respecto a los análogos de nucleótido existentes. Por ejemplo, los análogos de nucleótido provistos pueden dar como resultado un tiempo de ciclo de SBS más rápido, menores valores de ajuste y preajuste de fase, y mayor longitud de lectura de secuenciación.

Grupos protectores $-C(R)_2N_3$ de 3'-OH

55 Algunas realizaciones que se describen en la presente memoria se refieren a una molécula de nucleótido o nucleósido modificada que tiene un grupo protector $-C(R)_2N_3$ de 3'-hidroxi retirable, en donde R se selecciona entre el grupo que consiste en hidrógeno, $-C(R^1)_m(R^2)_n$, $-C(=O)OR^3$, $-C(=O)NR^4R^5$, $-C(R^6)_2O(CH_2)_pNR^7R^8$ y $-C(R^9)_2O-Ph-C(=O)NR^{10}R^{11}$, en donde R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 , R^{10} , R^{11} , m, n y p son como se definieron anteriormente.

En algunas realizaciones, uno de R es hidrógeno y el otro R es $-C(R^1)_m(R^2)_n$. En algunas de tales realizaciones, $-C(R^1)_m(R^2)_n$ se selecciona entre $-CHF_2$, $-CH_2F$, $-CHCl_2$ o $-CH_2Cl$. En una realización, $-C(R^1)_m(R^2)_n$ es $-CHF_2$. En otra realización, $-C(R^1)_m(R^2)_n$ es $-CH_2F$.

5 En algunas realizaciones, uno de R es hidrógeno y el otro R es $-C(=O)OR^3$. En algunas de tales realizaciones, R^3 es hidrógeno.

En algunas realizaciones, uno de R es hidrógeno y el otro R es $-C(=O)NR^4R^5$. En algunas de tales realizaciones, ambos R^4 y R^5 son hidrógeno. En algunas de tales otras realizaciones, R^4 es hidrógeno y R^5 es alquilo C_{1-6} . En algunas otras realizaciones más, ambos R^4 y R^5 son alquilo C_{1-6} . En una variante, R^5 es n-butilo. En otra variante, ambos R^4 y R^5 son metilo.

10 En algunas realizaciones, uno de R es hidrógeno y el otro R es $-C(R^6)_2O(CH_2)_pNR^7R^8$. En algunas de tales realizaciones, ambos R^6 son hidrógeno. En algunas de tales realizaciones, ambos R^7 y R^8 son hidrógeno. En algunas de tales realizaciones, p es 0. En alguna de tales otras realizaciones, p es 6.

15 En algunas realizaciones, uno de R es hidrógeno y el otro R es $-C(R^9)_2O-Ph-C(=O)NR^{10}R^{11}$. En algunas de tales realizaciones, ambos R^9 son hidrógeno. En algunas de tales realizaciones, ambos R^{10} y R^{11} son hidrógeno. En alguna de tales otras realizaciones, R^{10} es hidrógeno y R^{11} es un alquilo sustituido. En una realización, R^{11} es un alquilo sustituido con amino.

Desprotección de los grupos protectores de 3'-OH

20 En algunos aspectos, el grupo protector de 3'-OH se retira con una reacción de desprotección con una fosfina. El grupo azido en $-C(R)_2N_3$ se puede convertir en un grupo amino poniendo en contacto las moléculas de nucleótido o nucleósido modificadas con una fosfina. Alternativamente, el grupo azido en $-C(R)_2N_3$ se puede convertir en un grupo amino poniendo en contacto tales moléculas con los tioles, en particular tioles solubles en agua tales como ditiotreitil (DTT). En una variante, la fosfina es tris(hidroximetil)fosfina (THP). A menos que se indique de otro modo, la referencia a nucleótidos también se pretende que sea aplicable a nucleósidos.

Etiquetas detectables

25 Algunas realizaciones que se describen en la presente memoria se refieren al uso de etiquetas detectables convencionales. La detección se puede llevar a cabo mediante cualquier método adecuado, incluyendo espectroscopía de fluorescencia o mediante otros medios ópticos. La etiqueta preferente es un fluoróforo que, después de la absorción de energía, emite radiación a una longitud de onda definida. Se conocen numerosas etiquetas fluorescentes adecuadas. Por ejemplo, Welch *et al.* (*Chem. Eur. J.* 5(3):951-960, 1999) describen restos fluorescentes funcionalizados con dansilo que se usan en la presente invención. Zhu *et al.* (*Cytometry* 28:206-211, 1997) describen el uso de las etiquetas fluorescentes Cy3 y Cy5, que también se pueden usar en la presente invención. También se desvelan etiquetas adecuadas para su uso en Prober *et al.* (*Science* 238:336-341, 1987); Connell *et al.* (*BioTechniques* 5(4):342-384, 1987), Ansorge *et al.* (*Nucl. Acids Res.* 15(11):4593-4602, 1987) y Smith *et al.* (*Nature* 321:674, 1986). Otras etiquetas fluorescentes disponibles en el mercado incluyen, pero no se limitan a, fluoresceína, rodamina (incluyendo TMR, rojo Texas y Rox), alexa, bodipi, acridina, cumarina, pireno, benzantraceno y las cianinas.

30

35

También se pueden usar etiquetas múltiples en la presente solicitud, por ejemplo, casetes FRET de difluoróforo (*Tet. Let.* 46:8867-8871, 2000). También se pueden usar sistemas dendriméricos multifluor (*J. Am. Chem. Soc.* 123:8101-8108, 2001). Aunque son preferentes las etiquetas fluorescentes, otras formas de etiquetas detectables serán evidentes como útiles para los expertos habituales en la técnica. Por ejemplo, también se pueden usar micropartículas, incluyendo puntos cuánticos (Empodocles *et al.*, *Nature* 399:126-130, 1999), nanopartículas de oro (Reichrt *et al.*, *Anal. Chem.* 72:6025-6029, 2000) y microperlas (Lacoste *et al.*, *Proc. Natl. Acad. Sci USA* 97(17):9461-9466, 2000).

40

45 También se pueden usar etiquetas de múltiples componentes en la presente solicitud. Una etiqueta de múltiples componentes es la que es dependiente de la interacción con un compuesto adicional para su detección. La etiqueta de múltiples componentes más habitual usada en biología es el sistema de biotina-estreptoavidina. La biotina se usa como la etiqueta unida a la base del nucleótido. La estreptoavidina se añade a continuación por separado para permitir que se produzca la detección. Están disponibles otros sistemas de múltiples componentes. Por ejemplo, el dinitrofenol tiene un anticuerpo fluorescente disponible en el mercado que se puede usar para detección.

50 A menos que se indique de otro modo, la referencia a nucleótidos también se pretende que sea aplicable a nucleósidos. La presente solicitud también se describirá adicionalmente por referencia a ADN, aunque la descripción también será aplicable a ARN, APN, u otros ácidos nucleicos, a menos que se indique de otro modo.

Conectores

55 En algunas realizaciones que se describen en la presente memoria, la base de purina o pirimidina de las moléculas de nucleótido o nucleósido modificadas se pueden unir a una etiqueta detectable como se ha descrito anteriormente.

En algunas de tales realizaciones, los conectores usados son escindibles. El uso de un conector escindible asegura que la etiqueta se pueda retirar, si se requiere, después de la detección, evitando cualquier señal que interfiera con cualquier nucleótido o nucleósido etiquetado incorporado posteriormente.

5 En algunas de otras realizaciones, los conectores que se usan son no escindibles. Dado que en cada caso en donde se incorpora un nucleótido etiquetado de la invención no se necesita que se incorpore posteriormente ningún nucleósido y de ese modo la etiqueta no necesita retirarse del nucleósido.

10 Los expertos en la técnica serán conscientes de la utilidad de los trifosfatos de didesoxinucleósido en los denominados métodos de secuenciación de Sanger, y los protocolos relacionados (de tipo Sanger), que se basan en la terminación de cadena randomizada para un tipo particular de nucleótido. Un ejemplo de un protocolo de secuenciación de tipo Sanger es el método BASS que se describe por Metzker.

15 Los métodos de Sanger y de tipo Sanger operan en general mediante la realización de un experimento en el que se proporcionan ocho tipos de nucleótidos, cuatro de los cuales contienen un grupo 3'-OH; y cuatro de los cuales omiten el grupo OH y que están etiquetados de forma diferente entre sí. Los nucleótidos usados que omiten el grupo 3'-OH - didesoxinucleótidos (ddNTP). Como conoce el experto en la técnica, dado que los ddNTP están etiquetados de forma diferente, mediante la determinación de las posiciones de los nucleótidos terminales incorporados, y la combinación de esta información, se puede determinar la secuencia del oligonucleótido diana.

20 Los nucleótidos de la presente solicitud, como se reconocerá, pueden ser de utilidad en los métodos de Sanger y los protocolos relacionados dado que se puede conseguir el mismo efecto conseguido mediante el uso de ddNTP mediante el uso de los grupos protectores de 3'-OH que se describen en la presente memoria: ambos previenen la incorporación de los nucleótidos posteriores.

Además, se ha de entender que la monitorización de la incorporación de nucleótidos protegidos en 3'-OH se puede determinar mediante el uso de ^{32}P radiactivo en los grupos fosfato unidos. Estos pueden estar presentes en los propios ddNTP o en los cebadores que se usan para la extensión.

25 Los conectores escindibles se conocen en la técnica, y se puede aplicar química convencional para unir un conector a una base de nucleótido y una etiqueta. El conector se puede escindir mediante cualquier método adecuado, incluyendo exposición a ácidos, bases, nucleófilos, electrófilos, radicales, metales, agentes reductores u oxidantes, luz, temperatura, enzimas, etc. El conector que se discute en la presente memoria también se puede escindir con el mismo catalizador que se usa para escindir el enlace 3'-O-grupo protector. Se pueden adaptar conectores adecuados para grupos protectores químicos convencionales, como se describe en Greene & Wuts, Protective Groups in Organic Synthesis, John Wiley & Sons. Algunos conectores escindibles adecuados adicionales que se usan en síntesis en fase sólida se desvelan en Guillier *et al.* (Chem. Rev. 100:2092-2157, 2000).

35 El uso de la expresión "conector escindible" no pretende implicar que se requiera que se retire la totalidad del conector, por ejemplo, de la base de nucleótido. Cuando la etiqueta detectable está unida a la base, el sitio de escisión del nucleósido se puede situar en una posición en el conector que asegure que una parte del conector permanezca unida a la base de nucleótido después de la escisión.

40 Cuando la etiqueta detectable está unida a la base, el conector puede estar unido en cualquier posición en la base de nucleótido con la condición de que se pueda seguir llevando a cabo el emparejamiento de bases de Watson-Crick. En el contexto de bases de purina, es preferente que el conector esté unido a través de la posición 7 de la purina o el análogo de desazapurina preferente, a través de una purina modificada en 8, a través de una adenosina modificada en N-6 o una guanina modificada en N-2. Para las pirimidinas, la unión es preferentemente a través de la posición 5 en citosina, timidina o uracilo y la posición N-4 en la citosina.

A. Conectores escindidos con electrófilos

45 Los conectores escindidos con electrófilos se escinden por lo general mediante protones e incluyen escisiones sensibles a ácidos. Los conectores adecuados incluyen los sistemas bencílicos modificados tales como tritilo, ésteres de p-alcoxibencilo y p-alcoxibencil amidas. Otros conectores adecuados incluyen grupos terc-butiloxycarbonilo (Boc) y el sistema acetal.

El uso de metales tiófilos, tales como níquel, plata o mercurio, en la escisión de grupos protectores que contienen tioacetal u otros compuestos de azufre también se puede considerar para la preparación de moléculas de conector adecuadas.

50 B. Conectores escindidos con nucleófilos

La escisión nucleofílica también es un método bien reconocido en la preparación de moléculas de conector. Se pueden usar grupos tales como ésteres que son lábiles en agua (es decir, se pueden escindir simplemente a pH básico) y grupos que son lábiles para nucleófilos no acuosos. Se pueden usar iones fluoruro para escindir enlaces silicio-oxígeno en grupos tales como triisopropil silano (TIPS) o t-butildimetil silano (TBDMS).

C. Conectores fotoescindibles

Los conectores fotoescindibles se han usado ampliamente en la química de los carbohidratos. Es preferente que la luz requerida para activar la escisión no afecte a otros componentes de los nucleótidos modificados. Por ejemplo, si se usa un fluoróforo como la etiqueta, es preferente que este absorba luz en una longitud de onda diferente de la requerida para escindir la molécula de conector. Los conectores adecuados incluyen los que se basan en compuestos de O-nitrobencilo y compuestos de nitroveratrilo. También se pueden usar conectores basados en la química de la benzoína (Lee *et al.*, *J. Org. Chem.* 64:3454-3460, 1999).

D. Escisión en condiciones reductoras

Existen numerosos conectores conocidos que son susceptibles a escisión reductora. La hidrogenación catalítica que usa catalizadores basados en paladio se ha usado para escindir los grupos bencilo y benciloxicarbonilo. La reducción del enlace disulfuro también se conoce en la técnica.

E. Escisión en condiciones oxidantes

Los enfoques basados en oxidación se conocen bien en la técnica. Estos incluyen la oxidación de grupos p-alcoxibencilo y la oxidación de conectores de azufre y selenio. El uso de yodo acuoso para escindir disulfuros y otros conectores basados en azufre o selenio también está dentro del alcance de la invención.

F. Conectores de captura de seguridad

Los conectores de captura de seguridad son los que se escinden en dos etapas. En un sistema preferente, la primera etapa es la generación de un centro nucleófilo reactivo seguido de una segunda etapa que implica una ciclación intramolecular que da como resultado la escisión. Por ejemplo, las uniones de éster levulínico se pueden tratar con hidrazina o fotoquímica para liberar una amina activa, que a continuación se puede ciclar para escindir un éster en otra parte de la molécula (Burgess *et al.*, *J. Org. Chem.* 62:5165-5168, 1997).

G. Escisión mediante mecanismos de eliminación

También se pueden usar reacciones de eliminación. Por ejemplo, se puede usar la eliminación catalizada con base de grupos tales como Fmoc y cianoetilo, y la eliminación reductora catalizada con paladio de sistemas alílicos.

En algunas realizaciones, el conector puede comprender una unidad de espaciador. La longitud del conector se proporciona sin importar que la etiqueta se mantenga a una distancia suficiente del nucleótido para que no interfiera con cualquier interacción entre el nucleótido y una enzima.

En algunas realizaciones, el conector puede consistir en una funcionalidad similar a la del grupo protector de 3'-OH. Esto hará que la desprotección y el proceso de desprotección sean más eficaces, ya que solo se requerirá un tratamiento individual para retirar tanto la etiqueta como el grupo protector. Los conectores particularmente preferentes son conectores que contienen azida escindibles con fosfina.

Métodos de secuenciación

Los nucleósidos o nucleótidos modificados que se describen en la presente memoria se pueden usar junto con una diversidad de técnicas de secuenciación. En algunas realizaciones, el proceso para determinar la secuencia de nucleótidos de un ácido nucleico diana puede ser un proceso automatizado.

Los análogos de nucleótido que se presentan en la presente memoria se pueden usar en un procedimiento de secuenciación, tal como una técnica de secuenciación por síntesis (SBS). En resumen, SBS se puede iniciar poniendo en contacto los ácidos nucleicos diana con uno o más nucleótidos etiquetados, ADN polimerasa, etc. Las características donde un cebador se extiende usando el ácido nucleico diana como plantilla incorporarán un nucleótido etiquetado que se puede detectar. Opcionalmente, los nucleótidos etiquetados pueden incluir además una propiedad de terminación reversible que termina además la extensión del cebador una vez se ha añadido un nucleótido a un cebador. Por ejemplo, un análogo de nucleótido que tiene un resto de terminador reversible se puede añadir a un cebador de un modo tal que no se pueda producir la extensión posterior hasta que se suministre un agente de desbloqueo para retirar el resto. De ese modo, para realizaciones que usan terminación reversible, se puede suministrar un agente de desbloqueo para la célula de flujo (antes o después de que se produzca la detección). Se pueden llevar a cabo lavados entre las diversas etapas de suministro. El ciclo se puede repetir a continuación n veces para extender el cebador en n nucleótidos, detectando de ese modo una secuencia de longitud n. Algunos procedimientos, sistemas fluidos y plataformas de detección de SBS a modo de ejemplo que se pueden adaptar con facilidad para su uso con una agrupación producida mediante los métodos de la presente divulgación se describen, por ejemplo, en Bentley *et al.*, *Nature* 456:53-59 (2008), los documentos de Patente WO 04/018497; WO 91/06678; WO 07/123744; los documentos de Patente de Estados Unidos con números 7.057.026; 7.329.492; 7.211.414; 7.315.019 o 7.405.281, y el documento de Publicación de Solicitud de Patente de Estados Unidos n.º 2008/0108082 A1.

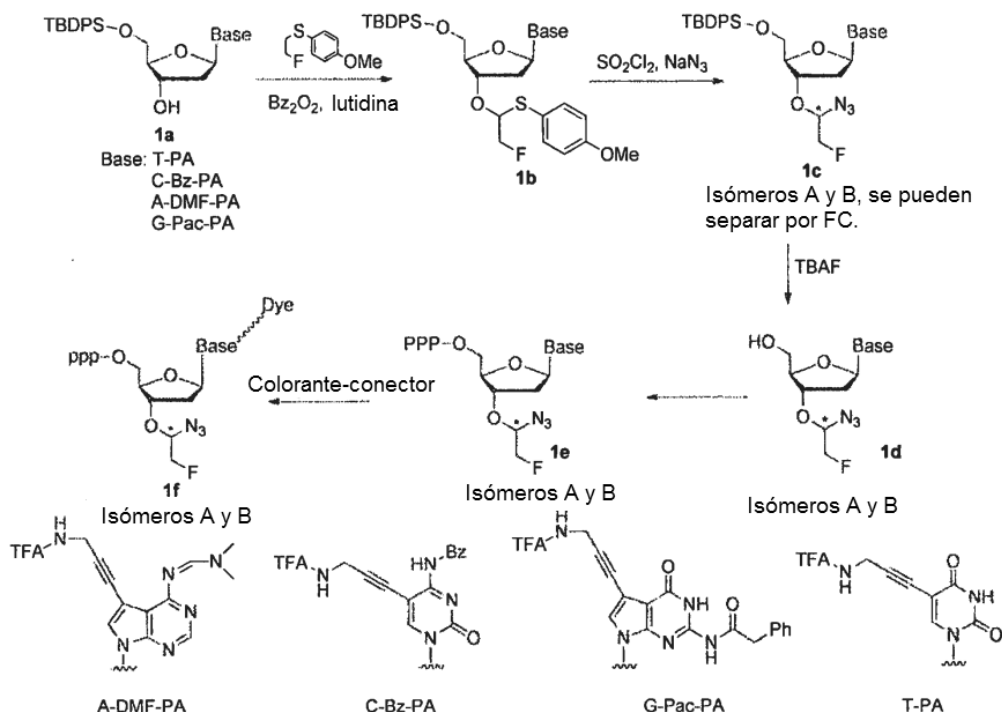
- Se pueden usar otros procedimientos de secuenciación que usan reacciones cíclicas, tales como pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) a medida que se incorporan nucleótidos particulares a una cadena de ácido nucleico naciente (Ronaghi, *et al.*, *Analytical Biochemistry* 242(1), 84-9 (1996); Ronaghi, *Genome Res.* 11(1), 3-11 (2001); Ronaghi *et al.* *Science* 281(5375), 363 (1998); documentos de Patente de Estados Unidos con números 6.210.891; 6.258.568 y 6.274.320). En la pirosecuenciación, se puede detectar el PPi liberado mediante su conversión en trifosfato de adenosina (ATP) mediante ATP sulfurilasa, y el ATP resultante se puede detectar a través de fotones producidos por luciferasa. De ese modo, la reacción de secuenciación se puede monitorizar a través de un sistema de detección de luminiscencia. Las fuentes de radiación de excitación que se usan para sistemas de detección basados en fluorescencia no son necesariamente para procedimientos de pirosecuenciación. Algunos sistemas fluidos, detectores y procedimientos útiles que se pueden usar para la aplicación de la pirosecuenciación a las agrupaciones de la presente divulgación se describen, por ejemplo, en el documento de Solicitud de Patente WIPO con número de serie PCT/US11/57111, el documento de Publicación de Solicitud de Patente de Estados Unidos n.º 2005/0191698 A1, el documento de Patente de Estados Unidos n.º 7.595.883, y el documento de Patente de Estados Unidos n.º 7.244.559.
- También son útiles las reacciones de secuenciación por ligadura que incluyen, por ejemplo, las que se describen en Shendure *et al.* *Science* 309:1728-1732 (2005); el documento de Patente de Estados Unidos n.º 5.599.675; y el documento de Patente de Estados Unidos n.º 5.750.341. Algunas realizaciones pueden incluir los procedimientos de secuenciación por hibridación que se describen, por ejemplo, en Bains *et al.*, *Journal of Theoretical Biology* 135(3), 303-7 (1988); Drmanac *et al.*, *Nature Biotechnology* 16, 54-58 (1998); Fodor *et al.*, *Science* 251(4995), 767-773 (1995); y el documento de Patente WO 1989/10977. En los procedimientos tanto de secuenciación por ligadura como de secuenciación por hibridación, los ácidos nucleicos que están presentes en paredes que contienen gel (u otros rasgos cóncavos) se someten a ciclos repetidos de suministro y detección de oligonucleótidos. Los sistemas fluidos para los métodos de SBS que se exponen en la presente memoria, o en las referencias que se citan en la presente memoria, se pueden adaptar con facilidad para el suministro de reactivos para los procedimientos de secuenciación por ligadura o secuenciación por hibridación. Por lo general, los oligonucleótidos están etiquetados fluorescentemente y se pueden detectar usando detectores de fluorescencia similares a los que se describen con respecto a los procedimientos de SBS en la presente memoria o en las referencias que se citan en la presente memoria.
- Algunos aspectos pueden utilizar métodos que implican la monitorización en tiempo real de la actividad de la ADN polimerasa. Por ejemplo, se pueden detectar las incorporaciones de nucleótidos a través de interacciones de transferencia de energía por resonancia de fluorescencia (FRET) entre una polimerasa que porta fluoróforo y nucleótidos etiquetados con γ -fosfato, o con guías de onda en modo cero. Se describen técnicas y reactivos para secuenciación basada en FRET, por ejemplo, en Levene *et al.* *Science* 299, 682-686 (2003); Lundquist *et al.* *Opt. Lett.* 33, 1026-1028 (2008); Korlach *et al.* *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008).
- Algunas variantes de SBS incluyen la detección de un protón liberado tras la incorporación de un nucleótido a un producto de extensión. Por ejemplo, la secuenciación basada en la detección de protones liberados puede usar un detector eléctrico y las técnicas asociadas que están disponibles en el mercado en Ion Torrent (Guilford, CT, una filial de Life Technologies) o los métodos y sistemas de secuenciación que se describen en los documentos de Publicación de Solicitud de Patente de Estados Unidos con números 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; o 2010/0282617 A1.

Ejemplos

Se desvelan realizaciones adicionales con mayor detalle en los siguientes ejemplos, que en ningún modo se debe interpretar que limitan el alcance de las reivindicaciones. Las síntesis de diversos nucleótidos modificados con el grupo 3'-hidroxil protegido se muestran en los Ejemplos 1-3.

Ejemplo 1

Síntesis de nucleótidos con grupo protector de 3'-OH



Esquema 1

- El Esquema 1 ilustra una ruta sintética para la preparación de los nucleótidos modificados con azidometilo sustituido con monofluorometilo como grupos protectores de 3'-OH. Los compuestos **1a-1f** emplean una timina modificada (T-PA) como base. Otros ejemplos no limitantes de las bases que se pueden usar incluyen Cbz-PA, ADMF-PA, y GPac-PA, cuyas estructuras se han mostrado anteriormente en el Esquema 1.

Procedimientos experimentales

- A una solución del nucleósido **1a** de partida (1,54 g, 2,5 mmol) en CH₃CN anhidro (25 ml) se añadió 2,6-lutidina (0,87 ml, 7,5 mmol), (2-fluoroetil)(4-metoxifenil)sulfano (MPSF) (3,26 g, 17,5 mmol) y a continuación Bz₂O₂ (50 % de pureza, 8,47 g, 17,5 mmol) a 4 °C. La mezcla de reacción se dejó calentar lentamente hasta la temperatura ambiente. La mezcla se agitó durante otras 6 horas. Se monitorizó por TLC (EtOAc:DCM = 2:8 v/v) para ver el consumo completo del nucleósido de partida. A continuación, la reacción se concentró a presión reducida hasta un residuo aceitoso. A esta mezcla, se añadió éter de petróleo (500 ml) y se agitó vigorosamente durante 10 min. La fase de éter de petróleo se decantó y se repitió el tratamiento del residuo con éter de petróleo (x 2). El residuo oleoso se repartió entre DCM/NaHCO₃ (1:1) (300 ml). La fase orgánica se separó y fase acuosa se extrajo adicionalmente en DCM (2 x 150 ml). Las fases orgánicas combinadas se secaron sobre MgSO₄, se filtraron y los compuestos volátiles se evaporaron a presión reducida. El producto en bruto **1c** se purificó mediante una columna de gel de sílice Biotag (50 g) usando un gradiente de éter de petróleo a éter de petróleo:EtOAc 1:1 (v/v) para proporcionar 1,63 g del nucleósido **1b** en forma de una espuma de color amarillo pálido (diastereómeros, 82 % de rendimiento). RMN ¹H (DMSO d₆, 400 MHz): δ, 0,95 (s, 9H, tBu), 2,16 - 2,28 (m, 2H, H-2'), 3,67 (s, OMe), 3,65 -3,85 (m, 2H, HH-5'), 3,77 (dd, J = 11,1, 4,5 Hz, 1H, HH-5'), 3,95-3,98 (m, 1H, H-4'), 4,04 (m, 2H, CH₂F), 4,63-4,64 (m, 1H, H-3'), 5,01-5,32 (s, 1H, CH), 6,00 (m, 1H, H-1'), 6,72-6,87 (m, 3H, Ar), 7,35-7,44 (m, 7H, Ar), 7,55-7,60 (m, 4H, Ar), 7,88 (s, 1H, H-6), 9,95 (t a, 1H, NH), 11,70 (s, 1H, NH).
- A una solución del nucleósido **1b** de partida (1,14 g, 1,4 mmol) en CH₂Cl₂ anhidro (14 ml) con tamices moleculares (4 Å) en atmósfera de N₂ se añadió ciclohexeno (1,44 ml, 14 mmol). La mezcla se enfrió con un baño de hielo seco/acetona a -78 °C. Se añadió lentamente una solución de cloruro de sulfurilo (580 µl, 7,2 mmol) en DCM (14 ml) durante 90 minutos en atmósfera de N₂. Después de 20 min a esa temperatura TLC (EtOAc: éter de petróleo = 1:1 v/v) indicó el consumo completo del nucleósido de partida. Los compuestos volátiles se evaporaron a presión reducida (y una temperatura ambiente de 25 °C) y el residuo aceitoso se sometió rápidamente a alto vacío durante un periodo adicional de 10 minutos hasta que se formó una espuma. El producto en bruto se purgó con N₂ y a continuación se disolvió en DMF anhidra (5 ml) y NaN₃ (470 mg, 7 mmol) añadido de una vez. La suspensión

resultante se agitó a temperatura ambiente durante 2 horas o hasta que TLC indicó la finalización de la reacción y la formación de **1c** en forma de dos isómeros (a y b). La mezcla de reacción se repartió entre EtOAc:NaHCO₃ (1:1) (200 ml). La fase orgánica se separó y la fase acuosa se extrajo adicionalmente con EtOAc (2 x 100 ml). Los extractos orgánicos combinados se secaron sobre MgSO₄, se filtraron y los compuestos volátiles se evaporaron a presión reducida. Los dos diaestereoisómeros de **1c** (A y B) se separaron mediante una columna de gel de sílice Biotag (25 g) usando un gradiente de éter de petróleo a éter de petróleo: EtOAc 1:1 (v/v) en forma de una espuma de color amarillo.

Isómero A (370 mg, rendimiento: 38 %). RMN ¹H (DMSO d₆, 400 MHz): δ 1,02 (s, 9H, tBu), 2,35 - 2,43 (m, 2H, H-2'), 3,76-3,80 (m, 1H, H-5'), 3,88 - 3,92 (m, 1H, H-5'), 4,10 - 4,12 (m, 1H, H-4'), 4,14 (d, *J* = 4,1 Hz 2H, NHCH₂), 4,46-4,60 (m, 3H, H-3', CH₂F), 5,05-5,09 (m, 1H, CHN₃), 6,11 (t, *J* = 6,1 Hz, 1H, H-1'), 7,47 - 7,51 (m, 6H, Ar), 7,64 - 7,68 (m, 4H, Ar), 7,97 (s, 1H, H-6), 10,03 (t a, 1H, *J* = 10,0 Hz, NH), 11,76 (s, 1H, NH). RMN ¹⁹F: -74,3 (CF₃), -230,2 (CH₂F).

Isómero B (253 mg, rendimiento: 26 %). RMN ¹H (DMSO d₆, 400 MHz): δ 1,01 (s, 9H, tBu), 2,38 - 2,42 (m, 2H, H-2'), 3,74-3,78 (m, 1H, H-5'), 3,86-3,90 (m, 1H, H-5'), 4,00-4,05 (m, 1H, H-4'), 4,12 (d, *J* = 4,1 Hz 2H, NHCH₂), 4,45-4,60 (m, 3H, H-3', CH₂F), 5,00-5,14 (m, 1H, CHN₃), 6,09 (t, *J* = 6,1 Hz, 1H, H-1'), 7,41 - 7,50 (m, 6H, Ar), 7,63-7,66 (m, 4H, Ar), 7,95 (s, 1H, H-6), 10,01 (s a, 1H, NH), 11,74 (s, 1H, NH). RMN ¹⁹F: -74,5 (CF₃), -230,4 (CH₂F).

El material de partida **1c** (isómero A) (500 mg, 0,71 mmol) se disolvió en THF (3 ml) y se enfrió a 4 °C en un baño de hielo. A continuación, se añadió lentamente TBAF (1,0 M en THF, 5 % en peso de agua, 1,07 ml, 1,07 mmol) durante un periodo de 5 min. La mezcla de reacción se calentó lentamente hasta la temperatura ambiente. El progreso de la reacción se controló por TLC (éter de petróleo: EtOAc 3:7 (v/v)). La reacción se detuvo después de 1 hora cuando ya no había más material de partida visible por TLC. La solución de reacción se disolvió en EtOAc (50 ml) y se añadió a NaHCO₃ (60 ml). Las dos fases se separaron y la fase acuosa se extrajo con una cantidad adicional de DCM (50 ml x 2). Las extracciones orgánicas se combinaron, se secaron (MgSO₄), se filtraron, y se evaporaron para dar un aceite de color amarillo. El producto en bruto **1d** (isómero A) se purificó mediante una columna de gel de sílice Biotag (10 g) usando un gradiente de éter de petróleo: EtOAc 8:2 (v/v) a EtOAc en forma de un sólido de color blanco (183 mg, rendimiento: 56 %).

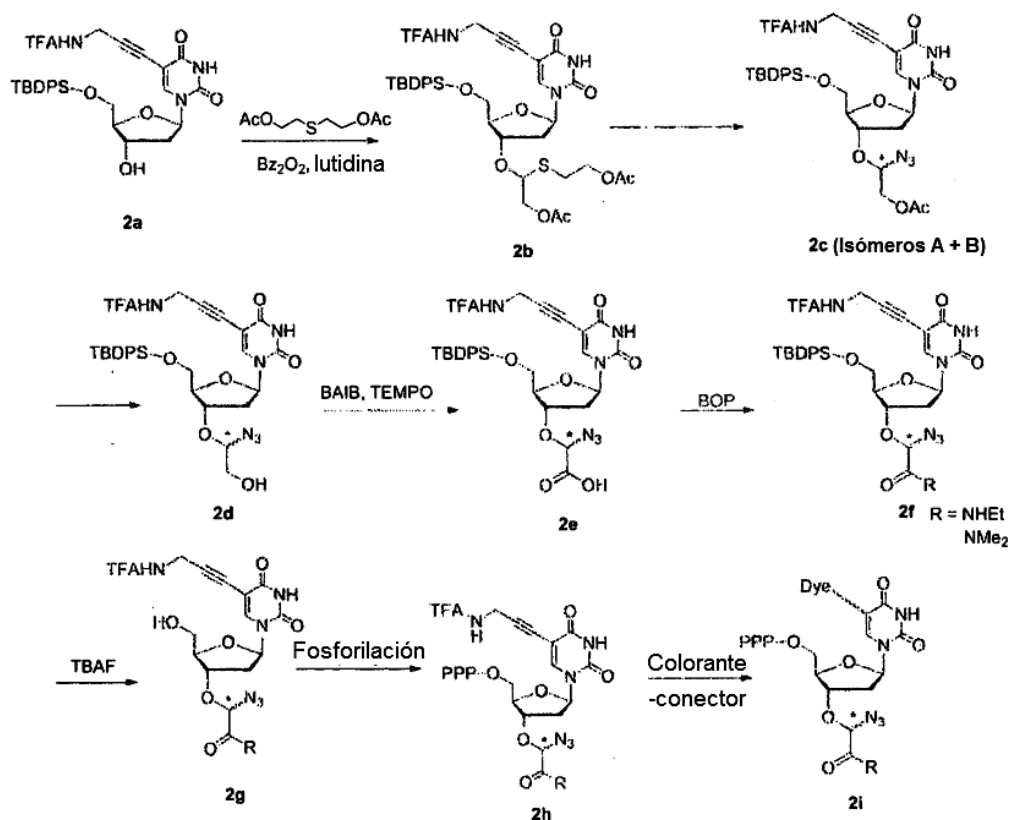
Isómero A: RMN ¹H (400 MHz, DMSO d₆): δ 2,24-2,35 (m, 2H, H-2'), 3,56-3,66 (m, 2H, H-5'), 3,96-4,00 (m, 1H, H-4'), 4,23 (s, 2H, CH₂NH), 4,33-4,37 (m, 1H, H-3'), 4,43-4,51 (m, CH₂F), 5,12 (s a, 1H, CHN₃), 5,23 (s a, 1H, 5'-OH), 6,07 (t, *J* = 6,7 Hz, 1H, H-1'), 8,26 (s, 1H, H-6), 10,11 (s a, 1H, NH), 11,72 (s a, 1H, NH). RMN ¹⁹F: -74,3 (CF₃), -230,5 (CH₂F).

Se llevó a cabo la misma reacción para **1c** (isómero B) a una escala de 360 mg y proporcionó el correspondiente producto **1d** (isómero B, 150 mg, 63 %). RMN ¹H (400 MHz, DMSO d₆): δ 2,24-2,37 (m, 2H, H-2'), 3,57-3,70 (m, 2H, H-5'), 3,97-4,01 (m, 1H, H-4'), 4,23 (s a, 2H, CH₂NH), 4,33-4,37 (m, 1H, H-3'), 4,44-4,53 (m, CH₂F), 5,11-5,21 (s a, 1H, CHN₃), 5,23 (s a, 1H, 5'-OH), 6,07 (t, *J* = 6,6 Hz, 1H, H-1'), 8,23 (s, 1H, H-6), 10,09 (s a, 1H, NH), 11,70 (s a, 1H, NH). RMN ¹⁹F: -74,1 (CF₃), -230,1 (CH₂F).

La preparación de los correspondientes trifosfatos **1e** y la unión adicional del colorante a la base nitrogenada para permitir el trifosfato de nucleósido completamente funcional (ffN) **1f** se ha informado en el documento de Patente WO 2004/018497 y se conoce en términos generales por el experto en la técnica.

Ejemplo 2

Síntesis de nucleótidos con el grupo protector de 3'-OH



Esquema 2

El Esquema 2 ilustra una ruta sintética para la preparación de los nucleótidos modificados con azidometilo sustituido con C-amido como grupos protectores de 3'-OH. Los compuestos **2a-2i** emplean una timina modificada (T-PA) como base. Otros ejemplos no limitantes de las bases que se pueden usar incluyen Cbz-PA, ADMF-PA, y GPac-PA, cuyas estructuras se han mostrado anteriormente en el Esquema 1. En el procedimiento experimental, se informan el compuesto **2f** con un grupo protector azidometilo (R = NMe₂) sustituido con N,N-dimetil-C(=O)- y las reacciones posteriores. También se prepararon compuestos con otros grupos C-amido, tales como N-etil-C(=O)- (R = NHEt).

10 Procedimientos experimentales

A una solución del nucleósido de partida **2a** (4,27 g, 6,9 mmol) en CH₃CN anhidro (50 ml) se añadió 2,6-lutidina (2,4 ml, 20,7 mmol), S(CH₂CH₂OAc)₂ (12,2 g, 69 mmol) y a continuación Bz₂O₂ (50 % de pureza, 33,4 g, 69 mmol) a 4 °C. La mezcla de reacción se dejó calentar lentamente hasta la temperatura ambiente. La mezcla se agitó durante otras 12 horas. Se realizó monitorización por TLC (EtOAc:DCM = 4:6 v/v) para ver el consumo completo del nucleósido de partida. A continuación, la reacción se concentró a presión reducida hasta un residuo aceitoso. A esta mezcla, se añadió éter de petróleo (800 ml) y se agitó vigorosamente durante 10 min. La fase de éter de petróleo se decantó y el residuo se trató repetidamente con éter de petróleo (x 2). El residuo oleoso se repartió a continuación entre DCM/NaHCO₃ (1:1) (1000 ml). La fase orgánica se separó y la fase acuosa se extrajo además en DCM (2 x 500 ml). Las fases orgánicas combinadas se secaron sobre MgSO₄, se filtraron y los compuestos volátiles se evaporaron a presión reducida. El producto en bruto **2b** se purificó mediante una columna de gel de sílice Biotag (100 g) usando un gradiente de éter de petróleo a éter de petróleo: EtOAc 2:8 (v/v) en forma de una espuma de color amarillo pálido (4,17 g, rendimiento: 74 %, diaestereoisómeros).

A una solución del nucleósido de partida **2b** (4,54 g, 5,56 mmol) en CH₂Cl₂ anhidro (56 ml) con tamices moleculares (4 Å) en atmósfera de N₂ se añadió ciclohexeno (5,62 ml, 56 mmol). La mezcla se enfrió con un baño de hielo a 4 °C. Se añadió lentamente una solución de cloruro de sulfurilo (1,13 ml, 13,9 mmol) en DCM (25 ml) durante 90 minutos en atmósfera de N₂. Después de 30 min a esa temperatura TLC (EtOAc: DCM = 4:6 v/v) indicó que quedaba un 10 % del nucleósido de partida **2b**. Se añadió una cantidad adicional de cloruro de sulfurilo (0,1 ml) a la mezcla de reacción. TLC indicó la conversión completa de **2b**. Los compuestos volátiles se evaporaron a presión reducida (y temperatura ambiente de 25 °C) y el residuo aceitoso se sometió rápidamente a alto vacío durante un periodo adicional de 10 minutos hasta que se formó una espuma. El producto en bruto se purgó con N₂ y a continuación se

5 disolvió en DMF anhidra (5 ml) y se añadió NaN_3 (1,8 g, 27,8 mmol) de una vez. La suspensión resultante se agitó a temperatura ambiente durante 2 horas o hasta que TLC indicó la finalización de la reacción y la formación de **2c** en forma de dos isómeros (A y B). La mezcla de reacción se repartió entre EtOAc: NaHCO_3 (1:1) (1000 ml). La fase orgánica se separó y la fase acuosa se extrajo además en EtOAc (2 x 300 ml). Los extractos orgánicos combinados se secaron a continuación sobre MgSO_4 , se filtraron y los compuestos volátiles se evaporaron a presión reducida. Los dos diaestereoisómeros **2c** (isómeros A y B) se separaron mediante una columna de gel de sílice Biotag (100 g) usando un gradiente de éter de petróleo a éter de petróleo: EtOAc 1:1 (v/v) en forma de una espuma de color amarillo pálido. Isómero A: 1,68 g, rendimiento: 40,7 %. Isómero B: 1,79 g, rendimiento: 43,2 %.

10 A una solución del nucleósido de partida **2c** (isómero A) (1,63 g, 2,2 mmol) en MeOH/THF (1:1) (20 ml) se añadió lentamente NaOH (1 M en agua) (2,2 ml, 2,2 mmol) y se agitó a 4 °C. El progreso de la reacción se controló por TLC (EtOAc: DCM = 4:6 v/v). La reacción se detuvo después de 1 hora cuando ya no hubo más material de partida visible por TLC. La mezcla de reacción se repartió entre DCM: NaHCO_3 (1:1) (150 ml). La fase orgánica se separó y la fase acuosa se extrajo además en DCM (2 x 70 ml). Los extractos orgánicos combinados se secaron sobre MgSO_4 , se filtraron y los compuestos volátiles se evaporaron a presión reducida. El producto en bruto **2d** se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de éter de petróleo: EtOAc (8:2) (v/v) a EtOAc en forma de una espuma de color amarillo pálido (1,1 g, rendimiento: 71 %).

Se repitió la misma reacción para **2c** (isómero B, 1,57 g) y proporcionó el correspondiente producto **2d** (isómero B, 1,01 g, 69 % de rendimiento).

20 A una solución del nucleósido de partida **2d** (isómero A) (700 mg, 1 mmol) en CH_3CN (10 ml) se trató con TEMPO (63 mg, 0,4 mmol) y BAIB (644 mg, 2 mmol) a temperatura ambiente. El progreso de la reacción se controló por TLC (EtOAc:DCM = 7:3 v/v). La reacción se detuvo después de 2 horas cuando ya no hubo más material de partida visible por TLC. La mezcla de reacción se repartió entre DCM: $\text{Na}_2\text{S}_2\text{O}_3$ (1:1) (100 ml). La fase orgánica se separó y la fase acuosa se extrajo además en DCM (2 x 70 ml). Los extractos orgánicos combinados se lavaron a continuación con NaCl (sat.). La fase orgánica se evaporó a presión reducida después de secado sobre MgSO_4 con el fin de prevenir que el producto precipitara. El producto en bruto **2e** se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de éter de petróleo: EtOAc (1:1) (v/v) a EtOAc a MeOH: EtOAc (1:9) en forma de una espuma de color amarillo pálido (isómero A, 482 mg, 68 % de rendimiento).

Se llevó a cabo la misma reacción para **2d** (isómero B, 700 mg) y proporcionó el correspondiente producto **2e** (isómero B, 488 mg, 69 % de rendimiento).

30 A una solución del nucleósido de partida **2e** (isómero A) (233 mg, 0,33 mmol) en CH_3CN (10 ml) se añadieron base de Hunig (173 μl , 1 mmol) y BOP (165 mg, 0,39 mmol) a temperatura ambiente. Después de agitar durante 5 min, la solución se trató con Me_2NH (2 M en THF) (0,41 ml, 0,82 mmol). El progreso de la reacción se controló por TLC (MeOH: DCM = 1:9 v/v). La reacción se detuvo después de 2 horas cuando ya no hubo más material de partida visible por TLC. La mezcla de reacción se repartió entre DCM: NaHCO_3 (1:1) (50 ml). La fase orgánica se separó y la fase acuosa se extrajo además en DCM (2 x 30 ml). Los extractos orgánicos combinados se secaron sobre MgSO_4 , se filtraron y los compuestos volátiles se evaporaron a presión reducida. El producto en bruto **2f** (R = NMe_2) se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de DCM:EtOAc (8:2) (v/v) a EtOAc en forma de una espuma de color amarillo pálido (isómero A, 220 mg, 90 % de rendimiento).

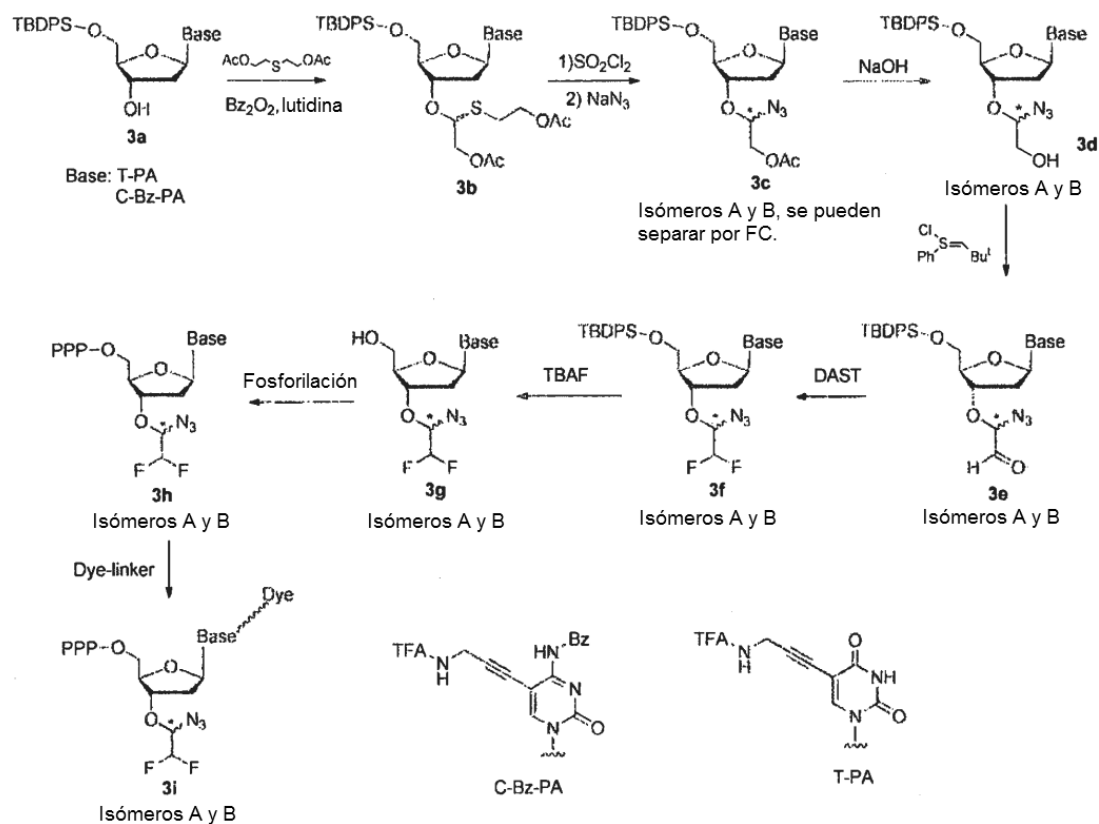
40 Se llevó a cabo la misma reacción para **2e** (isómero B, 249 mg) y proporcionó el correspondiente producto **2f** (isómero B, 240 mg, 92 % de rendimiento).

45 El material de partida **2f** (mezcla de los isómeros A y B) (455 mg, 0,61 mmol) se disolvió en THF (2 ml) y se enfrió a 4 °C con un baño de hielo. A continuación, se añadió lentamente TBAF (1,0 M en THF, 5 % en peso de agua, 1,0 ml, 1,0 mmol) durante un periodo de 5 min. La mezcla de reacción se calentó lentamente hasta la temperatura ambiente. El progreso de la reacción se controló por TLC (EtOAc). La reacción se detuvo después de 1 hora cuando ya no hubo más material de partida visible por TLC. La solución de reacción se disolvió en DCM (30 ml) y se añadió a NaHCO_3 (30 ml). Las dos fases se separaron y la fase acuosa se extrajo con una cantidad adicional de DCM (30 ml x 2). Las extracciones orgánicas se combinaron, se secaron (MgSO_4), se filtraron, y se evaporaron para dar un aceite de color amarillo. El producto en bruto **2g** se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de DCM: EtOAc 8:2 (v/v) a EtOAc a MeOH: EtOAc (2:8) en forma de un sólido de color blanco (52 % de rendimiento, 160 mg).

La preparación de los correspondientes trifosfatos **2h** y la unión adicional del colorante a la base nitrogenada para proporcionar el trifosfato de nucleósido completamente funcional (ffN) **2i** se han informado en el documento de patente WO 2004/018497 y se conocen en términos generales por el experto en la técnica.

Ejemplo 3

Síntesis de nucleótidos con un grupo protector de 3'-OH



Esquema 3

- 5 El Esquema 3 ilustra una ruta sintética para la preparación de nucleótidos modificados con los grupos protectores de 3'-OH azidometilo sustituido con difluorometilo. Los compuestos **3a-3i** emplean una timina modificada (T-PA) como base. Otros ejemplos no limitantes de las bases que se pueden usar incluyen Cbz-PA, ADMF-PA, y GPac-PA, cuyas estructuras se mostraron anteriormente en el Esquema 1. Procedimiento para la síntesis de **3b**, **3c** y **3d** se describieron en el Ejemplo 2.

Procedimientos experimentales

- 10 A una solución del nucleósido de partida **3d** (**isómero A**) (490 mg, 0,7 mmol) y DBU (209 μ l, 1,4 mmol) en DCM anhidro (5 ml) se añadió lentamente una solución de cloruro de N-terc-butil benceno sulfinimidoilo (181 mg, 0,84 mmol) en DCM anhidro (2 ml) a -78 °C. La mezcla de reacción se agitó durante 2 h a -78 °C. El progreso de la reacción se controló por TLC (EtOAc : DCM 4:6 v/v). La reacción se detuvo después de 2 horas cuando aún había un 10 % de material de partida restante por TLC, para prevenir la sobre-reacción. La mezcla de reacción se repartió entre DCM:NaHCO₃ (1:1) (50 ml). La fase acuosa se extrajo además en DCM (2 x 30 ml). Las extracciones orgánicas se combinaron, se secaron (MgSO₄), se filtraron, y se evaporaron para dar un aceite de color amarillo. El producto en bruto **3c** se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de éter de petróleo: EtOAc (8:2) (v/v) a éter de petróleo: EtOAc (2:8) (v/v) en forma de una espuma de color amarillo pálido (**isómero A**, 250 mg, 51 % de rendimiento).
- 15
- 20 Se llevó a cabo la misma reacción para **3d** (**isómero B**, 480 mg) y proporcionó el correspondiente producto **3e** (**isómero B**, 240 mg, 50 % de rendimiento).

- 25 A una solución del nucleósido de partida **3e** (**isómero A**) (342 mg, 0,49 mmol), se añadió lentamente EtOH (15 μ l, 0,25 mmol) en DCM (2,5 ml) a la solución de DAST (181 mg, 0,84 mmol) en DCM (2,5 ml) a 4 °C (baño de hielo). La mezcla de reacción se agitó durante 1 h a 4 °C. El progreso de la reacción se controló por TLC (EtOAc: éter de petróleo = 3:7 v/v). La reacción se detuvo después de 1 hora. La mezcla de reacción se repartió entre DCM: NaHCO₃ (1:1) (50 ml). La fase acuosa se extrajo además en DCM (2 x 30 ml). Las extracciones orgánicas se combinaron, se secaron (MgSO₄), se filtraron, y se evaporaron para dar un aceite de color amarillo. El producto en bruto **3f** se purificó mediante una columna de gel de sílice Bioatg (10 g) usando un gradiente de éter de petróleo: EtOAc (9:1) (v/v) a éter de petróleo: EtOAc (2:8) (v/v) en forma de una espuma de color amarillo pálido (**isómero A**, 100 mg, 28 %).
- 30

Se llevó a cabo la misma reacción para **3e** (isómero B, 480 mg) y proporcionó el correspondiente producto **3f** (isómero B, 240 mg, 50 % de rendimiento).

El material de partida **3f** (isómero A) (124 mg, 0,17 mmol) se disolvió en THF (2 ml) y se enfrió a 4 °C con un baño de hielo. A continuación, se añadió lentamente TBAF (1,0 M en THF, 5 % en peso de agua, 255 µl, 10,255 mmol) durante un periodo de 5 min. La mezcla de reacción se calentó lentamente hasta la temperatura ambiente. El progreso de la reacción se controló por TLC (EtOAc). La reacción se detuvo después de 1 hora cuando ya no hubo más material de partida visible por TLC. La solución de reacción se disolvió en DCM (30 ml) y se añadió a NaHCO₃ (30 ml). Las dos fases se separaron y la fase acuosa se extrajo con una cantidad adicional de DCM (30 ml x 2). Las extracciones orgánicas se combinaron, se secaron (MgSO₄), se filtraron, y se evaporaron para dar un aceite de color amarillo. El producto en bruto **3g** se purificó mediante una columna de gel de sílice Bioatg (4 g) usando un gradiente de DCM : EtOAc 8:2 (v/v) a EtOAc a MeOH: EtOAc (2:8) en forma de una espuma de color amarillo pálido (isómero A, 54 % de rendimiento, 44 mg).

Isómero A: RMN ¹H (400 MHz, DMSO d₆): δ 2,24-2,35 (m, 2H, H-2'), 3,56-3,66 (m, 2H, H-5'), 3,96-4,00 (m, 1H, H-4'), 4,23 (s, 2H, CH₂NH), 4,33-4,37 (m, 1H, H-3'), 4,85 (s, 2H, OCH₂N₃), 5,23 (t, J = 5,1 Hz, 1H, 5'-OH), 6,07 (t, J = 6,7 Hz, 1H, H-1'), 8,19 (s, 1H, H-6), 10,09 (s a, 1H, NH), 11,70 (s a, 1H, NH). RMN ¹⁹F: -74,4 (CF₃), -131,6 (CH₂F).

Se llevó a cabo la misma reacción para **3f** (isómero B, 133 mg) y proporcionó el correspondiente producto **3g** (isómero B, 48 mg, 54 % de rendimiento). RMN ¹H (400 MHz, DMSO d₆): δ 2,27-2,44 (m, 2H, H-2'), 3,58-3,67 (m, 2H, H-5'), 4,00-4,02 (m, 1H, H-4'), 4,24 (d, J = 4,1 Hz, 2H, CH₂NH), 4,57-4,58 (m, 1H, H-3'), 5,24-5,29 (m, 2H, 5'-OH, OCHN₃), 6,07-6,34 (m, 2H, H-1', CHF₂), 8,19 (s, 1H, H-6), 10,09 (s a, 1H, NH), 11,70 (s a, 1H, NH). RMN ¹⁹F: -74,2 (CF₃), -131,4 (CH₂F).

La preparación de los correspondientes trifosfatos **3h** y la unión adicional del colorante a la base nitrogenada para proporcionar el nucleótido completamente funcional (ffN) **3i** se han informado en el documento de Patente WO 2004/018497 y se conocen en términos generales por el experto en la técnica.

Ejemplo 4

25 Ensayo de estabilidad térmica de los grupos protectores de 3'-OH

Se investigaron una diversidad de grupos protectores de 3'-OH con respecto a su estabilidad térmica (**Figura 1A**). La estabilidad térmica se evaluó por calentamiento de cada nucleótido 3'-OH protegido 0,1 mM en un tampón a pH = 9 (tis-HCl 50 mM, NaCl 50 mM, tween al 0,05 %, Mg₂SO₄ 6 mM) a 60 °C. Se tomaron diversos puntos temporales y se usó HPLC para analizar la formación de materiales sin bloquear. Se descubrió que las estabilidades de -CH₂F y -C(O)NHBu era aproximadamente 2 veces mayor que la del grupo protector azidometilo estándar (-CH₂N₃). Se descubrió que la estabilidad del grupo -CF₂H era aproximadamente 10 veces mayor que la del estándar (**Figura 1B**).

Ejemplo 5

Desprotección de los grupos protectores de 3'-OH

También se estudiaron las velocidades de reacción de desprotección de varios grupos protectores 3'-OH. La velocidad desprotección del grupo protector azidometilo estándar se comparó con el azidometilo sustituido con -CH₂F y el azidometilo sustituido con -C(O)NHBu. Se observó que los dos grupos de bloqueo de 3'-OH térmicamente más estables se retiraron más rápido que el grupo protector azidometilo estándar usando fosfinas (THP 1 mM) como agente de desprotección. Véase la **Figura 2A**. Por ejemplo, las semividas de -CH₂F y -C(O)NHBu fueron de 8,9 minutos y 2,9 minutos respectivamente, en comparación con la semivida de 20,4 minutos del azidometilo (**Figura 2B**).

Ejemplo 6

Ensayo de secuenciación

Se prepararon nucleótidos modificados con el grupo protector de 3'-OH azidometilo sustituido con -CH₂F (mono-F) y se evaluó su rendimiento de secuenciación en plataformas Miseq. Se previó que la estabilidad térmica aumentada de los grupos protectores de 3'-OH conduciría a una mayor calidad de los nucleótidos para química de secuenciación con menos nucleótidos 3'-sin bloquear contaminados. Por lo tanto, la presencia de nucleótidos 3'-sin bloquear en los kits de secuenciación de SBS daría como resultado sucesos de preajuste de fase, que se numeraron como valores de preajuste de fase.

En primer lugar se usaron experimentos de secuenciación cortos de 12 ciclos para generar los valores de ajuste y preajuste de fase. Se usaron ffN protegidos con azidometilo sustituido con mono-F de acuerdo con las siguientes concentraciones: ffA-colorante 1 (2 µM); ffT-colorante 2 (10 µM), ffC-colorante 3 (2 µM) y ffG-colorante 4 (5 µM). El grupo azidometilo sustituido con mono-F comprende ambos isómeros A y B. Se usaron dos colorantes - el colorante 2 como en los kits de Miseq estándar y el colorante 5 para etiquetar ffT. La Tabla 1 muestra diversas combinaciones de nucleótidos con los isómeros A y B de azidometilo sustituido con mono-F que se evaluaron con respecto a los

impactos de ajuste y preajuste de fase. En todos los casos, los valores de preajuste de fase fueron básicamente menores que el control que usó nucleótidos de los kits V2 Miseq estándares (**Figura 3**).

Tabla 1.

Muestra	Grupo protector de 3'-OH	Ajuste de fase (%)	Preajuste de fase (%)
1	Control IMX Miseq V2 Est.	0,119	0,177
2	Mono-F-isómero A	0,11	0,085
3	Mono-F-isómero A (ffT-Colorante 5)	0,076	0,032
4	Mono-F-isómero A (A, C y G) + Mono-F-B-ffT-Colorante 5	0,095	0,083
5	Mono-F-A (G, ffT-Colorante 5) y Mono-F-B (A,C)	0,104	0,05
6	Mono-F-A (G, T) y Mono-F-B (A,C)	0,098	0,095
7	Control IMX Miseq V2 Est.	0,145	0,167

5 Ensayo de calidad de secuenciación

Se llevó a cabo la secuenciación de 2 x 400 pb en Miseq para evaluar el potencial de estos nucleótidos para la mejora de la calidad de secuenciación. El proceso de secuenciación se llevó a cabo de acuerdo con las instrucciones del fabricante (Illumina Inc., San Diego, CA). Se reemplazó el tampón de incorporación estándar con un tampón de incorporación que contenía todos los FFN bloqueados con mono-F, cada uno con una etiqueta de colorante distinta: ffA-colorante 1 (2 uM), ffT-colorante 2 (1 uM), ffC-colorante 3 (2 uM) y ffG-colorante 4 (5 uM). La biblioteca de ADN usada se realizó siguiendo el protocolo estándar TruSeq HT a partir del ADN genómico del *B. cereus*.

En ambos experimentos de secuenciación (con isómero A y B de bloqueo mono-F), se observaron valores muy bajos de preajuste de fase. Acoplada a los bajos valores de ajuste de fase, la aplicación de estos nuevos nucleótidos ha generado datos de secuenciación de 2 x 400 pb superiores con > 80 % de bases por encima de Q30 en ambos casos (véase la **Figura 4A** para la puntuación Q del isómero A y la **Figura 4B** para la gráfica de puntuación Q del isómero B). Estos resultados demuestran una gran mejora en comparación con los kits Miseq v2 (2 x 250 pb, 80 % de bases > Q30 en experimentos de secuenciación habituales de R&D, o 70 % de bases > Q30 que las especificaciones indicadas). Como se muestra a continuación, la Tabla 2 resume los datos de secuenciación cuando se usan todos los mono-F ffN-isómero A en la IMX. La Tabla 3 resume los datos de secuenciación usando todos los mono-F ffN-isómero B en la IMX.

Tabla 2.

Columna	Filas	Densidad (K/mm ²)	Agrupaciones PF (%)	Ajuste/preajuste de fase (%)	Lecturas (M)	Lecturas PF (M)	Tasa de desapareamiento % (PF)	% >= Q30	Rendimiento Total (G)
R1	28	690 +/- 14	93,1 +/-0,7	0,075 / 0,051	13,35	12,43	0,58 ± 0,11	89,7	5
R2	28	690 +/- 14	93,1 +/-0,7	0,092 / 0,078	13,35	12,43	1,31 ± 0,25	81,9	5
Total								85,8	9,9

Tabla 3.

Columna	Filas	Densidad (K/mm ²)	Agrupaciones PF (%)	Ajuste/preajuste de fase (%)	Lecturas (M)	Lecturas PF (M)	Tasa de desapareamiento % (PF)	% >= Q30	Rendimiento Total (G)
R1	28	816 +/- 9	92,7 +/-0,6	0,073 / 0,033	15,79	14,64	0,44 ± 0,11	91,2	5,9
R2	28	816 +/- 9	92,7 +/-0,6	0,078 / 0,059	15,79	14,64	1,03 ± 0,19	83,4	5,9
Total								87,3	11,7

25

REIVINDICACIONES

1. Molécula de nucleótido o nucleósido modificada que comprende una base de purina o pirimidina y un resto de azúcar de ribosa o desoxirribosa que tiene un grupo protector de 3'-hidroxi retirable que forma una estructura $-O-C(R)_2N_3$ unida covalentemente al átomo de carbono 3', en donde:
- 5 R se selecciona entre el grupo que consiste en hidrógeno, $-C(R^1)_m(R^2)_n$, $-C(=O)OR^3$, $-C(=O)NR^4R^5$, $-C(R^6)_2O(CH_2)_pNR^7R^8$ y $-C(R^9)_2O-Ph-C(=O)NR^{10}R^{11}$;
- cada R^1 y R^2 se selecciona independientemente entre hidrógeno, o halógeno;
- R^3 se selecciona entre hidrógeno o alquilo opcionalmente sustituido;
- 10 cada R^4 y R^5 se selecciona independientemente entre hidrógeno, alquilo opcionalmente sustituido, arilo opcionalmente sustituido, heteroarilo opcionalmente sustituido, o aralquilo opcionalmente sustituido;
- cada R^6 y R^9 se selecciona entre hidrógeno, alquilo opcionalmente sustituido o halógeno;
- cada R^7 , R^8 , R^{10} y R^{11} se selecciona independientemente entre hidrógeno, alquilo opcionalmente sustituido, arilo opcionalmente sustituido, heteroarilo opcionalmente sustituido, o aralquilo opcionalmente sustituido;
- m es un número entero de 0 a 3;
- 15 n es un número entero de 0 a 3; con la condición de que el total de $m + n$ es igual a 3; y
- p es un número entero de 0 a 6; con la condición de que
- R^1 y R^2 no pueden ser halógeno; y
- al menos un R no es hidrógeno.
- 20 2. La molécula de nucleótido o nucleósido modificada de la reivindicación 1, en donde uno de R es hidrógeno y el otro R es $-C(R^1)_m(R^2)_n$, y en donde $-C(R^1)_m(R^2)_n$ se selecciona entre $-CHF_2$, $-CH_2F$, $-CHCl_2$ o $-CH_2Cl$.
3. La molécula de nucleótido o nucleósido modificada de la reivindicación 1, en donde uno de R es hidrógeno y el otro R es $-C(=O)OR^3$, preferiblemente R^3 es hidrógeno.
4. La molécula de nucleótido o nucleósido modificada de la reivindicación 1, en donde uno de R es hidrógeno y el otro R es $-C(=O)NR^4R^5$, en donde cada uno de R^4 y R^5 se selecciona entre hidrógeno, o alquilo C_{1-6} .
- 25 5. La molécula de nucleótido o nucleósido modificada de la reivindicación 1, en donde uno de R es hidrógeno y el otro R es $-C(R^6)_2O(CH_2)_pNR^7R^8$, preferiblemente R^6 es hidrógeno, y cada R^7 y R^8 es hidrógeno.
6. La molécula de nucleótido o nucleósido modificada de la reivindicación 1, en donde uno de R es hidrógeno y el otro R es $-C(R^9)_2O-Ph-C(=O)NR^{10}R^{11}$, preferiblemente cada R^9 y R^{10} es hidrógeno, y R^{11} se selecciona entre hidrógeno o un alquilo sustituido con amino.
- 30 7. La molécula de nucleótido o nucleósido modificada de una cualquiera de las reivindicaciones 1 a 5, en donde el grupo protector de 3'-hidroxi se retira en una reacción de desprotección con una fosfina.
8. La molécula de nucleótido o nucleósido modificada de la reivindicación 7, en donde la fosfina es tris(hidroximetil)fosfina (THP).
9. La molécula de nucleótido o nucleósido modificada de una cualquiera de las reivindicaciones 1 a 8, en donde dicha base está unida a una etiqueta detectable a través de un conector escindible o un conector no escindible, y en donde la etiqueta detectable es un fluoróforo.
- 35 10. La molécula de nucleótido o nucleósido modificada de una cualquiera de las reivindicaciones 1 a 8, en donde dicho grupo protector de 3'-hidroxi está unido a una etiqueta detectable a través de un conector escindible o un conector no escindible, y en donde la etiqueta detectable es un fluoróforo.
- 40 11. Método de preparación de un polinucleótido en crecimiento complementario a un polinucleótido monocatenario diana en una reacción de secuenciación, que comprende incorporar una molécula de nucleótido modificada de una cualquiera de las reivindicaciones 1 a 10 al polinucleótido complementario en crecimiento, en donde la incorporación del nucleótido modificado previene la introducción de cualquier nucleótido posterior al polinucleótido complementario en crecimiento.
- 45 12. El método de la reivindicación 11, en donde la incorporación de la molécula de nucleótido modificada se consigue mediante una transferasa terminal, una polimerasa terminal o una transcriptasa inversa.

13. Método para determinar la secuencia de un polinucleótido monocatenario diana, que comprende
monitorizar la incorporación secuencial de nucleótidos complementarios, en donde al menos un nucleótido complementario incorporado es una molécula de nucleótido modificada de la reivindicación 9 o 10; y
detectar la identidad de la molécula de nucleótido modificada.
- 5 14. El método de la reivindicación 13, en donde la identidad del nucleótido modificado se determina por detección de la etiqueta detectable unida a la base.
15. Kit que comprende una pluralidad de moléculas de nucleótido o nucleósido modificadas de una cualquiera de las reivindicaciones 1 a 10, y que se envasa con materiales para las mismas.

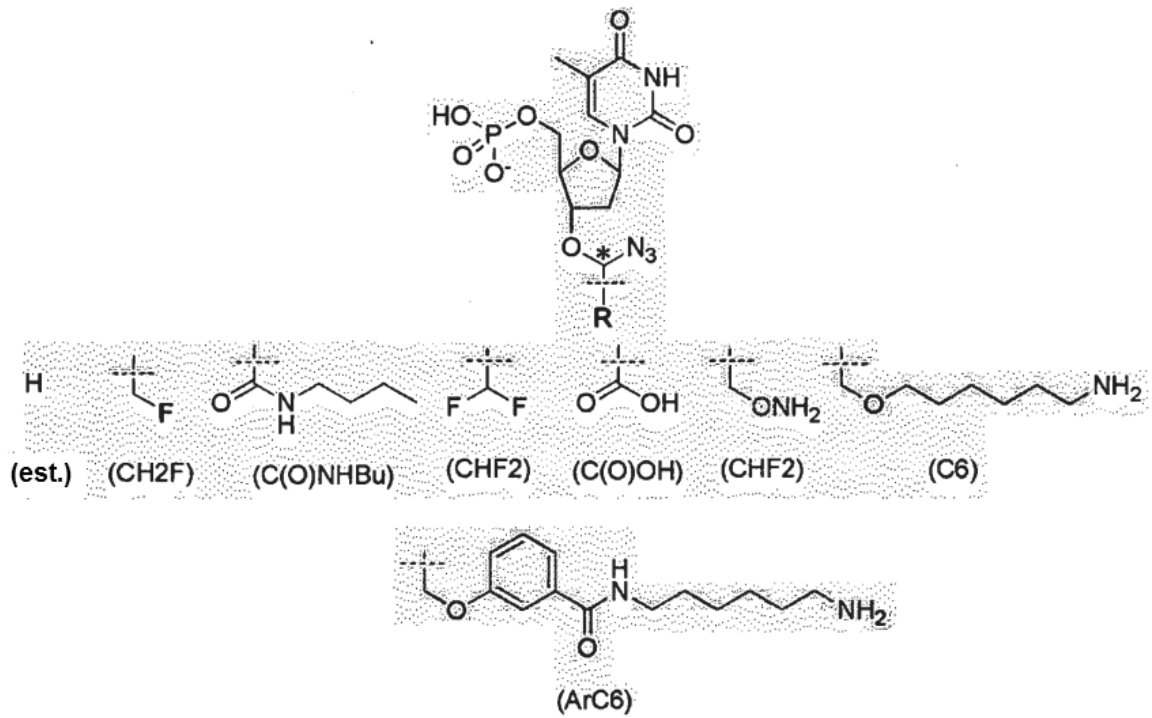


FIG. 1A

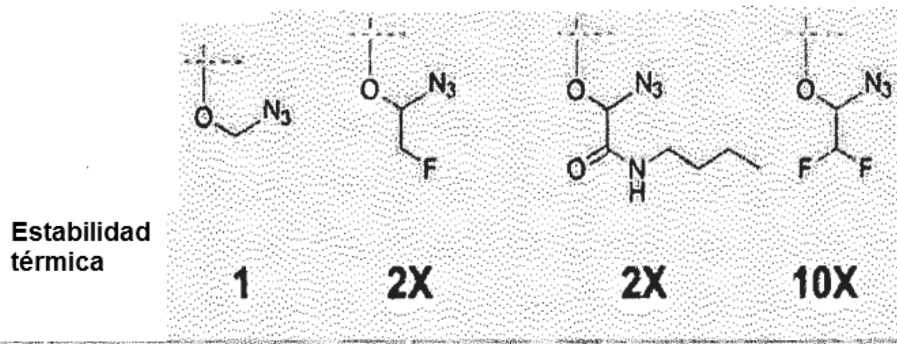


FIG. 1B

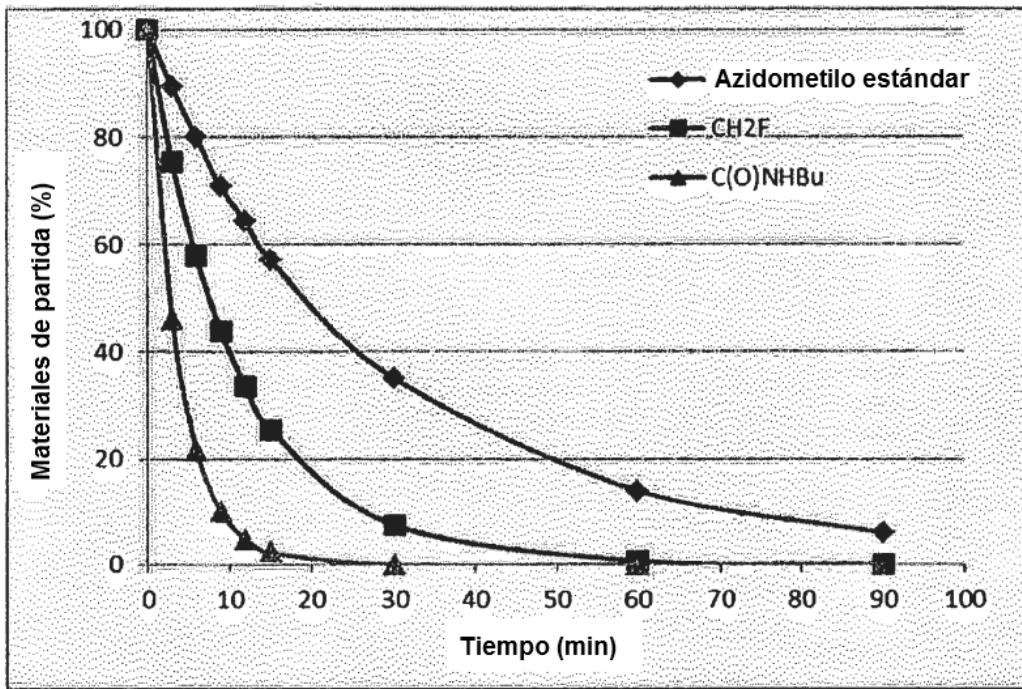


FIG. 2A

	Semivida (min)	Veces de mejora
Est.	20,4	1
-CH ₂ F	8,9	2,3
-C(O)NHBu	2,9	7

FIG. 2B

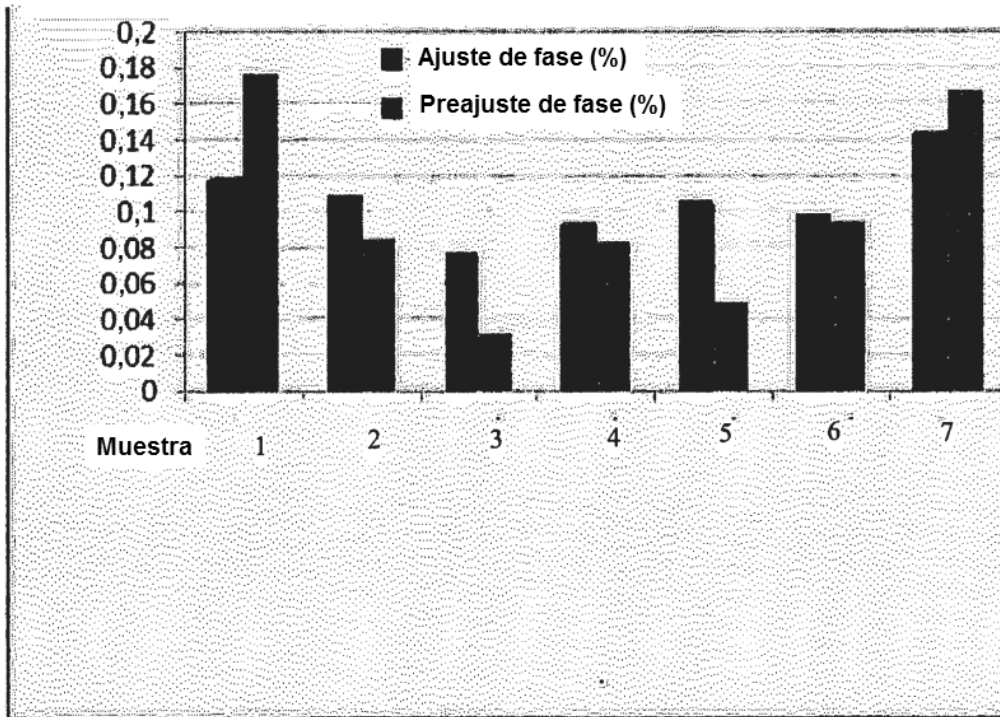


FIG. 3

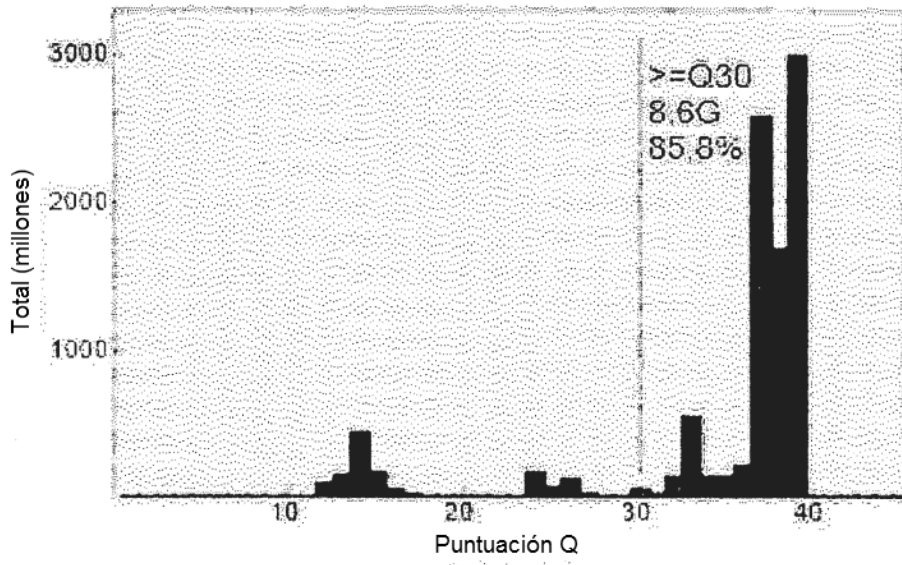


FIG. 4A

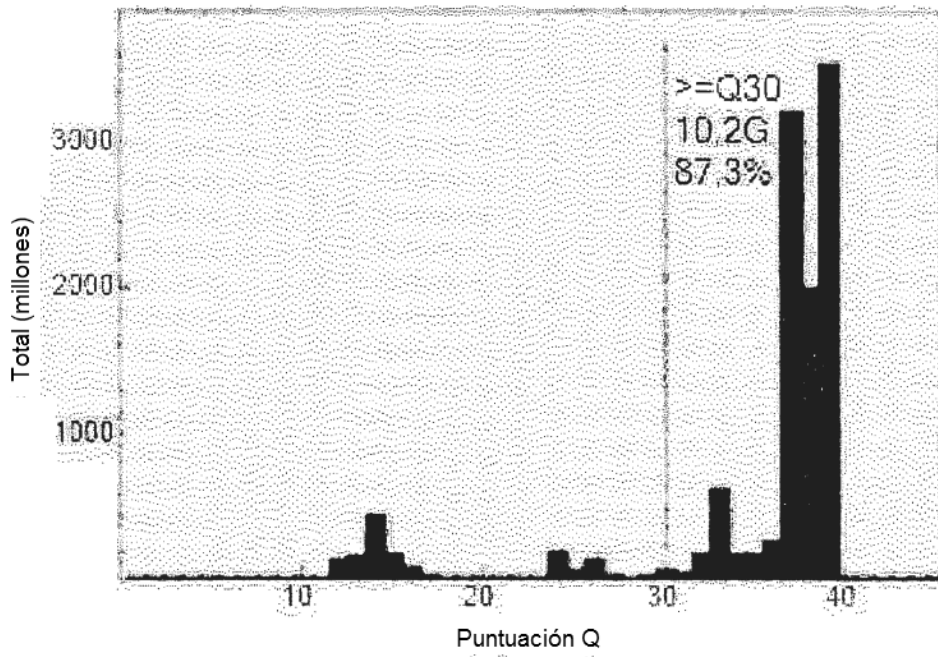


FIG. 4B