(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 773 438**

(51) Int. CI.:

C07H 19/14 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **06.08.2015 PCT/GB2015/052282**

(87) Fecha y número de publicación internacional: **11.02.2016 WO16020691**

(96) Fecha de presentación y número de la solicitud europea: **06.08.2015 E 15750100 (8)**

(97) Fecha y número de publicación de la concesión europea: **18.12.2019 EP 3177634**

(54) Título: **Enlazadores de nucleótidos modificados**

(30) Prioridad:

08.08.2014 GB 201414098

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
13.07.2020

(73) Titular/es:

ILLUMINA CAMBRIDGE LIMITED (100.0%)
19 Granta Park, Great Abington
Cambridge CB21 6DF, GB

(72) Inventor/es:

WU, XIAOLIN y
LIU, XIAOHAI

(74) Agente/Representante:

ELZABURU, S.L.P

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

Enlazadores de nucleótidos modificados

Campo

Algunas realizaciones de la presente solicitud se refieren a nuevos enlazadores de nucleótidos o nucleósidos para el aumento de la incorporación de nucleótidos en la secuenciación de ADN y otras aplicaciones de diagnóstico, por ejemplo, la secuenciación por síntesis.

Antecedentes

Se han llevado a cabo avances en el estudio de las moléculas, en parte por la mejora de las tecnologías utilizadas 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 utilizadas para el análisis de secuencias y el estudio de sucesos de hibridación.

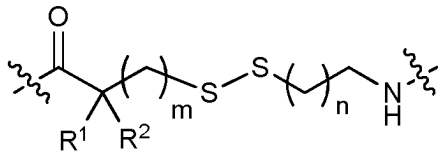
Un ejemplo de las tecnologías que han mejorado el estudio de los ácidos nucleicos es el desarrollo de matrices fabricadas de ácidos nucleicos inmovilizados. Estas matrices tienen normalmente una matriz de polinucleótidos de alta densidad inmovilizada sobre un material sólido de soporte. Véase, p. ej., Fodor *et al., Trends Biotech.* 12: 19-26, 1994, que describen formas de ensamblar distintos ácidos nucleicos utilizando una superficie de vidrio químicamente sensibilizada protegida por una máscara, pero expuestos en áreas definidas para permitir la unión de nucleótidos fosforamiditas adecuadamente modificados. Las matrices fabricadas también se pueden fabricar mediante la técnica de "aplicación puntual" de polinucleótidos conocidos sobre un soporte sólido en posiciones predeterminadas (p. ej., Stimpson *et al.*, *Proc. Natl. Acad. Sci.* 92: 6379-6383, 1995).

Una forma de determinar la secuencia de nucleótidos de un ácido nucleico unido a una matriz se denomina "secuenciación por síntesis" o "SPS". Esta técnica para determinar la secuencia de nucleótidos del ADN idealmente precisa la incorporación controlada (es decir, uno a la vez) del nucleótido complementario correcto opuesto al ácido nucleico que se está secuenciando. Esto permite una secuenciación precisa mediante la adición de nucleótidos en múltiples ciclos a medida que cada resto nucleotídico se secuencia uno a la vez, evitando así la incorporación de una serie no controlada de nucleótidos. Cada nucleótido incorporado se lee utilizando un marcador apropiado unido al mismo, antes de eliminar la fracción de marcador, y el siguiente ciclo de secuenciación posterior. El documento WO 02/088381 describe nucleósidos o nucleótidos unidos covalentemente a un fluoróforo a través de un enlazador.

Por consiguiente, en el contexto de las reacciones de secuenciación de ácidos nucleicos, sería conveniente poder aumentar la velocidad de incorporación de nucleótidos durante la secuenciación por síntesis, de forma que se pueda mejorar la eficacia del método de secuenciación.

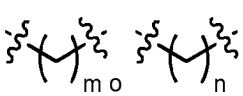
Compendio

Algunas realizaciones descritas en la presente memoria se refieren a un nucleósido o nucleótido unido covalentemente a un fluoróforo a través de un enlazador, en donde dicho enlazador comprende una estructura de fórmula (I):



(I),

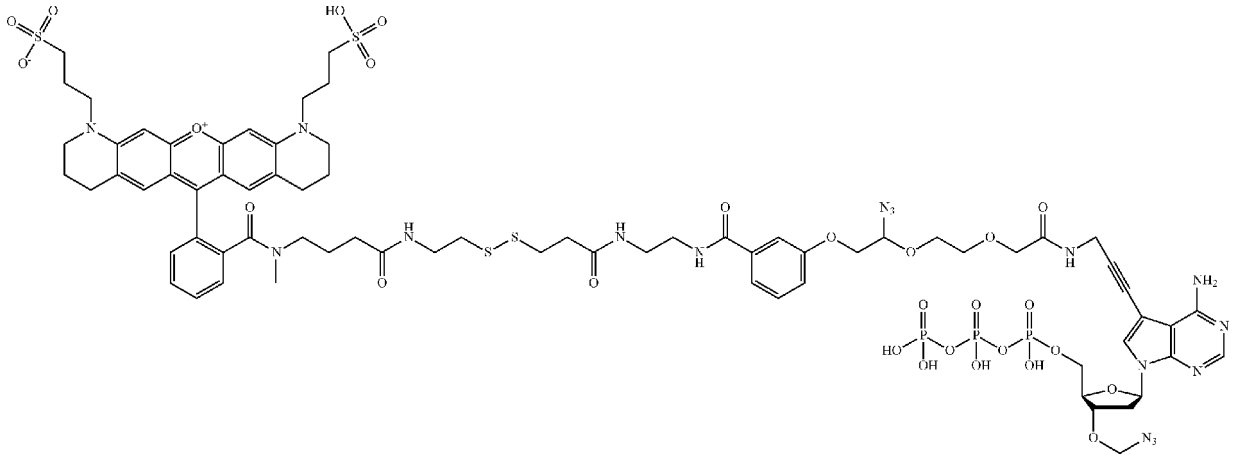
cada R^1 es alquilo C_{1-6} opcionalmente sustituido; y R^2 es alquilo C_{1-6} opcionalmente sustituido;

cada una de las unidades de repetición de metileno en  está opcionalmente sustituida;

m es un número entero de 0 a 20; y

n es un número entero de 1 a 20.

En algunas realizaciones, el nucleósido o nucleótido marcado con fluoróforo comprende una estructura de fórmula (I) que no tiene la estructura:



Algunas realizaciones descritas en la presente memoria se refieren a kits que comprenden un nucleósido o nucleótido marcado que comprende un enlazador entre el fluoróforo y el nucleósido o nucleótido, en donde el enlazador comprende una estructura de fórmula (I).

Algunas realizaciones descritas en la presente memoria se refieren a reactivos para modificar un nucleósido o un nucleótido que comprende un fluoróforo y un enlazador, en donde el enlazador comprende una estructura de fórmula (I).

Algunas realizaciones descritas en la presente memoria se refieren a métodos para la detección de un nucleósido que se ha incorporado en un polinucleótido, que comprenden: (a) incorporar un nucleósido o nucleótido marcado que comprende un enlazador en un polinucleótido; y (b) detectar una señal fluorescente procedente de dicho nucleósido o nucleótido marcado que se incorporó en la etapa (a), en donde el enlazador comprende una estructura de fórmula (I). En algunas realizaciones, el método comprende además: proporcionar una cadena de ácido nucleico molde y una cadena de ácido nucleico parcialmente hibridada, en donde la etapa (a) incorpora en la cadena hibridada al menos un nucleósido o nucleótido que es complementario con un nucleósido o nucleótido en la posición correspondiente de la cadena molde, y en donde la etapa (b) identifica la base del nucleósido o nucleótido incorporado, indicando de este modo la identidad del nucleósido o nucleótido complementario de la cadena molde.

Algunas realizaciones descritas en la presente memoria se refieren a métodos de secuenciación de una molécula de ácido nucleico molde, que comprenden: incorporar uno o más nucleótidos marcados en una cadena de ácido nucleico complementaria con el ácido nucleico molde; determinar la identidad de la base presente en uno o más nucleótidos marcados incorporados para determinar la secuencia de la molécula de ácido nucleico molde; en donde la identidad de la base presente en el uno o más nucleótidos marcados se determina detectando una señal fluorescente producida por dichos nucleótidos marcados; y en donde al menos un nucleótido marcado incorporado comprende un enlazador como se describe anteriormente, en donde el enlazador comprende una estructura de fórmula (I). En algunas realizaciones, la identidad de la base presente en uno o más nucleótidos se determina después de cada etapa de incorporación de nucleótidos.

Breve descripción de los dibujos

La **FIG. 1A** ilustra una estructura de un grupo de enlace parcial de un nucleótido marcado convencional.

La **FIG. 1B** ilustra el nucleótido marcado de la **FIG. 1A** con dos posibles enlazadores **125** y **130** a insertar en el grupo de enlace convencional de la **FIG. 1A.**

FIG. 2 demuestra una representación de la velocidad de incorporación de nucleótidos utilizando el nucleótido marcado de la **FIG. 1A** y los nucleótidos marcados modificados de la **FIG. 1B.**

Las **FIG. 3A-3E** ilustran las fórmulas estructurales de enlazadores adicionales no según la presente invención a insertar en el grupo de enlace convencional de la **FIG. 1A.**

La **FIG. 4** muestra una tabla de datos de una ejecución de secuenciación de dos colorantes utilizada para evaluar el efecto del inserto **125** de la **FIG.1B** y el inserto **315** de la **FIG. 3B** sobre la calidad de la secuenciación.

Las **FIG. 5A** y **5B** muestran una representación de la velocidad de errores para la lectura 1 y una representación de la velocidad de errores para la lectura 2 de la ejecución de secuenciación de la **FIG. 4** utilizando los insertos enlazadores **125** y **315.**

La **FIG. 6** muestra una tabla de datos de una ejecución de secuenciación utilizada para evaluar el efecto del inserto

125 de la **FIG.1B** y el inserto **310** de la **FIG. 3B** sobre la calidad de la secuenciación.

Las **FIG. 7A** y **7B** muestran una representación de la velocidad de errores para la lectura 1 y una representación de la velocidad de errores para la lectura 2 de la ejecución de secuenciación de la **FIG. 6** utilizando el inserto enlazador **125.**

La **FIG. 8A** muestra un ejemplo de un enlazador LN_3 convencional.

Las **FIG. 8B, 8C** y **8D** muestran tres ejemplos de estructuras modificadas del enlazador LN_3 de la **FIG. 8A.**

La **FIG. 8E** ilustra la inserción de una fracción protectora en el enlazador de la **FIG. 8D.**

La **FIG 9A** es un cromatograma que muestra la aparición de una impureza en un ffA con enlazador SS. La **FIG 9B** es una tabla que compara la estabilidad de los ffA con el enlazador SS y el enlazador AEDI. La **FIG 9C** es un cromatograma que muestra una comparación de los ffA con enlazador SS y enlazador AEDI en IMX 60° durante 22 horas, nuevamente mostrando una impureza con el enlazador SS.

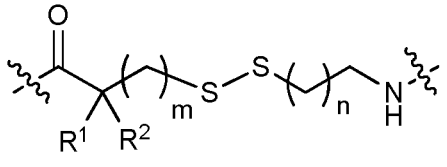
Las FIG. 10A, 10B y 10C muestran el aumento inesperado de la velocidad de incorporación de nucleótidos en solución con los cambios de enlazador. La **FIG 10A** es un gráfico que muestra la velocidad de incorporación a 1 uM, mostrando **FIG 10B** los resultados tabulados. La **FIG 10C** muestra de forma esquemática los enlazadores AEDI y SS con NR550S0.

La **FIG 11A** muestra diagramas de dispersión para combinaciones de V10 con distintos A-550S0 (misma concentración). La **FIG 11B** muestra la Kcat de enlazadores FFA en solución.

Las **FIGS 12A y 12B** muestran los parámetros de secuenciación sobre M111, Human550, 2x151 ciclos.

Descripción detallada de las realizaciones preferidas

Algunas realizaciones descritas en la presente memoria se refieren a un nucleósido o nucleótido unido covalentemente a un fluoróforo a través de un enlazador, en donde el enlazador comprende una estructura de la siguiente fórmula (I), en donde las definiciones de las variables se definen anteriormente.



(I).

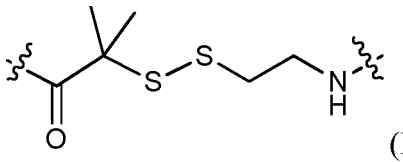
En algunas otras realizaciones, R^1 es alquilo C_{1-6} opcionalmente sustituido. En algunas de tales realizaciones, R^1 es metilo.

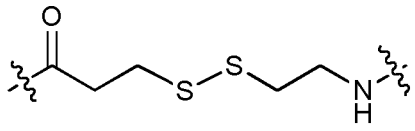
En cualquiera de las realizaciones de R^1, como se describe en la presente memoria, de fórmula (I), R^2 es alquilo C_{1-6} opcionalmente sustituido. En algunas de tales realizaciones, R^2 es metilo. En una realización, R^1 y R^2 son metilo ambos.

En algunas de las realizaciones de la estructura de fórmula (I), m es 0. En algunas otras realizaciones, m es 1.

En algunas de las realizaciones de la estructura de fórmula (I), n es 1.

En algunas de las realizaciones de la estructura de fórmula (I), la estructura de fórmula (I) también se puede representar por la fórmula (Ia) o (Ib):



(Ia),


(Ib).

En algunas realizaciones descritas en la presente memoria, la fórmula (Ia) se denomina "AEDI" y la fórmula (Ib) se denomina "SS".

En cualquiera de las realizaciones del nucleósido o nucleótido marcado con fluoróforo a través de un enlazador que comprende una estructura de fórmula (I) como se describe en la presente memoria, el nucleósido o nucleótido se puede unir al lado izquierdo del enlazador, ya sea directamente o a través de una fracción de enlace adicional.

En cualquiera de las realizaciones descritas en la presente memoria con respecto a un enlazador que comprende una estructura de fórmula (I), cuando el término "opcionalmente sustituido" se utiliza para definir una variable, tal variable puede no estar sustituida.

Definiciones

A menos que se definan de otro modo, todos los términos técnicos y científicos que se utilizan en la presente memoria tienen el mismo significado que entiende habitualmente un experto en la técnica. El uso del término "que incluye", así como otras formas, tales como "incluyen", "incluye" e "incluido", no es limitante. El uso del término "que tiene", así como otras formas, tales como "tienen", "tiene" y tuvo", no es limitante. Como se emplea en esta memoria descriptiva, ya sea en una expresión de transición o en el cuerpo de la reivindicación, los términos "comprende (comprenden)" y "que comprende" deben interpretarse como que tienen un significado abierto. Es decir, los términos anteriores se deben interpretar como sinónimos de las frases "que tiene al menos" o "que incluye al menos". Por ejemplo, cuando se utiliza en el contexto de un procedimiento, el término "que comprende" significa que el procedimiento incluye al menos las etapas indicadas, pero puede incluir etapas adicionales. Cuando se utiliza 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 indicados, pero también puede incluir características o componentes adicionales.

Los encabezados de sección utilizados en la presente memoria tienen fines organizativos solamente, y no deben considerarse como limitantes de la materia objeto descrita.

Como se emplea en esta memoria, las abreviaturas de los orgánicos comunes se definen de la siguiente manera:

Ac	Acetilo
Ac_2O	Anhídrido acético
ac.	Acuoso
Bn	Bencilo
Bz	Benzoilo
BOC o Boc	*terc*-Butoxicarbonilo
Bu	n-butilo
cat.	Catalítico
ºC	Temperatura en grados centígrados
CHAPS	3-[(3-colamidopropil)dimetilamonio]-1-propanosulfonato
dATP	Desoxiadenosín trifosfato
dCTP	Desoxicitidín trifosfato
dGTP	Desoxiguanosín trifosfato
dTTP	Desoxitimidín trifosfato
ddNTP (o ddNTPs)	Didesoxinucleótido (o didesoxinucleótidos)
DCM	Cloruro de metileno
DMA	Dimetilacetamida
DMAP	4-dimetilaminopiridina
DMF	N,N'-dimetilformamida
DMSO	Dimetilsulfóxido
DSC	Carbonato de N,N'-disuccinimidilo
EDTA	Ácido etilendiaminotetraacético
Et	Etilo
EtOAc	Acetato de etilo
ffN	Nucleótido completamente funcional
ffA	Nucleótido adenosina completamente funcionalizado
g	Gramo (o gramos)
GPC	Cromatografía de permeación en gel
h o hr	Hora (u horas)
Base de Hunig	N,N-diisopropiletilamina
iPr	Isopropilo
KPi	Tampón de fosfato de potasio 10 mM a pH 7,0
IPA	Alcohol isopropílico

LCMS	Cromatografía líquida-espectrometría de masas
LDA	Diisopropilamida de litio
m o min	Minuto (o minutos)
MeCN	Acetonitrilo
ml	Mililitro (o mililitros)
PEG	Polietilenglicol
GP	Grupo protector
Ph	Fenilo
pNB	p-nitro-bencilo
ppt	Precipitado
rt	Temperatura ambiente
SPS	Secuenciación por síntesis
$-S(O)_2OH$	Hidróxido de sulfonilo
TEA	Trietilamina
TEAB	Bromuro de tetraetilamonio
TFA	Ácido trifluoracético
Terc, t	terciario
THF	Tetrahidrofurano
TLC	Cromatografía de capa fina
TSTU	Tetrafluoroborato de O-(N-succinimidil)-N,N,N',N'-tetrametiluranio
µl	Microlitro (o microlitros)

Como se emplea en esta memoria, el término "matriz" se refiere a una población de distintas moléculas de sonda que están unidas a uno o más sustratos de modo que las distintas moléculas de sonda pueden diferenciarse entre sí según la ubicación relativa. Una matriz puede incluir distintas moléculas de sonda que estén ubicadas en una ubicación direccionable diferente sobre un sustrato. De manera alternativa o adicional, una matriz puede incluir sustratos separados, portando cada uno una molécula de sonda distinta, en donde las distintas moléculas de sonda pueden identificarse según las ubicaciones de los sustratos sobre una superficie a la que están unidos los sustratos o según las ubicaciones de los sustratos en un líquido. Las matrices ilustrativas en las que se ubican sustratos distintos sobre una superficie incluyen, sin limitación, los que incluyen perlas en pocillos como se describe, por ejemplo, en la Patente de Estados Unidos N.º 6.355.431 B1, el documento US 2002/0102578 y Publicación PCT N.º WO 00/63437. Los formatos ilustrativos que pueden utilizarse en la invención para distinguir perlas en una matriz líquida, por ejemplo, utilizando un dispositivo microfluídico, tal como un separador de células activadas por fluorescencia (FACS), se describen, por ejemplo, en la patente de Estados Unidos N.° 6.524.793. Otros ejemplos de matrices que pueden utilizarse en la invención incluyen, sin limitación, los descritos en las Patentes de Estados Unidos N.º 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; 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 documentos WO 93/17126; documentos WO 95/11995; documentos WO 95/35505; patentes europeas EP 742 287; y EP 799 897.

Como se emplea en esta memoria, la expresión "unido covalentemente" o "enlazado covalentemente" se refiere a la formación de un enlace químico que se caracteriza por compartir parejas de electrones entre los átomos. Por ejemplo, un recubrimiento polimérico unido covalentemente se refiere a un recubrimiento polimérico que forma enlaces químicos con una superficie funcionalizada de un sustrato, en comparación con la unión a la superficie por otros medios, por ejemplo, adhesión o interacción electrostática. Se apreciará que los polímeros que están unidos covalentemente a una superficie también pueden unirse por medios además de la unión covalente.

Como se emplea en esta memoria, "C_a a C_b" o "C_{a-b}", en que "a" y "b" son números enteros, se refiere al número de átomos de carbono en el grupo especificado. Es decir, el grupo puede contener de "a" a "b", inclusive, átomos de carbón. Por lo tanto, por ejemplo, un grupo "alquilo C_1 a C_4" o "alquilo C_{1-4}" se refiere a todos los grupos alquilo que tienen de 1 a 4 carbonos, es decir, CH_3-, CH_3CH_2-, $CH_3CH_2CH_2$-, $(CH_3)_2CH$-, $CH_3CH_2CH_2CH_2$-, $CH_3CH_2CH(CH_3)$- y $(CH_3)_3C$-.

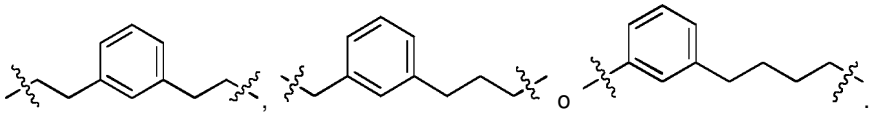
El término "halógeno" o "halo", como se emplea en esta memoria, significa uno cualquiera de los átomos radioestables de la columna 7 de la Tabla Periódica de los Elementos, p. ej., flúor, cloro, bromo o yodo, siendo preferidos flúor y cloro.

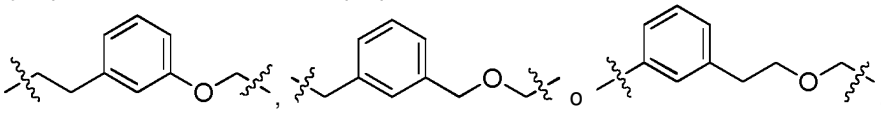
Como se emplea en esta memoria, "alquilo" se refiere a una cadena de hidrocarburo lineal o ramificada que está completamente saturada (es decir, no contiene enlaces dobles o triples). El grupo alquilo puede tener de 1 a 20 átomos

de carbono (siempre que aparezca en la presente memoria un intervalo numérico tal como "1 a 20" se refiere a cada número entero en el intervalo dado; p. ej., "1 a 20 átomos de carbono" significa que el grupo alquilo puede consistir en 1 átomo de carbono, 2 átomos de carbono, 3 átomos de carbono, etc., hasta, e incluido, 20 átomos de carbono, aunque la presente definición también cubre la aparición del término "alquilo" cuando no se designa ningún intervalo numérico). El grupo alquilo puede ser también un alquilo de tamaño medio que tiene de 1 a 9 átomos de carbono. El grupo alquilo puede ser también un alquilo inferior que tiene de 1 a 4 átomos de carbono. El grupo alquilo puede designarse como "alquilo C_{1-4}" o designaciones similares. Únicamente a modo de ejemplo, "alquilo C_{1-4}" indica que hay de uno a cuatro átomos de carbono en la cadena de alquilo, es decir, la cadena de alquilo se selecciona del grupo que consiste en metilo, etilo, propilo, iso-propilo, n-butilo, iso-butilo, sec-butilo y t-butilo. Los grupos alquilo típicos incluyen, pero no se limitan de ningún modo a, metilo, etilo, propilo, isopropilo, butilo, isobutilo, butilo terciario, pentilo, hexilo y similares. El grupo alquilo puede estar sustituido o no sustituido.

Como se emplea en esta memoria, "alcoxi" se refiere a la fórmula -OR en donde R es un alquilo como se define anteriormente, tal como "alcoxi C_{1-9}", incluyendo, pero no se limita a, metoxi, etoxi, n-propoxi, 1-metiletoxi (isopropoxi), n-butoxi, iso-butoxi, sec-butoxi y *terc*-butoxi, y similares.

Como se emplea en esta memoria, "heteroalquilo" se refiere a una cadena de hidrocarburo lineal o ramificada que contiene uno o más heteroátomos, es decir, un elemento distinto del carbono, incluyendo, pero no se limita a, nitrógeno, oxígeno y azufre, en la cadena principal. El grupo heteroalquilo puede tener de 1 a 20 átomos de carbono, aunque la presente definición también cubre la aparición del término "heteroalquilo" cuando no se designa ningún intervalo numérico. El grupo heteroalquilo puede ser también un heteroalquilo de tamaño medio que tiene de 1 a 9 átomos de carbono. El grupo heteroalquilo puede ser también un heteroalquilo inferior que tiene de 1 a 4 átomos de carbono. El grupo heteroalquilo puede designarse como "heteroalquilo C_{1-4}" o designaciones similares. El grupo heteroalquilo puede contener uno o más heteroátomos. Únicamente a modo de ejemplo, "heteroalquilo C_{1-4}" indica que hay de uno a cuatro átomos de carbono en la cadena de heteroalquilo y adicionalmente uno o más heteroátomos en la estructura principal de la cadena.

Como se emplea en esta memoria, "alquileno" significa un grupo químico di-radical completamente saturado de cadena ramificada o lineal que contiene solo carbono e hidrógeno, que está unido al resto de la molécula a través de dos puntos de unión (es decir, un alcanodiilo). El grupo alquileno puede tener de 1 a 20 átomos de carbono, aunque la presente definición también cubre la aparición del término "alquileno" cuando no se designa ningún intervalo numérico. El grupo alquileno puede ser también un alquileno de tamaño medio que tiene de 1 a 9 átomos de carbono. El grupo alquileno puede ser también un alquileno inferior que tiene de 1 a 4 átomos de carbono. El grupo alquileno puede designarse como "alquileno C_{1-4}" o designaciones similares. Únicamente a modo de ejemplo, "alquileno C_{1-4}" indica que hay de uno a cuatro átomos de carbono en la cadena de alquileno, es decir, la cadena de alquileno se selecciona del grupo que consiste en metileno, etileno, etan-1,1-diilo, propileno, propan-1,1-diilo, propan-2,2-diilo, 1-metil-etileno, butileno, butan-1,1-diilo, butan-2,2-diilo, 2-metil-propan-1,1-diilo, 1-metil-propileno, 2-metil-propileno, 1,1-dimetil-etileno, 1,2-dimetil-etileno y 1-etil-etileno. Como se emplea en esta memoria, cuando un alquileno está interrumpido por un grupo aromático, se refiere a la inserción de un grupo aromático entre un enlace carbono-carbono de la cadena de alquileno a través de dos puntos de unión, o la unión de un grupo aromático a un terminal de la cadena de alquileno a través de un punto de unión. Por ejemplo, cuando un n-butileno está interrumpido por un grupo fenilo, las estructuras ilustrativas incluyen  .

Como se emplea en esta memoria, el término "heteroalquileno" se refiere a una cadena de alquileno en la que uno o más átomos esqueléticos del alquileno se seleccionan de un átomo distinto del carbono, p. ej., oxígeno, nitrógeno, azufre, fósforo o combinaciones de los mismos. La cadena de heteroalquileno puede tener una longitud de 2 a 20.000. Las estructuras ilustrativas incluyen, pero no se limitan a, $-OCH_2-$, $-OCH(CH_3)-$, $-OC(CH_3)_2-$, $-OCH_2CH_2-$, $-CH(CH_3)O-$, $-CH_2OCH_2-$, $-CH_2OCH_2CH_2-$, $-SCH_2-$, $-SCH(CH_3)-$, $-SC(CH_3)_2-$, $-SCH_2CH_2-$, $-CH_2SCH_2CH_2-$, $-NHCH_2-$, $-NHCH(CH_3)-$, $-NHC(CH_3)_2-$, $-NHCH_2CH_2-$, $-CH_2NHCH_2-$, $-CH_2NHCH_2CH_2-$, y similares. Como se emplea en esta memoria, cuando un heteroalquileno está interrumpido por un grupo aromático, se refiere a la inserción de un grupo aromático entre un enlace carbono-carbono o carbono-heteroátomo de la cadena de heteroalquileno a través de dos puntos de unión, o la unión de un grupo aromático a un terminal de la cadena de heteroalquileno a través de un punto de unión. Por ejemplo, cuando un óxido de n-propileno está interrumpido por un grupo fenilo, las estructuras ilustrativas incluyen  .

Como se emplea en esta 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 no sustituido o sustituido.

Como se emplea en esta 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 no sustituido o sustituido.

Como se emplea en esta memoria, "cicloalquilo" se refiere a un sistema de anillos de hidrocarburo mono- o multicíclico completamente saturado (sin dobles o triples enlaces). Cuando está compuesto por dos o más anillos, los anillos pueden estar unidos uno a otro de forma condensada. Los grupos cicloalquilo pueden contener de 3 a 10 átomos en el anillo (o anillos). En algunas realizaciones, los grupos cicloalquilo pueden contener de 3 a 8 átomos en el anillo (o anillos). Un grupo cicloalquilo puede estar no sustituido o sustituido. Los grupos cicloalquilo típicos incluyen, pero no se limitan de ningún modo a, ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo, cicloheptilo y ciclooctilo.

El término "aromático" se refiere a un anillo o sistema de anillos que tiene un sistema de electrones pi conjugado e incluye grupos aromáticos carbocíclicos (p. ej., fenilo) y heterocíclicos (p. ej., piridina). El término incluye grupos monocíclicos o policíclicos de anillos condensados (es decir, anillos que comparten parejas de átomos adyacentes) con la condición de que todo el sistema de anillos sea aromático.

Como se emplea en esta memoria, "arilo" se refiere a un anillo o sistema de anillos aromático (es decir, dos o más anillos condensados que comparten dos átomos de carbono adyacentes) que solo contiene carbono en la estructura principal del anillo. Cuando el arilo es un sistema de anillos, cada anillo del sistema es aromático. El grupo arilo puede tener de 6 a 18 átomos de carbono, aunque la presente definición también cubre la aparición del término "arilo" cuando no se designa ningún intervalo numérico. En algunas realizaciones, el grupo arilo tiene de 6 a 10 átomos de carbono. El grupo arilo puede designarse como "arilo C_{6-10}", "arilo C_6 o C_{10}", o designaciones similares. Los ejemplos de grupos arilo incluyen, pero no se limitan a, fenilo, naftilo, azulenilo y antracenilo.

Un "aralquilo" o "arilalquilo" es un grupo arilo conectado, como sustituyente, a través de un grupo alquileno, tal como "aralquilo C_{7-14}" y similares, incluyendo, pero no se limita a, bencilo, 2-feniletilo, 3-fenilpropilo y naftilalquilo. En algunos casos, el grupo alquileno es un grupo alquileno inferior (es decir, un grupo alquileno C_{1-4}).

Como se emplea en esta memoria, "heteroarilo" se refiere a un anillo o sistema de anillos aromático (es decir, dos o más anillos condensados que comparten dos átomos de adyacentes) que contiene (o contienen) uno o más heteroátomos, es decir, un elemento distinto del carbono, incluyendo, pero no se limita a, nitrógeno, oxígeno y azufre, en a estructura principal del anillo. Cuando el heteroarilo es un sistema de anillos, cada anillo del sistema es aromático. El grupo heteroarilo puede tener 5-18 miembros en el anillo (es decir, el número de átomos que constituyen a estructura principal del anillo, incluidos los átomos de carbono y heteroátomos), aunque la presente definición también cubre la aparición del término "heteroarilo" cuando no se designa ningún intervalo numérico. En algunas realizaciones, el grupo heteroarilo tiene 5 a 10 miembros en el anillo o 5 a 7 miembros en el anillo. El grupo heteroalquilo puede designarse como "heteroalquilo de 5-7 miembros", "heteroarilo de 5-10 miembros", o designaciones similares. Los ejemplos de anillos heteroarilo incluyen, pero no se limitan a, furilo, tienilo, ftalazinilo, pirrolilo, oxazolilo, tiazolilo, imidazolilo, pirazolilo, isoxazolilo, isotiazolilo, triazolilo, tiadiazolilo, piridinilo, piridazinilo, pirimidinilo, pirazinilo, triazinilo, quinolinilo, isoquinlinilo, benzoimidazolilo, benzoxazolilo, benzotiazolilo, indolilo, isoindolilo y benzotienilo.

Un "heteroaralquilo" o "heteroarilalquilo" es un grupo heteroarilo conectado, como sustituyente, a través de un grupo alquileno. Los ejemplos incluyen, pero no se limitan a, 2-tienilmetilo, 3-tienilmetilo, furilmetilo, tieniletilo, pirrolilalquilo, piridilalquilo, isoxazolilalquilo e imidazolilalquilo. En algunos casos, el grupo alquileno es un grupo alquileno inferior (es decir, un grupo alquileno C_{1-4}).

Como se emplea en esta memoria, "cicloalquilo" significa un anillo o sistema de anillos carbociclilo completamente saturado. Los ejemplos incluyen ciclopropilo, ciclobutilo, ciclopentilo y ciclohexilo.

Un grupo "O-carboxi" se refiere a un grupo "-OC(=O)R" en el que R se selecciona de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria.

Un grupo "C-carboxi" se refiere a un grupo "-C(=O)OR" en el que R se selecciona de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria. Un ejemplo no limitante incluye carboxilo (es decir, -C(=O)OH).

Un grupo "ciano" se refiere a un grupo "-CN".

Un grupo "azido" se refiere a un grupo "-N3".

Un grupo "O-carbamilo" se refiere a un grupo "-OC(=O)NR_AR_B" en el que R_A y R_B se seleccionan cada uno de forma independiente de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria.

Un grupo "N-carbamilo" se refiere a un grupo "-$N(R_A)$OC(=O)R_B" en el que R_A y R_B se seleccionan cada uno de forma independiente de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria.

Un grupo "C-amido" se refiere a un grupo "-C(=O)NR_AR_B" en el que R_A y R_B se seleccionan cada uno de forma independiente de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria.

Un grupo "N-amido" se refiere a un grupo "-$N(R_A)C(=O)R_B$" en el que R_A y R_B se seleccionan cada uno de forma independiente de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria.

Un grupo "amino" se refiere a un grupo "-NR_AR_B" en el que R_A y R_B se seleccionan cada uno de forma independiente de hidrógeno, alquilo C_{1-6}, alquenilo C_{2-6}, alquinilo C_{2-6}, carbociclilo C_{3-7}, arilo C_{6-10}, heterarilo de 5-10 miembros y heterociclilo de 5-10 miembros, como se define en la presente memoria. Un ejemplo no limitante incluye amino libre (es decir, -NH_2).

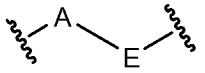
Como se emplea en esta memoria, el término "Trolox" se refiere al ácido 6-hidroxi-2,5,7,8-tetrametilcroman-2-carboxílico.

Como se emplea en esta memoria, el término "ascorbato" se refiere a la sal de ácido ascórbico.

Como se emplea en esta memoria, el término "ácido gálico" se refiere al ácido 3,4,5-trihidroxibenzoico.

Como se emplea en esta memoria, un grupo sustituido se obtiene a partir del grupo precursor no sustituido en el que ha habido un intercambio de uno o más átomos de hidrógeno por otro átomo o grupo. A menos que se indique otra cosa, cuando se considera que un grupo está "sustituido", se entiende que el grupo está sustituido con uno o más sustituyentes seleccionados de forma independiente de alquilo C_1-C_6, alquenilo C_1-C_6, alquinilo C_1-C_6, heteroalquilo C_1-C_6, carbociclilo C_3-C_7 (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), carbociclil C_3-C_7-alquilo C_1-C_6 (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), heterociclilo de 5-10 miembros (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), heterociclil-alquilo C_1-C_6 de 5-10 miembros (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), arilo (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), arilalquilo C_1-C_6 (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), heteroarilo de 5-10 miembros (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), heteroarilalquilo C_1-C_6 de 5-10 miembros (opcionalmente sustituido con halo, alquilo C_1-C_6, alcoxi C_1-C_6, haloalquilo C_1-C_6 y haloalcoxi C_1-C_6), halo, ciano, hidroxi, alcoxi C_1-C_6, alcoxi C_1-C_6-alquilo C_1-C_6 (es decir, éter), ariloxi, sulfhidrilo (mercapto), haloalquilo C_1-C_6 (p. ej. -CF_3), haloalcoxi C_1-C_6 (p. ej. -OCF_3), alquiltio C_1-C_6, ariltio, amino, aminoalquilo C_1-C_6, nitro, O-carbamilo, N-carbamilo, O-tiocarbamilo, N-tiocarbamilo, C-amido, N-amido, S-sulfonamido, N-sulfonamido, C-carboxi, O-carboxi, acilo, cianato, isocianato, tiocianato, isotiocianato, sulfinilo, sulfonilo y oxo (=O). Cuando un grupo se describe como "opcionalmente sustituido", ese grupo puede estar sustituido con los sustituyentes anteriores.

Debe apreciarse que determinadas convenciones de denominación de radicales pueden incluir un mono-radical o un di-radical, dependiendo del contexto. Por ejemplo, cuando un sustituyente precisa dos puntos de unión al resto de la molécula, se entiende que el sustituyente es un di-radical. Por ejemplo, un sustituyente identificado como alquilo que precisa dos puntos de unión incluye di-radicales tales como -CH_2-, -CH_2CH_2-, -$CH_2CH(CH_3)CH_2$-, y similares. De manera similar, un grupo identificado como amino que precisa dos puntos de unión incluye di-radicales tales como -NH-, -$N(CH_3)$- y similares. Otras convenciones de denominación de radicales indican claramente que el radical es un di-radical tal como "alquileno" o "alquenileno".

Cuando un sustituyente se representa como un di-radical (es decir, tiene dos puntos de unión al resto de la molécula), debe entenderse que el sustituyente puede unirse en cualquier configuración direccional a menos que se indique otra cosa. Por lo tanto, por ejemplo, un sustituyente representado como -AE- o 

incluye el sustituyente orientado de tal manera que el **A** está unido en el punto de unión más a la izquierda de la molécula, así como el caso en que **A** está unido en el punto de unión más a la derecha de la molécula.

Cuando los compuestos descritos en esta memoria tienen al menos un estereocentro, pueden existir como enantiómeros y diastereómeros individuales o como mezclas de tales isómeros, incluyendo racematos. La separación de los isómeros individuales o la síntesis selectiva de los isómeros individuales se lleva a cabo mediante la aplicación de diversos métodos que son bien conocidos por los expertos en la técnica. A menos que se indique otra cosa, todos tales isómeros y mezclas de los mismos están incluidos en el alcance de los compuestos descritos en la presente memoria.

Como se emplea en esta 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 es 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 purínica o pirimidínica. Las bases purínicas incluyen adenina (A) y guanina (G), y derivados modificados o análogos de las mismas. Las bases pirimidínicas incluyen citosina (C), timina (T) y uracilo (U), y derivados modificados o análogos de las mismas. El átomo C-1 de la desoxirribosa está enlazado al N-1 de una pirimidina o al N-9 de una purina.

Como se emplea en esta memoria, un "nucleósido" es estructuralmente similar a un nucleótido, pero le faltan las fracciones fosfato. Un ejemplo de un análogo de nucleósido sería uno en el que el marcador está unido a la base y no

hay un grupo fosfato unido a la molécula de azúcar. El término "nucleósido" se usa en la presente memoria en su sentido ordinario tal como lo entienden los expertos en la técnica. Los ejemplos incluyen, pero no se limitan a, un ribonucleósido que comprende una fracción ribosa y un desoxirribonucleósido que comprende una fracción desoxirribosa. Una fracción pentosa modificada es una fracción pentosa en la 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 y/o un resto de azúcar sustituidos. Adicionalmente, se puede incorporar un nucleósido en polímeros y oligómeros de ADN y/o ARN más grandes.

Como se emplea en esta memoria, el término "polinucleótido" se refiere a los ácidos nucleicos en general, incluyendo el ADN (p. ej. ADNc de ADN genómico), ARN (p. ej., ARNm), oligonucleótidos sintéticos y análogos de ácidos nucleicos sintéticos. Los polinucleótidos pueden incluir bases naturales o no naturales, o combinaciones de las mismas, y enlaces de la estructura principal naturales o no naturales, p. ej., fosforotioatos, PNA o 2'-O-metil-ARN, o combinaciones de los mismos.

Como se emplea en esta memoria, la expresión "escalonamiento (*phasing*)" se refiere a los fenómenos en la SPS provocados por la eliminación incompleta de los terminadores 3' y los fluoróforos, y la imposibilidad de las polimerasas de completar la incorporación de una porción de cadenas de ADN dentro de los grupos (*clusters*) en un ciclo de secuenciación dado. El pre-escalonamiento (*pre-phasing*) es producido por la incorporación de nucleótidos sin terminadores 3' eficaces y el acontecimiento de incorporación continua 1 ciclo. El escalonamiento y el pre-escalonamiento hacen que las intensidades extraídas para un ciclo específico consistan en la señal del ciclo actual, así como en el ruido de los ciclos anteriores y siguientes. A medida que aumenta el número de ciclos, aumenta la fracción de secuencias por grupo (*cluster*) afectada por el escalonamiento, obstaculizando la identificación de la base correcta. El pre-escalonamiento puede estar provocado por la presencia de una cantidad mínima de nucleótidos 3'-OH no protegidos o no bloqueados durante la secuenciación por síntesis (SPS). Los nucleótidos 3'-OH no protegidos podrían generarse durante los procedimientos de fabricación o posiblemente durante los procedimientos de almacenamiento y manipulación de reactivos. Por consiguiente, las modificaciones de los análogos de nucleótidos o los grupos de enlace que dan como resultado un tiempo de ciclo de SPS más rápido, valores más bajos de escalonamiento y pre-escalonamiento, y una longitud de lectura de secuenciación más larga, proporcionan mayores ventajas en aplicaciones de SPS.

Como se emplea en esta memoria, la expresión "fracción protectora" incluye, pero no se limita a, moléculas que pueden proteger frente a daños en el ADN (p. ej., foto-daño u otros daños químicos). Algunos ejemplos específicos incluyen antioxidantes, tales como la vitamina C, derivados de vitamina E, ácido fenólico, polifenoles, y derivados y análogos de los mismos. Debe entenderse que en determinados contextos donde se define la expresión "fracción protectora", se refiere a la fracción resultante de la reacción entre uno o más grupos funcionales de la fracción protectora con el grupo funcional correspondiente del enlazador, como se describe en la presente memoria. Por ejemplo, cuando la fracción protectora es "ácido gálico" puede referirse a las amidas y los ésteres del ácido gálico en lugar de al propio ácido gálico con el grupo carboxilo libre.

Marcadores detectables

Algunas realizaciones descritas en la presente memoria se refieren al uso de marcadores detectables convencionales. La detección puede llevarse a cabo por cualquier método adecuado, incluyendo espectroscopía de fluorescencia o por otros medios ópticos. El marcador es un fluoróforo, que, después de la absorción de energía, emite radiación a una longitud de onda definida. Se conocen muchos marcadores fluorescentes adecuados. Por ejemplo, Welch *et al.* (*Chem. Eur. J.* 5(3):951-960, 1999) describen fracciones fluorescentes funcionalizadas con dansilo que pueden usarse en la presente invención. Zhu *et al.* (*Cytometry* 28:206-211, 1997) describen el uso de los marcadores fluorescentes Cy3 y Cy5, los cuales también se puede usar en la presente invención. Las etiquetas adecuadas para su uso también se describen 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). Otros marcadores fluorescentes disponibles en el mercado incluyen, pero no se limitan a, fluoresceína, rodamina (incluyendo TMR, texas red y Rox), alexa, bodipy, acridina, cumarina, pireno, benzantraceno y las cianinas.

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

Además se pueden utilizar en la presente solicitud marcadores multicomponente. Un marcador multicomponente es uno que depende de la interacción con un compuesto adicional para la detección. El marcador multicomponente más común utilizado en biología es el sistema biotina-estreptavidina. La biotina se utiliza como el marcador unido a la base nucleotídica. Después, se añade estreptavidina por separado para permitir que se produzca la detección. Están disponibles otros sistemas multicomponente. Por ejemplo, el dinitrofenol tiene un anticuerpo fluorescente disponible en el mercado que puede utilizarse para la detección.

A menos que se indique otra cosa, la referencia a los nucleótidos también pretende ser aplicable a los nucleósidos. La presente solicitud también se describirá adicionalmente con referencia al ADN, aunque la descripción también será aplicable al ARN, el PNA y otros ácidos nucleicos, a menos que se indique otra cosa.

Métodos de secuenciación

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

Los análogos de nucleótidos presentados en la presente memoria pueden utilizarse en un procedimiento de secuenciación, tal como una técnica de secuenciación por síntesis (SPS). Brevemente, la SPS puede iniciarse poniendo en contacto los ácidos nucleicos diana con uno o más nucleótidos marcados, ADN polimerasa, etc. Las casillas en que se extiende un cebador utilizando el ácido nucleico diana como molde incorporarán un nucleótido marcado que puede detectarse. Opcionalmente, los nucleótidos marcados pueden incluir además una propiedad de terminación reversible que termina la extensión adicional del cebador una vez que se ha añadido un nucleótido a un cebador. Por ejemplo, se puede añadir a un cebador un análogo de nucleótido que tiene una fracción terminadora reversible, de modo que la extensión posterior no pueda producirse hasta que se suministre un agente de desbloqueo para eliminar la fracción. Por lo tanto, para realizaciones que utilizan terminación reversible, se puede suministrar un reactivo de desbloqueo a la cubeta de lectura (antes o después de que se produzca la detección). Los lavados se pueden llevar a cabo entre las diversas etapas de suministro. El ciclo entonces puede repetirse n veces para extender el cebador en n nucleótidos, detectando de este modo una secuencia de longitud n. Los procedimientos ilustrativos de SPS, los sistemas fluídicos y las plataformas de detección que pueden adaptarse fácilmente para su uso con una matriz producida por los métodos de la presente descripción se describen, por ejemplo, en Bentley *et al*., *Nature* 456:53-59 (2008), documentos WO 04/018497; WO 91/06678; WO 07/123744; las patentes de Estados Unidos N.° 7.057.026; 7.329.492; 7.211.414; 7.315.019 o 7.405.281, y la solicitud de patente de EE.UU. Pub. N.º 2008/0108082 A1,.

Se pueden usar otros procedimientos de secuenciación que utilizan reacciones cíclicas, tales como la pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) a medida que se incorporan nucleótidos particulares en 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); patentes de Estados Unidos N.° 6.210.891; 6.258.568 y 6.274.320). En la pirosecuenciación, el PPi liberado puede detectarse al convertirse en adenosín trifosfato (ATP) por la ATP sulfurilasa, y el ATP resultante puede detectarse a través de fotones producidos por la luciferasa. Por lo tanto, la reacción de secuenciación puede controlarse a través de un sistema de detección de luminiscencia. Las fuentes de radiación de excitación utilizadas para los sistemas de detección basados en fluorescencia no son necesarias para los procedimientos de pirosecuenciación. Los sistemas fluídicos, los detectores y los procedimientos útiles que se pueden utilizar para la aplicación de pirosecuenciación a matrices de la presente descripción se describen, por ejemplo, en la Sol. de Patente WIPO N.º de Ser. PCT/US11/57111, Sol. de Pat. Pub. N.º 2005/0191698 A1, Pat. de EE.UU. N.º 7.595.883 y Pat. de EE.UU. N.º 7.244.559.

Las reacciones de secuenciación por ligamiento también son útiles, incluyendo, por ejemplo, las descritas en Shendure *et al*. *Science* 309:1728-1732 (2005); patente de Estados Unidos N.° 5.599.675; y patente de Estados Unidos N.° 5.750.341. Algunas realizaciones pueden incluir procedimientos de secuenciación por hibridación como se describe, 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 documento WO 1989/10977. Tanto en los procedimientos de secuenciación por ligamiento como en los de secuenciación por hibridación, los ácidos nucleicos que están presentes en los pocillos que contienen gel (u otras casillas cóncavas) se someten a ciclos repetidos de suministro de oligonucleótidos y detección. Los sistemas fluídicos para los métodos de SPS como se exponen en la presente memoria, o en las referencias citadas en la presente memoria, puede adaptarse fácilmente para el suministro de reactivos para los procedimientos de secuenciación por ligamiento o de secuenciación por hibridación. Normalmente, los oligonucleótidos están marcados con fluorescencia y pueden detectarse utilizando detectores de fluorescencia similares a los descritos con respecto a los procedimientos de SPS en la presente memoria o en las referencias citadas en la presente memoria.

Algunas realizaciones pueden utilizar métodos que implican el control en tiempo real de la actividad de la ADN polimerasa. Por ejemplo, las incorporaciones de nucleótidos se pueden detectar mediante interacciones de transferencia de energía de resonancia de fluorescencia (FRET, forma siglada del inglés *fluorescence resonance energy transfer*) entre una polimerasa que porta un fluoróforo y nucleótidos marcados con γ-fosfato, o con guías de onda de modo cero. Se describen técnicas y reactivos para la 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 realizaciones de SPS 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 utilizar un detector eléctrico y técnicas asociadas que están disponibles en el mercado en Ion Torrent (Guilford, CT, una filial de Life Technologies) o métodos y sistemas de secuenciación descritos en la Sol. de patente de Estados Unidos Pub.

N.os 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; o 2010/0282617 A1.

Enlazadores modificados ilustrativos

Se describen realizaciones adicionales con más detalle en los siguientes ejemplos.

La **FIG. 1A** ilustra una fórmula estructural parcial de un nucleótido marcado **100.** El nucleótido marcado **100** incluye un nucleótido adenosina completamente funcionalizado (ffA) **110,** una fracción enlazadora convencional **115,** y un colorante fluorescente **120.** La fracción enlazadora convencional **115** puede ser una fracción enlazadora usada normalmente en la síntesis de nucleótidos marcados para la secuenciación por síntesis (SPS). En un ejemplo, el colorante fluorescente **120** es NR550S0. En este ejemplo, el nucleótido marcado **100** puede describirse como "ffA-NR550S0". En otro ejemplo, el colorante fluorescente **120** es SO7181 y el nucleótido marcado **100** puede describirse como "ffA-SO7181".

La **FIG.1B** ilustra el nucleótido marcado **100** de la **FIG. 1A** con dos posibles modificaciones estructurales en la fracción enlazadora convencional **115.** En un ejemplo, la fracción enlazadora convencional **115** incluye un inserto AEDI **125** entre la porción carbonilo (es decir, -C(=O)-) y la porción amino (es decir, -NH-) de la fracción amido. En este ejemplo, el nucleótido marcado modificado puede describirse como "ffA-AEDI-NR550S0". En otro ejemplo, la fracción enlazadora convencional **115** incluye un inserto SS **130** y el nucleótido marcado modificado puede denominarse "ffA-SS-NR550S0".

La **FIG. 2** muestra una representación de la velocidad de incorporación de nucleótidos utilizando el nucleótido marcado **100** de la **FIG. 1A** y los nucleótidos marcados **100** modificados de la **FIG. 1B.** El ensayo se llevó a cabo a 55 ºC, en etanolamina 40 mM (pH 9,8), MgCl 9 mM, NaCl 40 mM, EDTA 1 mM, CHAPS al 0,2 % con cebador:molde de ADN 20 mM y polimerasa 812 30 ug/ml (MiSeq Kit V2), nucleótido 1 mM. La enzima se une al ADN y después se mezcla rápidamente con el nucleótido en una máquina de flujo de desactivación por un corto tiempo (hasta 10 s) antes de desactivar con EDTA 500 mM. Se toman varios puntos de tiempo para cada nucleótido. Las muestras generadas se analizan después en un gel desnaturalizante y se determina el porcentaje de ADN que se convierte en ADN+1, y se representa frente al tiempo para determinar la constante de velocidad de primer orden para cada nucleótido. Los datos se resumen a continuación en la Tabla 1. Los datos muestran que la velocidad de incorporación de un nucleótido marcado que comprende el inserto SS **130** (ffa-SS-NR550S0) fue aproximadamente 2 veces más rápida en comparación con la velocidad de incorporación de un nucleótido marcado que comprende un enlazador convencional **115** (ffa-NR550S0). La velocidad de incorporación de un nucleótido marcado que comprende un inserto AEDI **125** (ffa-AEDI-NR550S0) fue aproximadamente 4 veces más rápida en comparación con la velocidad de incorporación de ffa-NR550S0. Los datos también muestran que la velocidad de incorporación de un nucleótido marcado que comprende un enlazador convencional **115** y el colorante fluorescente SO7181 (ffa-SO7181) fue aproximadamente 4 veces más rápida en comparación con la velocidad de incorporación de ffa-NR550S0.

Tabla 1.	
ffA	K (μM/min)
ffA-NR550S0	22 (1X)
ffA-SS-NR550S0	54 (2X)
ffA-AEDI-NR550S0	97 (4X)
ffA-SO7181	99 (4X)

Las **FIG. 3A-3F** ilustran las fórmulas estructurales de los insertos adicionales no según la invención **310, 315, 320, 325, 330** y **335** para la fracción enlazadora convencional **115** de la **FIG. 1A.** En un inserto ACA **310,** las sustituciones de dimetilo se eliminan y el enlace azufre-azufre (SS) se reemplaza por un enlace carbono-carbono en comparación con el inserto **125.** El enlace azufre-azufre no es necesario para la SPS (p. ej., SPS de 2 colorantes o 4 colorantes).

En un inserto AcLys **315,** se utiliza lisina protegida con acetilo para reemplazar el inserto **125.**

En un inserto BocLys **320,** se utiliza lisina protegida con *terc*-butoxicarbonilo para reemplazar el inserto **125.**

En un inserto dMeO **325,** el enlace azufre-azufre (SS) se reemplaza por un enlace oxígeno-carbono (O-CH_2) en comparación con el inserto **125.**

En un inserto dMeS **330,** el enlace azufre-azufre (SS) se reemplaza por un enlace azufre-carbono (S-CH_2) en comparación con el inserto **125.**

En un inserto DMP **335,** el enlace azufre-azufre (SS) se reemplaza por un enlace azufre-carbono (CH_2) en comparación con el inserto **125.**

En diversos ejemplos descritos en la presente memoria, los insertos que incluyen un patrón de sustitución de dimetilo (p. ej., un inserto AEDI **125,** un inserto dMeO **325,** y un inserto dMeS **330)** se encontró que tenían una velocidad

aumentada de incorporación de nucleótidos durante la SPS.

En diversos ejemplos descritos en la presente memoria, la longitud de las cadenas de carbono en los insertos también puede variar.

La **FIG. 4** muestra una tabla de datos de una ejecución de secuenciación de dos colorantes utilizada para evaluar el efecto del inserto AEDI **125** de la **FIG.1B** y del inserto AcLys **315** de la **FIG. 3B** sobre la calidad de la secuenciación. La secuenciación se ejecutó en una plataforma híbrida Miseq con un molde humano de 550 pb y 2 veces 150 ciclos. El nuevo conjunto de colorantes, V10/cian-peg4 A-AEDI550S0, V10/cian-peg4 A-AcLys550S0 y V10/cian A-AcLys 550S0 se compararon con el conjunto de colorantes comerciales convencionales V4 y un conjunto de colorantes mejorado de la plataforma Nova V5.75. Para cada una de las muestras V10/cian-peg4 A-AEDI550S0, V10/cian-peg4 A-AcLys550S0 y VIO/cian A-AcLys 550S0, el valor de escalonamiento (Ph R1) fue menor que los valores de escalonamiento de las muestras sin los insertos AEDI o AcLys. Por lo tanto, los nucleótidos marcados que comprenden los insertos **125** y **315** adicionales demostraron mejoras en la calidad de secuenciación.

Las **FIG. 5A y 5B** muestran respectivamente una representación de la tasa de errores para la lectura 1 y una representación de la tasa de errores para la lectura 2 de la ejecución de secuenciación de la **FIG. 4**. Para la lectura 1, las tasas de error de V10/cian-peg4 A-AEDI550S0 y V10/cian-peg4 A-AcLys550S0 fueron más bajas que el mismo conjunto de colorante V10/Cyan-peg4 sin el inserto. Para la lectura 2, es aún más pronunciado cuando se utiliza el inserto AcLys **315**, donde la tasa de error final se redujo en un 30 % en comparación con el conjunto de colorante sin inserto. Por lo tanto, las insertos **125** y **315** han demostrado mejorar significativamente la calidad de secuenciación.

La **FIG. 6** muestra una tabla de datos de una ejecución de secuenciación utilizada para evaluar el efecto del inserto AEDI **125** y del inserto ACA **310** sobre la calidad de la secuenciación. La secuenciación se ejecutó en una plataforma híbrida Miseq con un molde humano de 550 pb y 2 veces 150 ciclos. El nuevo conjunto de colorantes, V10/cian-peg4 A-AEDI550S0, V10/cian-peg4 A-ACALys550S0, se compararon con el conjunto de colorantes comerciales convencionales V4 y un conjunto de colorantes mejorado de la plataforma Nova V5.75. Nuevamente, cada una de las muestras V10/cian-peg4 A-AEDI550S0 y V10/cian-peg4 A-ACA550S0 tiene un valor de escalonamiento (Ph R1) más bajo en comparación con las muestras sin los insertos AEDI o ACA y, por lo tanto, mostraron una mejora en la calidad de la secuenciación.

Las **FIG. 7A y 7B** muestran respectivamente una representación de la tasa de errores para la lectura 1 y una representación de la tasa de errores para la lectura 2 de la ejecución de secuenciación de la **FIG. 6**. Las tasas de error para la lectura 1 para el conjunto que comprende el nuevo inserto AEDI (V10/cian-peg4 A-AEDI550S0) fueron más bajas que el mismo conjunto de colorante sin el inserto, V10/Cian-peg4. El inserto ACA (V10/Cian-peg4 ACA550S0) arrojó una representación de la tasa de error similar en ambas lecturas al V10/Cian-peg4 convencional. AEDI demostró, nuevamente, mejorar la calidad de la secuenciación. Los datos mostraron también que la estructura del inserto en sí está influyendo en la mejora de la calidad de la secuenciación.

La **FIG. 8A** muestra la fórmula estructural de un enlazador LN_3 **800** convencional. El enlazador LN_3 **800** incluye una primera fracción funcional amido sustituida **810**, una segunda fracción funcional PEG sustituida con azido **815,** y una tercera fracción funcional éster **820** que puede desearse en una estructura de enlazador para unir una molécula de colorante **825** a un nucleótido **830.** La primera fracción funcional **810** puede, por ejemplo, utilizarse para unir la molécula de colorante **825** al enlazador LN_3 **800.** La segunda fracción funcional **815** puede, por ejemplo, ser un grupo funcional escindible que pueda usarse para escindir la molécula de colorante **825** del enlazador LN_3 **800.** La tercera fracción funcional **820** puede, por ejemplo, utilizarse para unir el nucleótido **830** al enlazador LN_3 **800.** La **FIG. 8B** ilustra alguna modificación al enlazador LN_3 convencional, donde la fracción fenoxi **850** está sustituida por uno a cuatro sustituyentes seleccionados de $-NO_2$, -CN, halo o $-SO_3H$. Además, la fracción éster **820** se reemplaza por una fracción amido **855.**

La Figura 9A es un gráfico que muestra la aparición de una impureza en un ffA con enlazador SS. Con ffA-SS-NR550S0, después de la purificación por HPLC: la impureza apareció durante la noche en una condición ligeramente básica (pH 8-9). T.a. durante una noche en TEAB 0,1 M/CH_3CN. La Figura 9B compara la estabilidad de los ffA con el enlazador SS y el enlazador AEDI, mostrando el enlazador AEDI una estabilidad significativamente mejorada en comparación con los enlazadores SS. Se usó ffA-LN3-NR550S0 como una referencia interna. El producto secundario disulfuro: (NR550S0-S-)2. La Figura 9C muestra una comparación de los ffA con enlazador SS y enlazador AEDI en IMX 60° durante 22 horas, nuevamente mostrando una impureza con el enlazador SS. El control interno es ffA-LN3-NR550S0. Di-P: difosfato.

Figuras 10A, 10B y 10C muestran el aumento inesperado de la velocidad de incorporación de nucleótidos en solución con los cambios de enlazador. La figura 10A muestra la velocidad de incorporación a 1 uM. Los resultados muestran que el colorante y el enlazador tuvieron un efecto significativo sobre la cinética de incorporación, mostrando claramente la tabla 10B el beneficio del enlazador AEDI en la cinética de incorporación. La Figura 10C muestra de forma esquemática los enlazadores AEDI y SS con NR550S0.

La Figura 11A muestra diagramas de dispersión para combinaciones de V10 con distintos A-550S0 (misma concentración). Los diagramas de dispersión son para el mosaico 1 ciclo 2. La Figura 11B muestra la Kcat de

enlazadores FFA en solución. Se puede observar que en la superficie, en términos de velocidad de incorporación, el enlazador AEDI es más rápido que ningún enlazador: La nube 'A' (en el diagrama de dispersión) se mueve ligeramente hacia el centro. El enlazador AcLys es más lento que AEDI y que sin enlazador: La nube 'A' se mueve hacia el eje x. El enlazador BocLys es similar a sin enlazador y no muy lejos del enlazador AEDI. En solución se puede observar que el enlazador ACA fue más lento, mientras que AEDI y ACLys tienen una Kcat similar, seguido por BocLys.

Las Figuras 12A y 12B muestran los parámetros de secuenciación sobre M111, Human550, 2x151 ciclos. Uso de los ffN en combinación con distintos A-550S0 (misma concentración). Se puede observar que tanto el enlazador AEDI como el BocLys arrojaron buenos resultados de secuencia similares. Aunque la Kcat en solución de AEDI y la de AcLys son similares, el resultado de la secuenciación de AEDI es ligeramente mejor.

El enlazador LN_3 **800** incluye una fracción fenoxi **835** opcional que puede eliminarse del enlazador LN_3 **800**, como se muestra en la **FIG. 8C.** El enlazador LN_3 **800** también incluye una fracción amido **840** opcional y una fracción éter **845** opcional, ambas pueden eliminarse del enlazador LN_3 **800** como se muestra la **FIG. 8D.** El fin de eliminar determinadas fracciones funcionales como la fracción amido **810,** la fracción fenoxi **835**, es analizar si tienen interacciones negativas con la enzima durante la incorporación de nucleótidos, lo que podría reducir la eficacia de la incorporación.

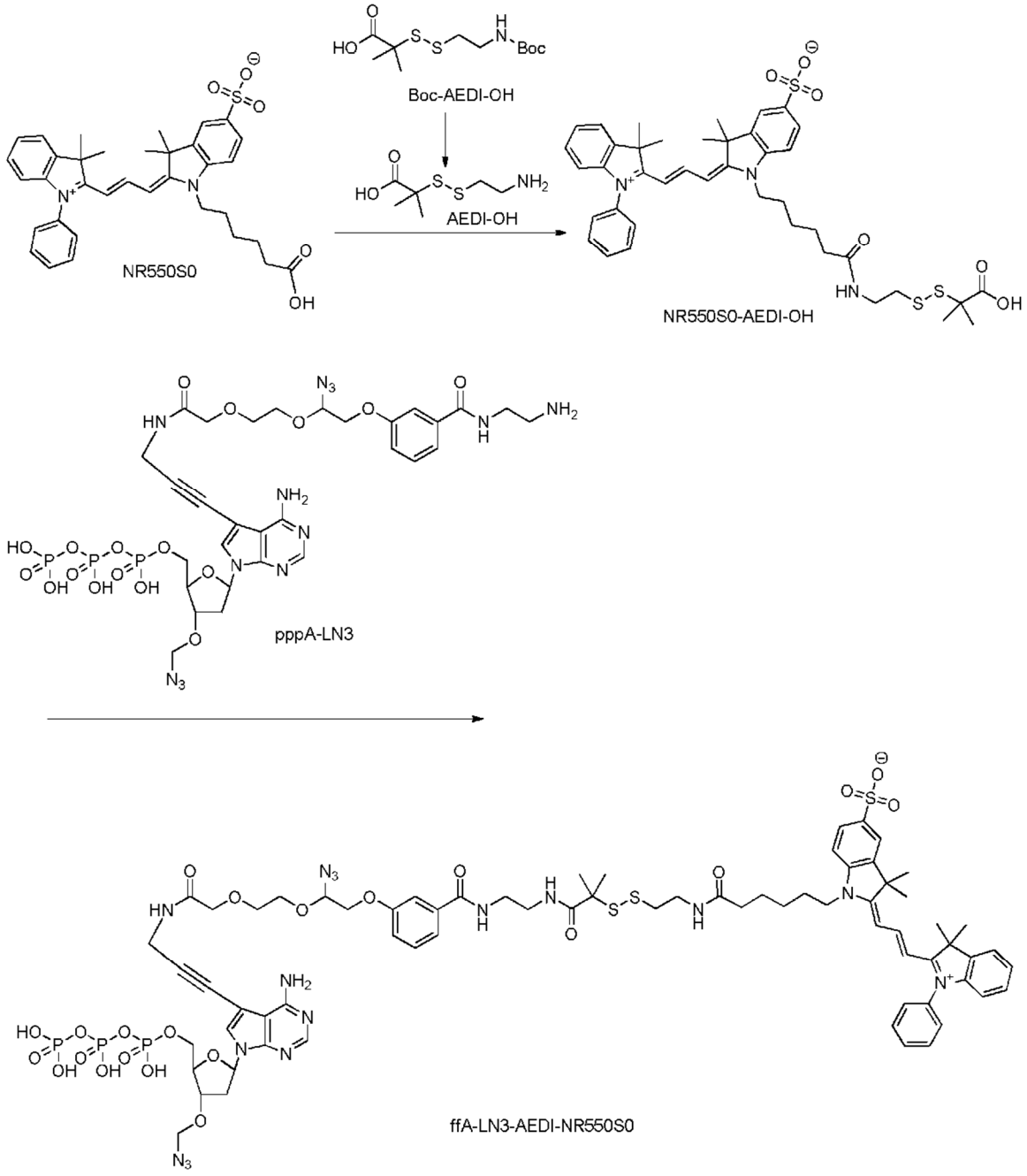
La **FIG. 8E** ilustra la inserción o adición de una fracción protectora **860** en el enlazador. La fracción protectora **860** se inserta entre la fracción funcional **810** (o se puede unir a la fracción fenoxi **835** u **850** en las **FIG. 8A-8C)** y la molécula de colorante **825.** La fracción protectora **860** puede, por ejemplo, ser una molécula que proteja frente el daño del ADN. El daño en el ADN, incluido el fotodaño u otros daños químicos, es uno de los efectos acumulativos (es decir, ciclo por ciclo) de la SPS. Reducir o eliminar sustancialmente el daño del ADN puede proporcionar una SPS más eficaz y lecturas de secuenciación más largas. En algunas realizaciones, la fracción protectora **860** se puede seleccionar de un desactivador de estado de triplete tal como Trolox, ácido gálico, 2-mercaptoetanol (BME), etc. En algunas otras realizaciones, el resto protector **860** puede seleccionarse de un reactivo de desactivación o de protección tal como alcohol 4-nitrobencílico o una sal de ácido ascórbico, tal como el ascorbato de sodio. En algunas otras realizaciones, se puede mezclar físicamente un agente protector en el tampón en lugar de formar un enlace covalente con el nucleósido o nucleótido marcado. Sin embargo, esta estrategia puede precisar una mayor concentración del agente protector y puede ser menos eficaz. Como alternativa, la fracción protectora unida covalentemente a los nucleósidos o nucleótidos puede proporcionar una mejor protección frente el daño del ADN. En algunas realizaciones adicionales de las **FIG. 8C, 8D** y **8E,** la fracción éster **820** también se puede reemplazar por la fracción amido **855** y sustituirse la fracción fenoxi **835** adicionalmente.

En cualquiera de los ejemplos demostrados en las **FIG. 8A-8E,** el inserto AEDI **125** y el inserto **SS 130** de la **FIG.1B** y los insertos **300** de las **FIG. 3A-3F** pueden, por ejemplo, insertarse entre la primera fracción funcional **810** y la molécula de colorante **825** de enlazador **800.**

Ejemplos

Se describen realizaciones adicionales con más detalle en los siguientes ejemplos.

Procedimiento general de reacción para preparar ffA-LN$_3$-AEDI-NR550S0:



En un matraz de fondo redondo de 50 ml, se disolvió Boc-AEDI-OH (1 g, 3,4 mmol) en DCM (15 ml) y se añadió TFA (1,3 ml, 17 ml) a la solución a ta. La mezcla de reacción se agitó durante 2 horas. La TLC (DCM:MeOH=9:1) indicó el consumo completo del Boc-AEDI-OH. La mezcla de reacción se evaporó hasta sequedad. Después, se añadió TEAB (2 M, ~15 ml) al residuo y se controló el pH hasta que fuera neutro. La mezcla se disolvió a continuación en H_2O/CH_3CN (1:1, ~15 ml) y se evaporó hasta sequedad. El procedimiento se repitió 3 veces para eliminar el exceso de sal TEAB. El residuo sólido blanco se trató con CH_3CN (20 ml) y se agitó durante 0,5 horas. Se filtró la solución y se lavó el sólido con CH_3CN, se obtuvo sal de AEDI-OH-TFA pura (530 mg, 80 %). RMN ^{1}H (400 MHz, D2O, δ (ppm)): 3,32 (*t,* J=6,5 Hz, 2H, NH_2-CH_2); 2,97 (*t,* J=6,5 Hz, 2H, S-CH_2); 1,53 *(s,* 6H, 2x CH_3). RMN ^{13}C (400 MHz, D_2O, δ (ppm)): 178,21 *(s,* CO); 127,91, 117,71 (2*s*, TFA); 51,87 *(s,* C-$(CH_3)_2$); 37,76 (*t,* CH_2-NH_2); 34,23 (*t,* S-CH_2); 23,84 (*q,* 2x CH_3). RMN ^{19}F (400 MHz, D_2O, δ (ppm)): -75,64.

En un matraz de fondo redondo de 50 ml, se disolvió el colorante NR550S0 (114 mg, 176 umol) en DMF (anhidro, 20 ml) y se evaporó hasta sequedad. El procedimiento se repitió 3 veces. Después, se pipeteó DMA anhidra (10 ml) y la base de Hunig (92 µl, 528 µmol, 3 equivalentes) en el matraz de fondo redondo. Se añadió TSTU (69 mg, 228 umol, 1,3 equivalentes) en una parte. La mezcla de reacción se mantuvo a ta. Después de 30 minutos, El análisis por TLC

($CH_3CN:H_2O$=85:15) indicó que la reacción había finalizado. Se añadió a la mezcla de reacción AEDI-OH (68 mg, 352 µmol, 2 equivalentes) en TEAB 0,1 M y se agitó a ta durante 3 h. La TLC ($CH_3CN:H_2O$=8:2) mostró un consumo completo del éster activado y apareció una mancha roja debajo del éster activado. Por otra parte, la HPLC analítica también indicó el consumo completo del éster activado y la formación del producto. La reacción se desactivó con tampón TEAB (0,1 M, 10 ml) y el disolvente volátil se eliminó mediante evaporación a presión reducida (HV) y se purificó en una columna Axia para obtener NR550S0-AEDI-OH. Rendimiento: 60 %.

En un matraz de fondo redondo de 25 ml, se disolvió NR550S0-AEDI-OH (10 umol) en DMF (anhidro, 5 ml) y se evaporó hasta sequedad. El procedimiento se repitió 3 veces. Después, se añadió DMA anhidra (5 ml) y DMAP (1,8 ml, 15 µmol, 1,5 equivalentes) en el matraz de fondo redondo. Se añadió DSC (5,2 mg, 20 µmol, 2 equivalentes) en una parte. La mezcla de reacción se mantuvo a temperatura ambiente. Después de 30 minutos, El análisis por TLC ($CH_3CN:H_2O$=8:2) indicó que la reacción había finalizado. La base de Hunig (3,5 ul, 20 umol) se pipeteó en la mezcla de reacción. Después, se añadió a la mezcla de reacción una solución de pppA-LN3 (20 µmol en 0,5 ml de H_2O, 2 equivalentes) y Et_3N (5 ul), y se agitó a ta durante una noche. La TLC ($CH_3CN:H_2O$=8:2) mostró un consumo completo del éster activado y apareció una mancha roja sobre punto de partida. Por otra parte, la HPLC analítica también indicó el consumo completo del éster activado y la formación del producto final. La reacción se desactivó con tampón TEAB (0,1 M, 10 ml) y se cargó en una columna DEAE Sephadex (columna de Biotage de 25 g). La columna se eluyó con gradiente como se muestra a continuación en la Tabla 2.

A: tampón TEAB 0,1 M (CH_3CN al 10 %)

B: tampón TEAB 1 M (CH_3CN al 10 %).

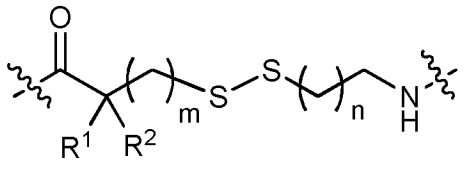
Gradiente:

Tabla 2

ETAPA	Mezcla de disolventes (% de B)	Longitud (ml)
1	0	100
2	0 - 45	50
3	45	100
4	45 - 100	50
5	100	100

El producto deseado se eluyó del 45 % al 100 % de tampón TEAB 1 M. La fracción que contenía el producto se combinó, se evaporó y se purificó por HPLC (columna YLC, 8 ml/min). Rendimiento: 53 %.

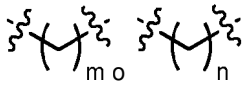
En resumen, la presente invención se refiere a un nucleósido o nucleótido unido covalentemente a un fluoróforo a través de un enlazador, en donde dicho enlazador comprende una estructura de fórmula (I):



(I),

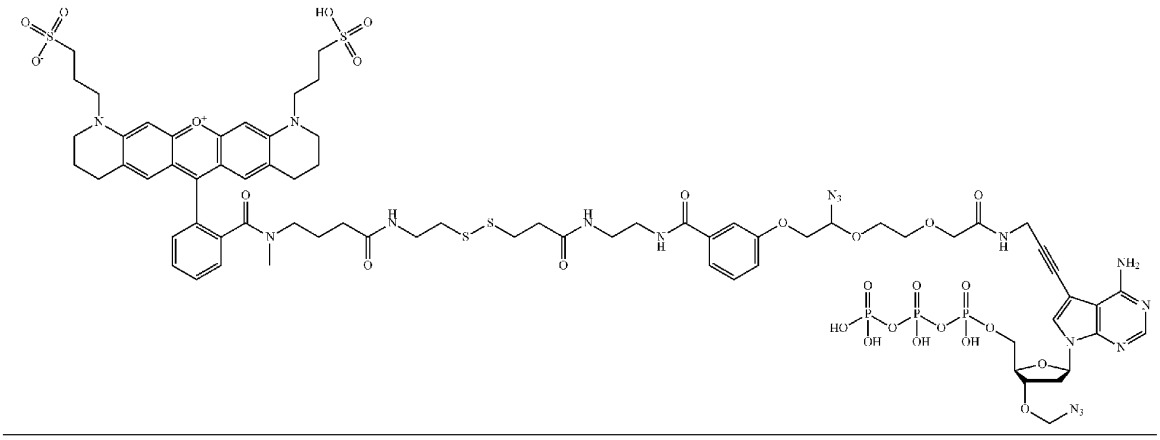
en donde

cada R^1 y R^2 es alquilo C_{1-6} opcionalmente sustituido de manera independiente;

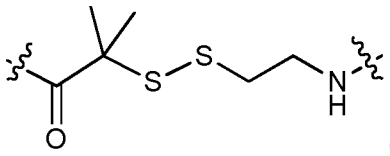
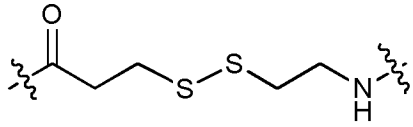
cada una de las unidades de repetición de metileno en  está opcionalmente sustituida;

m es un número entero de 0 a 20; y

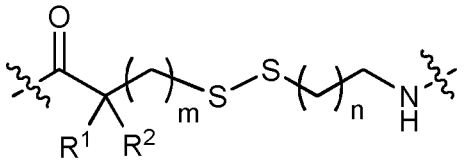
n es un número entero de 1 a 20, con la condición de que el nucleósido o nucleótido marcado con fluoróforo no tenga la estructura:



En algunos casos, el nucleósido o nucleótido al que se hace referencia anteriormente, la estructura de fórmula (I) también se representa por la fórmula (Ia) o (Ib):

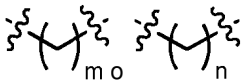

(Ia),

(Ib).

Más particularmente, la presente invención puede referirse a un nucleósido o nucleótido unido covalentemente a un fluoróforo a través de un enlazador, en donde dicho enlazador comprende una estructura de fórmula (I):


(I), en donde

R^1 se selecciona de alquilo C_{1-6} opcionalmente sustituido;

R^2 se selecciona de hidrógeno o alquilo C_{1-6} opcionalmente sustituido;

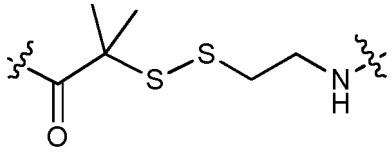
cada una de las unidades de repetición de metileno en  está opcionalmente sustituida;

X se selecciona de metileno (CH_2), oxígeno (O) o azufre (S);

m es un número entero de 0 a 20; y

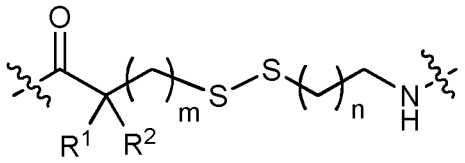
n es un número entero de 1 a 20.

Se prefiere que la estructura de fórmula (I) también se represente por la fórmula (Ia):


(Ia).

REIVINDICACIONES

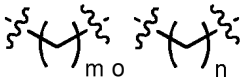
1. Un nucleósido o nucleótido unido covalentemente a un fluoróforo a través de un enlazador, en donde dicho enlazador comprende una estructura de fórmula (I):



(I), en donde

R^1 es alquilo C_{1-6} opcionalmente sustituido;

R^2 es alquilo C_{1-6} opcionalmente sustituido;

cada una de las unidades de repetición de metileno en  está opcionalmente sustituida;

m es un número entero de 0 a 20; y

n es un número entero de 1 a 20.

2. El nucleósido o nucleótido de la reivindicación 1, en donde R^1 es metilo.

3. El nucleósido o nucleótido de las reivindicaciones 1 o 2, en donde R^2 es alquilo C_{1-6} opcionalmente sustituido.

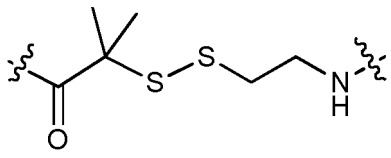
4. El nucleósido o nucleótido de las reivindicaciones 1 a 3, en donde R^2 es metilo.

5. El nucleósido o nucleótido de una cualquiera de las reivindicaciones 1 a 4, en donde m es 0.

6. El nucleósido o nucleótido de una cualquiera de las reivindicaciones 1 a 5, en donde m es 1.

7. El nucleósido o nucleótido de una cualquiera de las reivindicaciones 1 a 6, en donde n es 1.

8. El nucleósido o nucleótido de la reivindicación 1, en donde la estructura de fórmula (I) también se representa por la fórmula (Ia):



(Ia).

9. El nucleósido o nucleótido de una cualquiera de las reivindicaciones 1 a 8, en donde el nucleósido o nucleótido está unido al lado izquierdo del enlazador.

10. Un kit que comprende un nucleósido o nucleótido según una cualquiera de las reivindicaciones 1 a 9.

11. Un reactivo para modificar un nucleósido o un nucleótido, que comprende un fluoróforo y un enlazador según una cualquiera de las reivindicaciones 1 a 9.

12. Un método para la detección de un nucleósido o nucleótido que se ha incorporado en un polinucleótido, que comprende:

(a) la incorporación de un nucleósido o nucleótido según una cualquiera de las reivindicaciones 1 a 9 en un polinucleótido; y

(b) la detección de una señal fluorescente procedente de dicho nucleósido o nucleótido que se incorporó en la etapa (a).

13. El método de la reivindicación 12, que comprende además proporcionar una cadena de ácido nucleico molde y una cadena de ácido nucleico parcialmente hibridada, en donde la etapa (a) incorpora en la cadena hibridada al menos un nucleósido o nucleótido que es complementario con un nucleósido o nucleótido en la posición correspondiente de la cadena molde, y en donde la etapa (b) identifica la base del nucleósido o nucleótido incorporado, indicando de este modo la identidad del nucleósido o nucleótido complementario de la cadena molde.

14. Un método de secuenciación de una molécula de ácido nucleico molde, que comprende:

incorporar uno o más nucleótidos marcados en una cadena de ácido nucleico complementaria con el ácido nucleico molde;

determinar la identidad de la base presente en uno o más nucleótidos marcados incorporados para determinar la secuencia de la molécula de ácido nucleico molde;

en donde la identidad de la base presente en dichos uno o más nucleótidos marcados se determina detectando una señal fluorescente producida por dichos nucleótidos marcados; y en donde al menos un nucleótido marcado incorporado es un nucleótido según una cualquiera de las reivindicaciones 1 a 9.

15. El método de la reivindicación 14, en donde la identidad de la base presente en dichos uno o más nucleótidos se determina después de cada etapa de incorporación de nucleótidos.

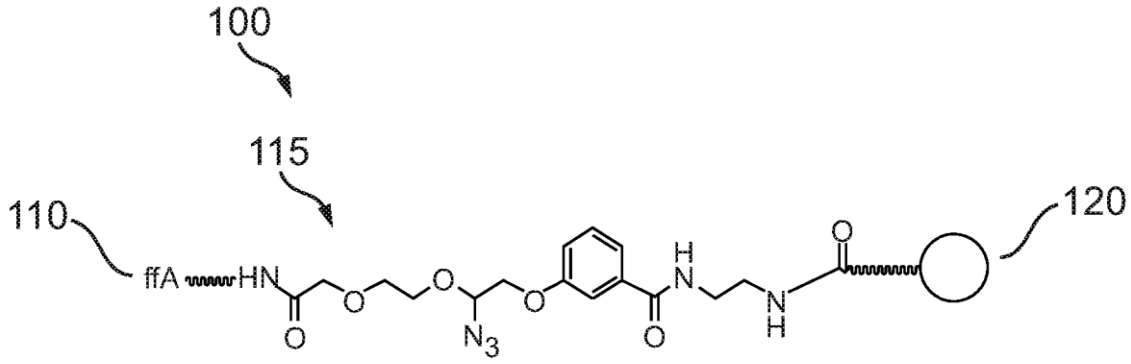

100
115
110
120
ffA
HN
O
O
O
O
O
N_3
H
N
O
N
H

Fig. 1A

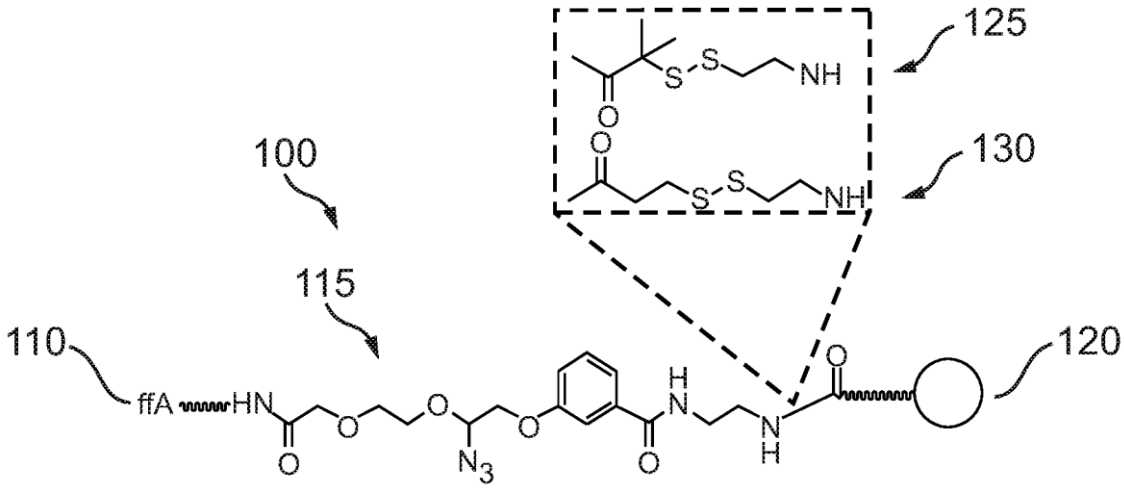

125
130
100
115
110
120
S
S
NH
O
ffA
HN
N_3

Fig. 1B

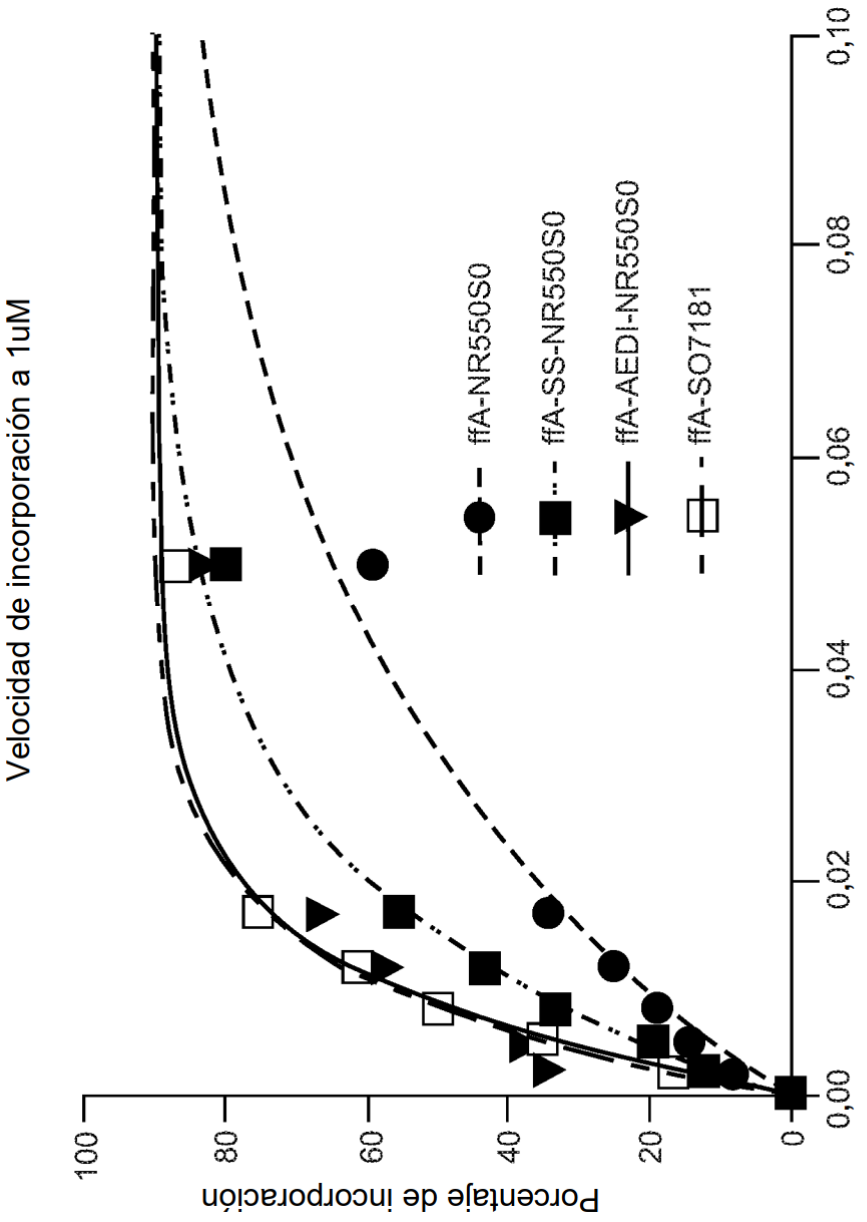

Velocidad de incorporación a 1uM
Porcentaje de incorporación
100
80
60
40
20
0
0,00
0,02
0,04
0,06
0,08
0,10
Tiempo (min)
ffA-NR550S0
ffA-SS-NR550S0
ffA-AEDI-NR550S0
ffA-SO7181

Fig. 2

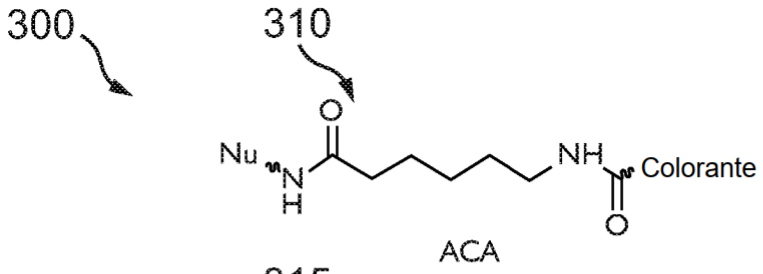


Fig. 3A

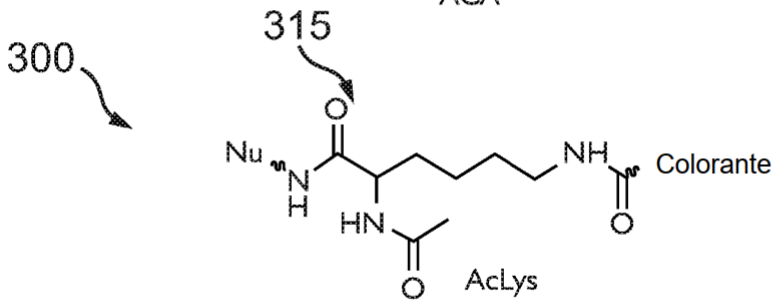


Fig. 3B

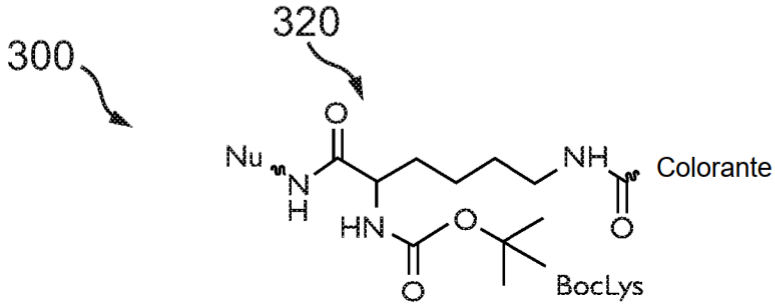


Fig. 3C

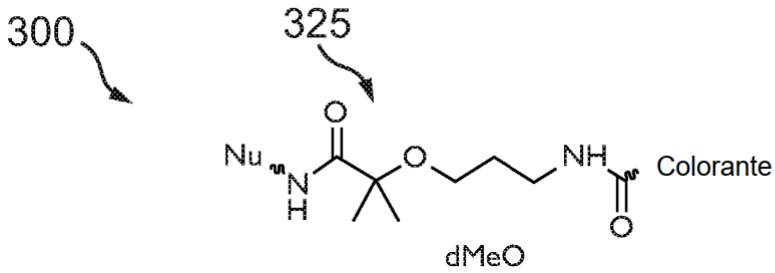


Fig. 3D

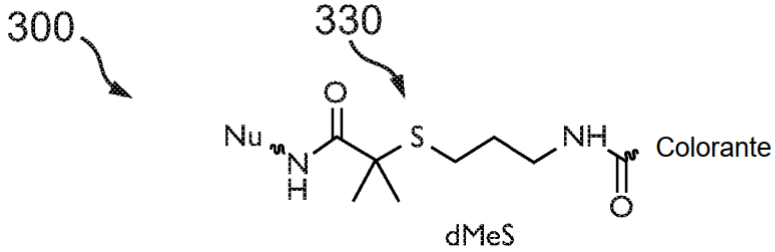


Fig. 3E

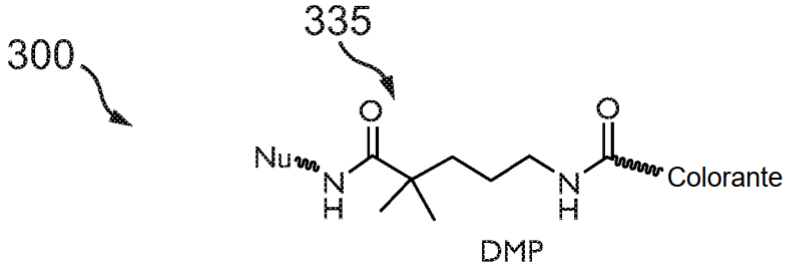


Fig. 3F

Conjunto de colorante	Densidad	%PF	Ph R1	PPh R1	Ph R2	PPh R2	% Alineam. R1	% Alineam. R2	error R1	error R2
V4	230	79,6	0,205	0,111	0,35	0,149	82,2	79,3	1,41	3,19
V5,75	187	77,65	0,371	0,048	0,442	0,164	80,06	79,99	2,16	3,94
V10/cian-peg4	222	79,88	0,305	0,084	0,53	0,179	81,89	79,45	1,42	3,48
V10/cian-peg4, A-AEDI550S0	223	81,83	0,14	0,123	0,294	0,197	81,61	67,12	0,85	5,84
V10/cian-peg4, A-AcLys550S0	227	78,52	0,2	0,128	0,319	0,272	82,23	78,52	1,44	3,48
V10/cian, A-AcLys550S0	145	86,27	0,178	0,211	0,283	0,29	82,4	79,84	1,33	3,33

Fig. 4

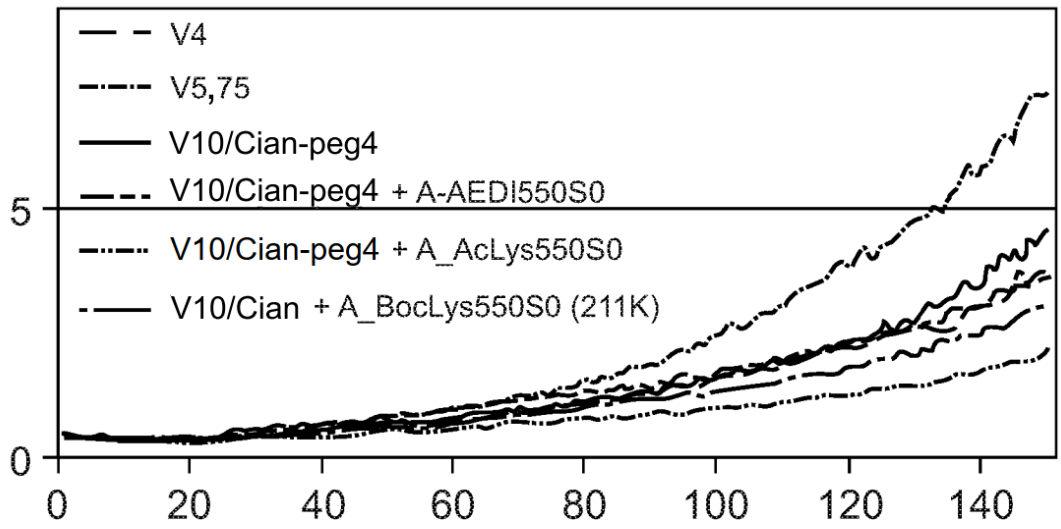

V4
V5,75
V10/Cian-peg4
V10/Cian-peg4 + A-AEDI550S0
V10/Cian-peg4 + A_AcLys550S0
V10/Cian + A_BocLys550S0 (211K)
5
0
0
20
40
60
80
100
120
140

Fig. 5A

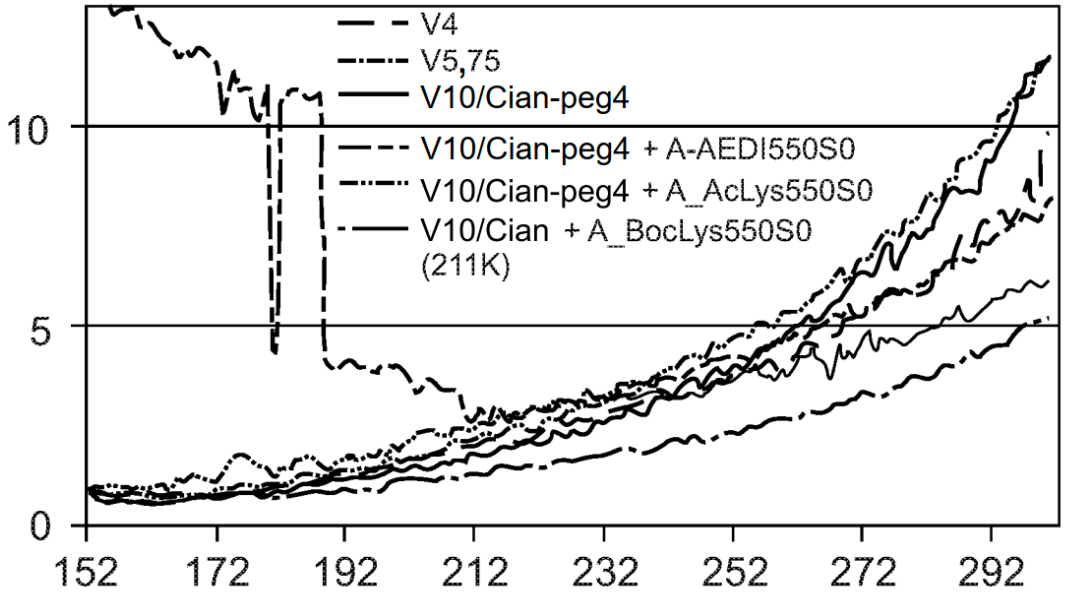

V4
V5,75
V10/Cian-peg4
V10/Cian-peg4 + A-AEDI550S0
V10/Cian-peg4 + A_AcLys550S0
V10/Cian + A_BocLys550S0 (211K)
10
5
0
152
172
192
212
232
252
272
292

Fig. 5B

Conjunto de colorante	Densidad	%PF	Ph R1	PPh R1	Ph R2	PPh R2	% Alineam. R1	% Alineam. R2	error R1	error R2
V4	191	82,62	0,132	0,151	0,242	0,163	81,91	79,59	1,34	2,43
V5,75	235	79,02	0,212	0,05	0,304	0,187	82,54	78,74	1,4	4,2
V10/Cian-peg4,	213	84,7	0,2	0,063	0,348	0,179	83,3	81,19	0,96	2,02
V10/Cian-peg4,	220	83,03	0,188	0,115	0,317	0,207	82,92	80,9	1,11	2,16
V10/Cian-peg4, A-AEDI-550S0	252	76,37	0,102	0,097	0,132	0,412	80,1	34,9	0,84	25,58
V10/Cian-peg4, A-ACA550S0	204	85,88	0,145	0,076	0,271	0,222	83,21	80,58	0,96	2,39

Fig. 6

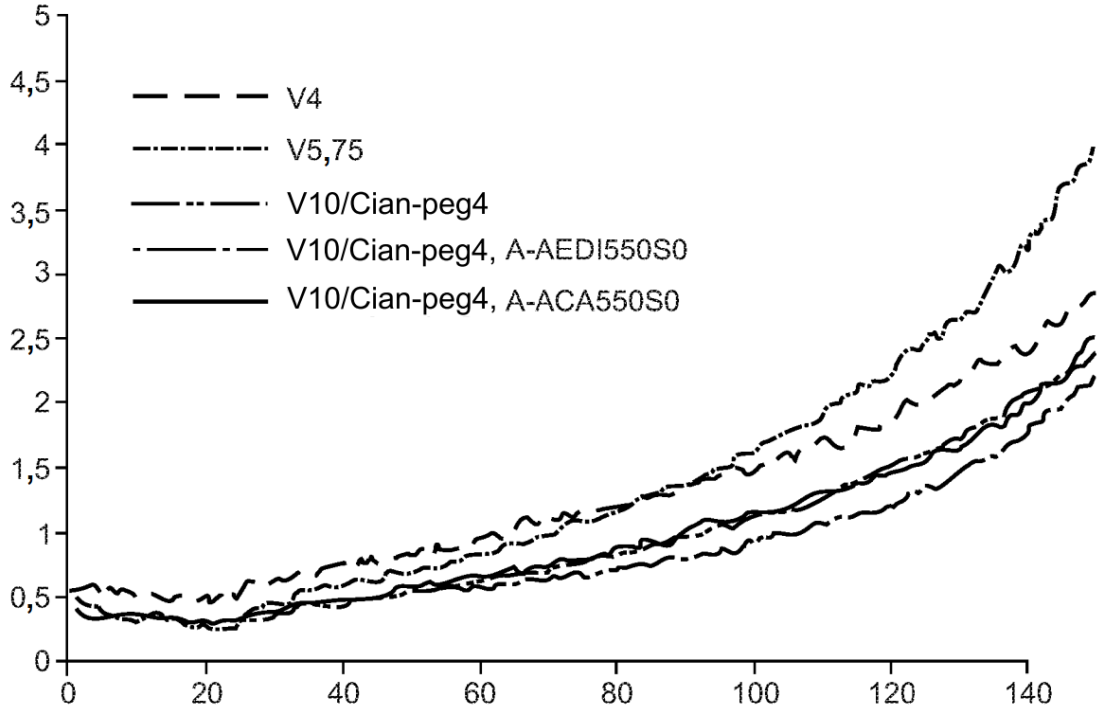

5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
0
20
40
60
80
100
120
140
V4
V5,75
V10/Cian-peg4
V10/Cian-peg4, A-AEDI550S0
V10/Cian-peg4, A-ACA550S0

Fig. 7A

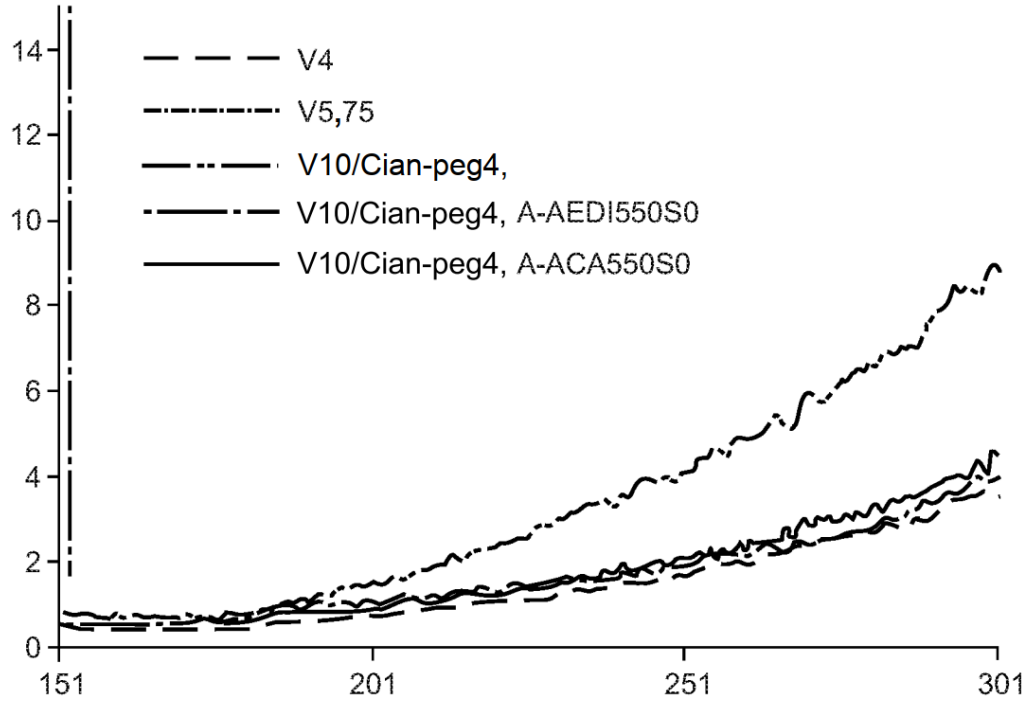

14
12
10
8
6
4
2
0
151
201
251
301
V4
V5,75
V10/Cian-peg4,
V10/Cian-peg4, A-AEDI550S0
V10/Cian-peg4, A-ACA550S0

Fig. 7B

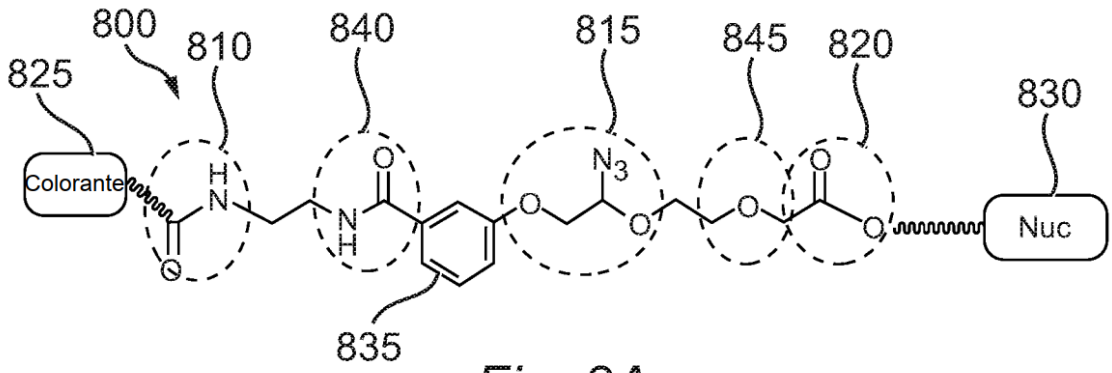


Fig. 8A

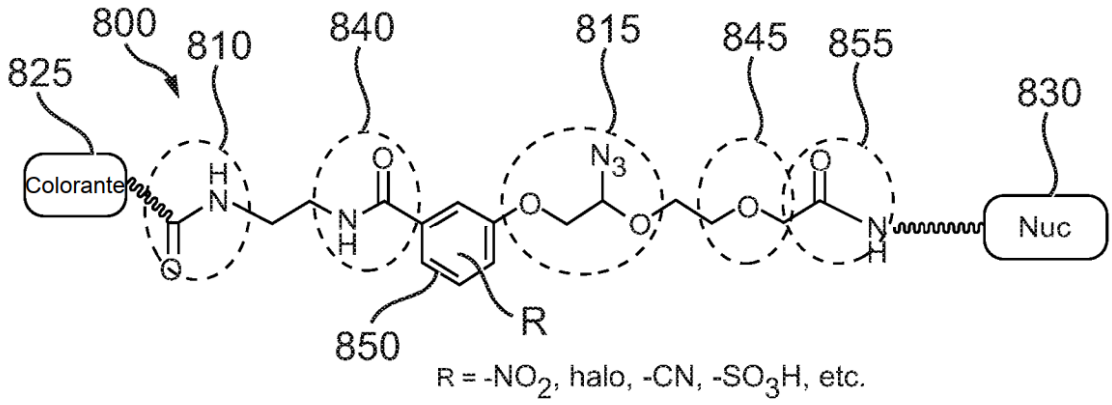


Fig. 8B

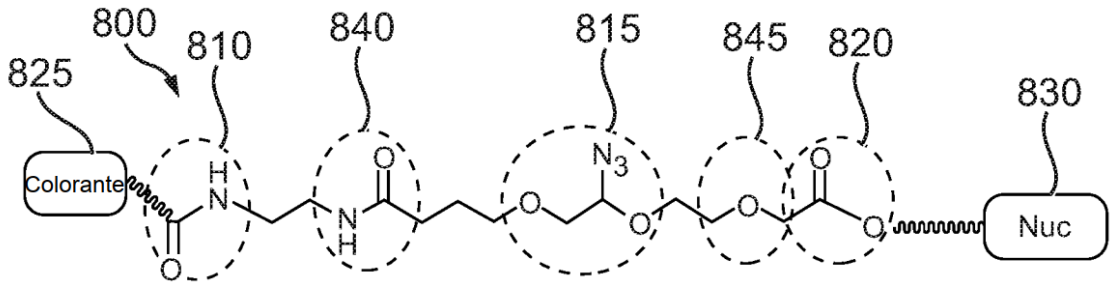


Fig. 8C

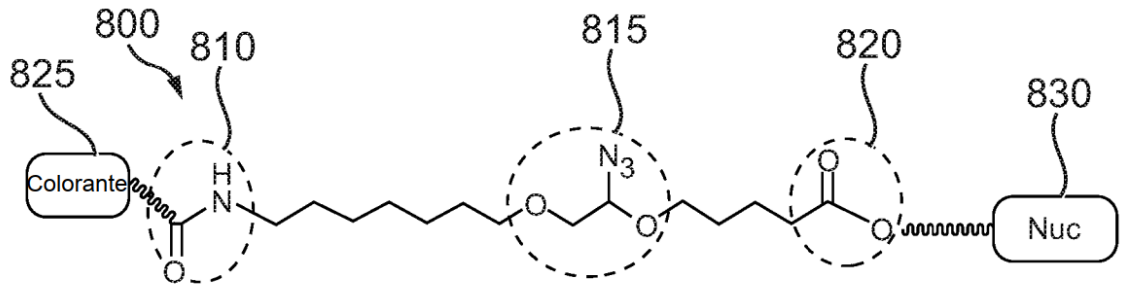

800
810
815
820
825
830
Colorante
H
N
O
N_3
O
O
O
O
Nuc

Fig. 8D

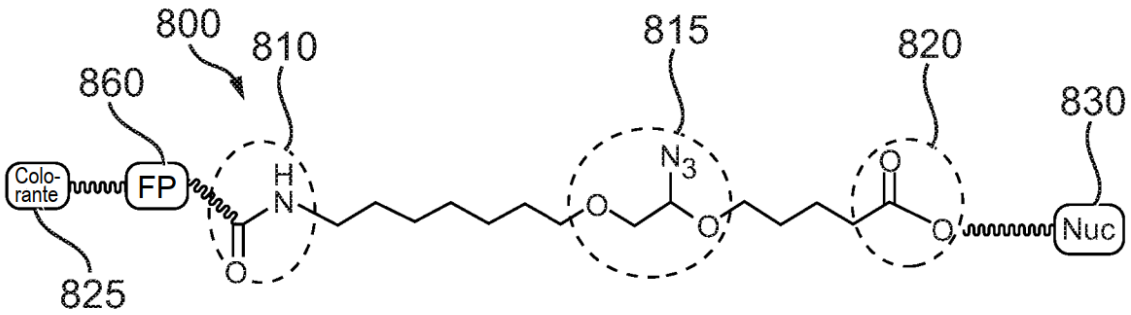

800
810
815
820
860
830
Colo-
rante
FP
H
N
O
N_3
O
O
O
O
Nuc
825

Fig. 8E

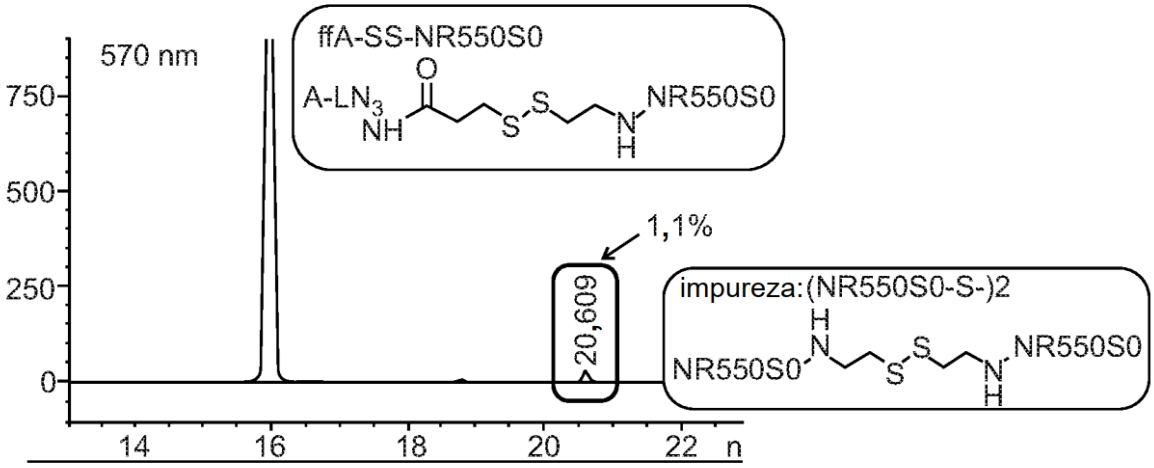


Fig. 9A

	Tampones	Resultados*	
	21 horas a 60 ºC	ffA-SS-NR550S0	ffA-AEDI-NR550S0
1	Tampón PR2	~50 % de producto secundario de disulfuro	~6 % de producto secundario de disulfuro
2	Tampón de incorporación (pH 9,9) (sin Mg2+, sin Chaps)	~10 % de producto secundario de disulfuro	Sin producto secundario de disulfuro
3	IMX	~38 % de producto secundario de disulfuro	Sin producto secundario de disulfuro
4	IMX + Pol 812	~34 % de producto secundario de disulfuro	Sin producto secundario de disulfuro
5	SRE	Sin producto secundario de disulfuro	Sin producto secundario de disulfuro
6	TEAB (0,1M): CH3CN (1,1)	~20 % de producto secundario de disulfuro	Sin producto secundario de disulfuro

Fig. 9B

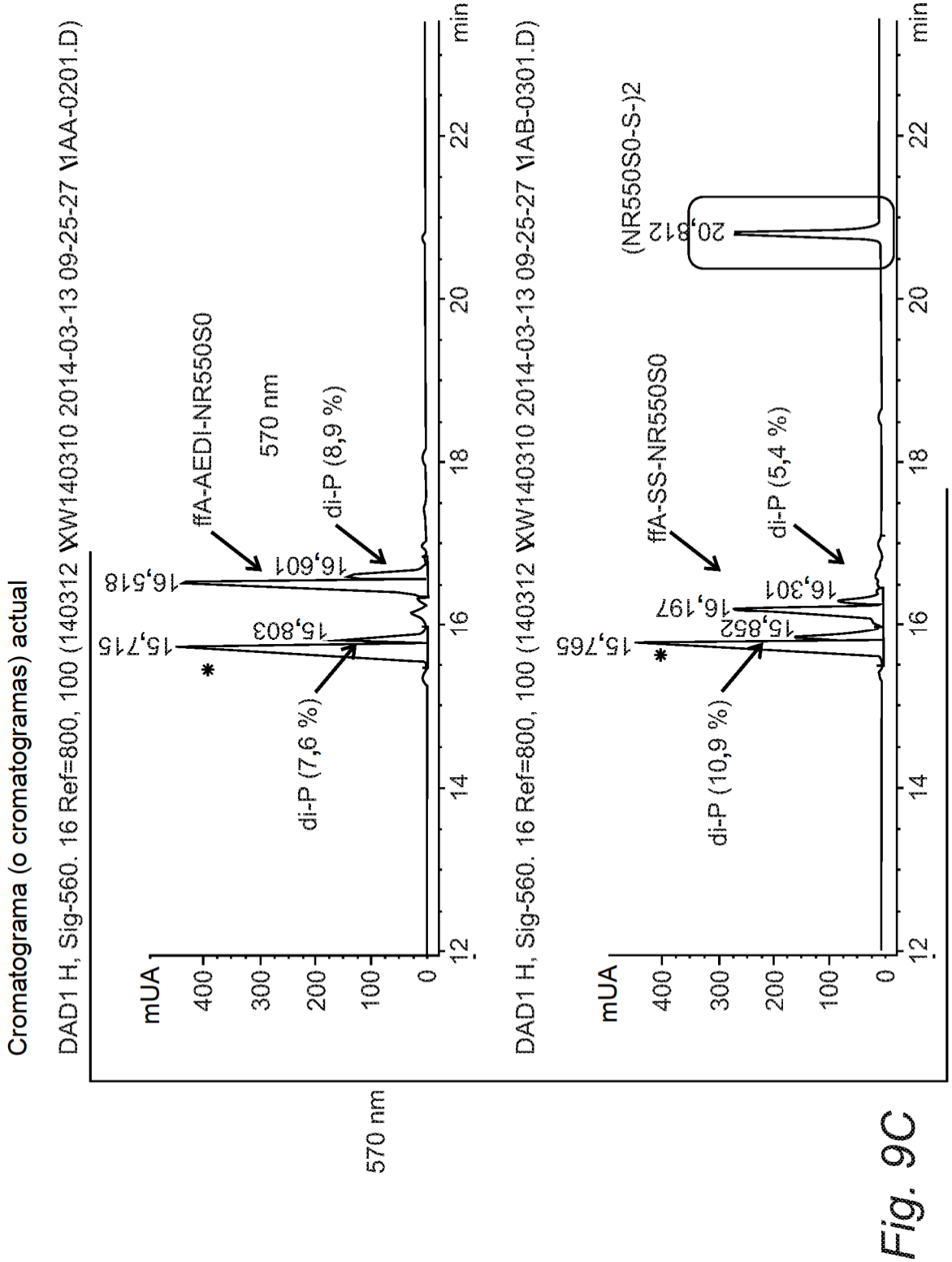

Cromatograma (o cromatogramas) actual
DAD1 H, Sig-560. 16 Ref=800, 100 (140312 \KW140310 2014-03-13 09-25-27 \1AA-0201.D)
570 nm
mUA
400
300
200
100
0
12
14
16
18
20
22
min
15,715
15,803
16,518
16,601
ffA-AEDI-NR550S0
570 nm
di-P (7,6 %)
di-P (8,9 %)
*
DAD1 H, Sig-560. 16 Ref=800, 100 (140312 \KW140310 2014-03-13 09-25-27 \1AB-0301.D)
mUA
400
300
200
100
0
12
14
16
18
20
22
min
15,765
15,852
16,197
16,301
20,812
ffA-SS-NR550S0
(NR550S0-S-)2
di-P (10,9 %)
di-P (5,4 %)
*

Fig. 9C

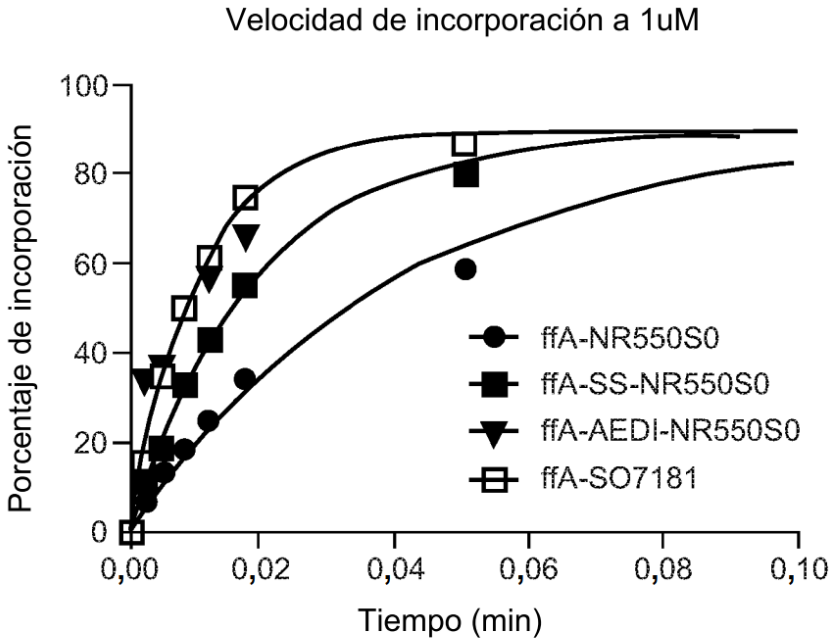


Fig. 10A

ffA	K (uM/min)
NR550S0	22 (1X)
-SS-NR550S0	54 (2X)
-AEDI-NR550S0	97 (4X)
-SO7181 conv.	99 (4X)

Fig. 10B

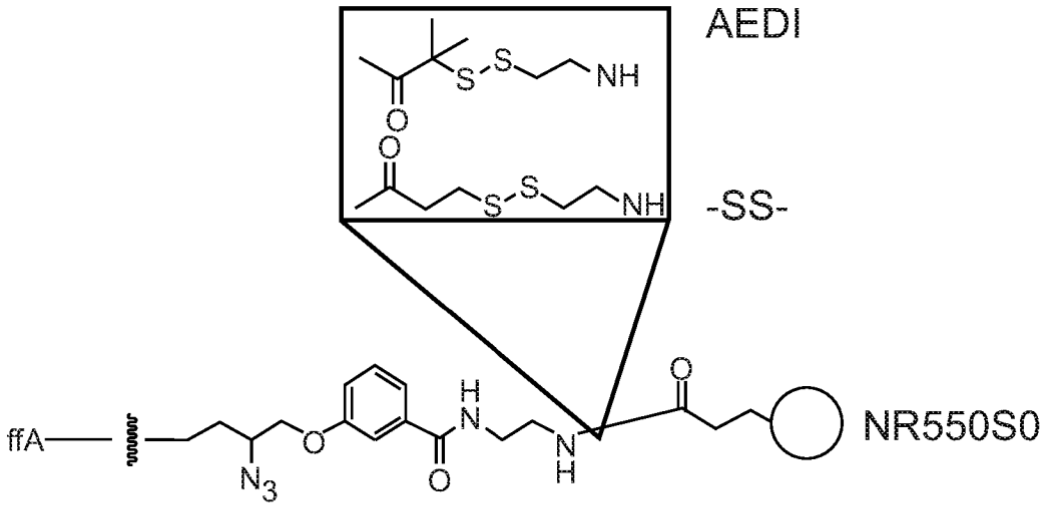


Fig. 10C

Combinación de V10 con distintos A-550S0
(misma concentración)
Diagrama de dispersión para mosaico 1 ciclo 2

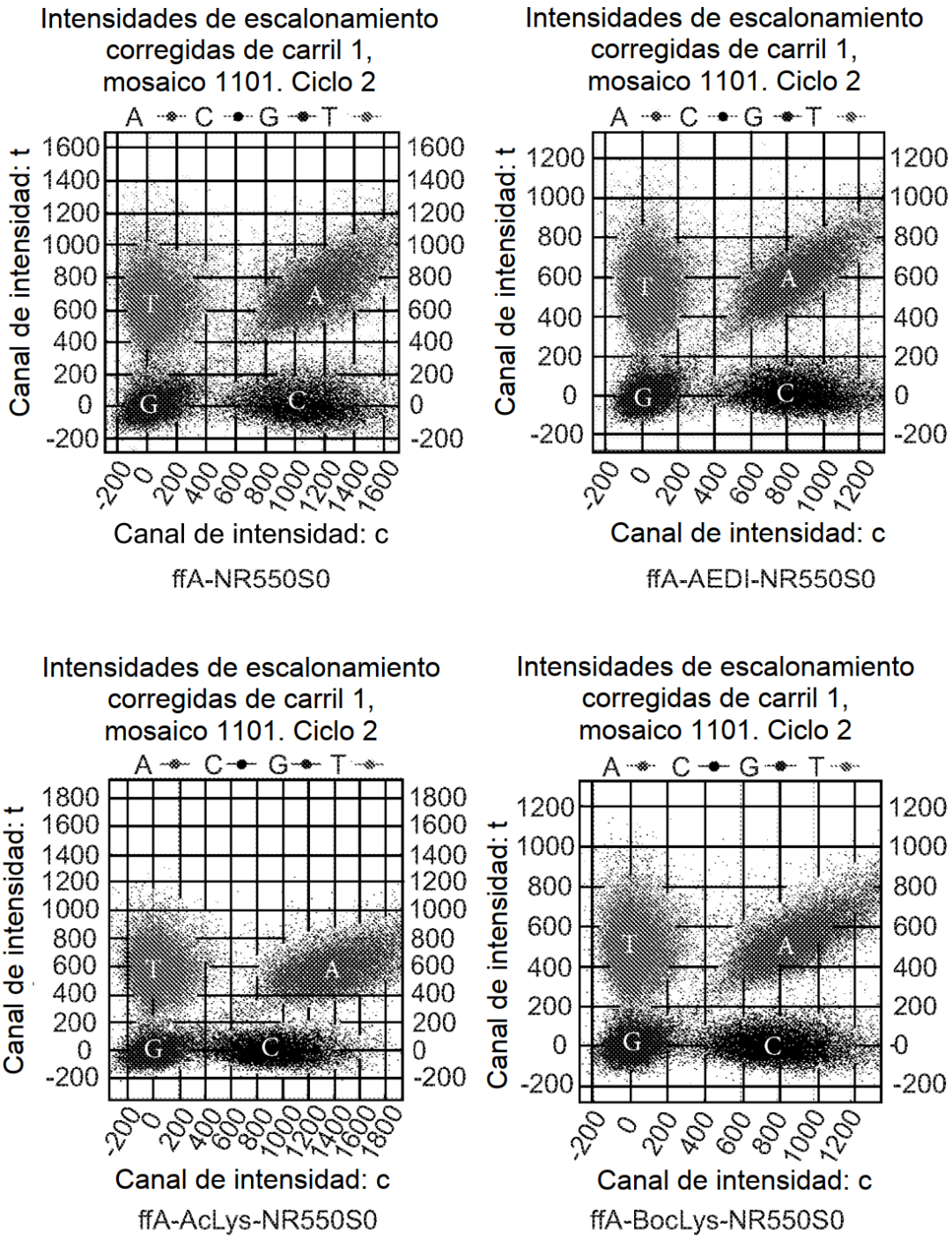


Fig. 11A

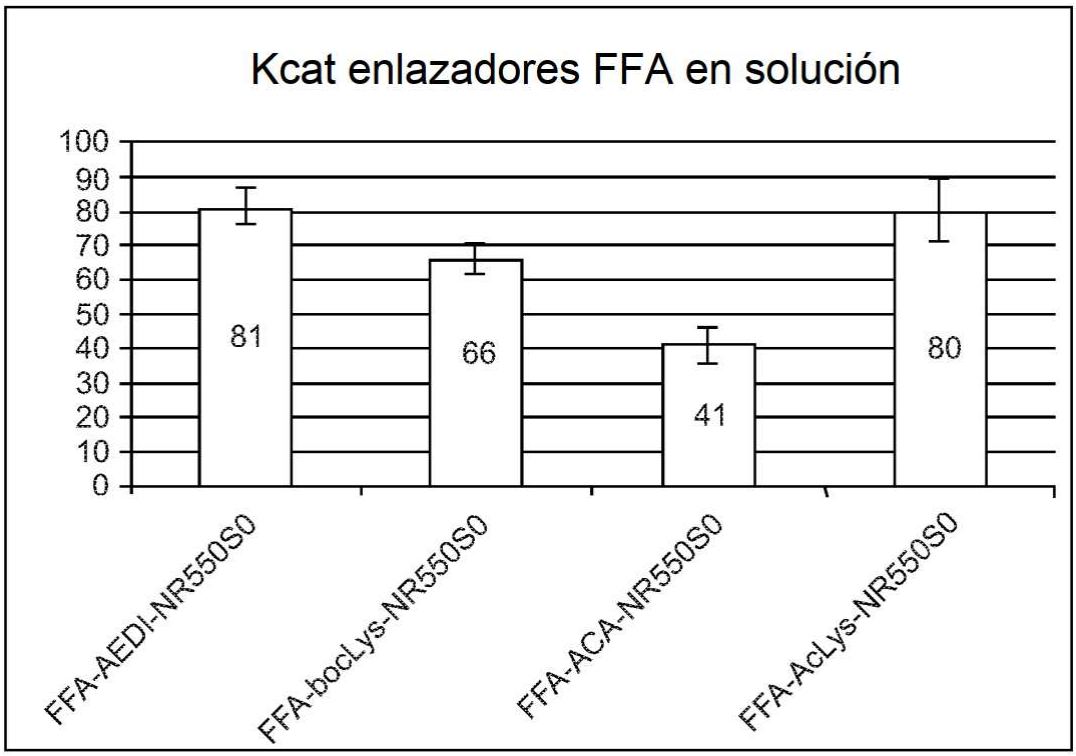

Kcat enlazadores FFA en solución
100
90
80
70
60
50
40
30
20
10
0
81
66
41
80
FFA-AEDI-NR550S0
FFA-bocLys-NR550S0
FFA-ACA-NR550S0
FFA-AcLys-NR550S0

Fig. 11B

Conjunto de colorante	Densidad	%PF	Ph R1	PPh R1	Ph R2	PPh R2	% Alineam. R1	% Alineam. R2	error R1	error R2
ffNs + ffA-NR550S0	222	79,88	0,305	0,084	0,53	0,179	81,89	79,45	1,42	3,48
ffNs + ffA-AEDI-NR550S0	223	81,83	0,14	0,123	0,294	0,197	81,61	67,12*	0,85	5,84*
ffNs + ffA-AcLys-NR550S0	227	78,52	0,2	0,128	0,319	0,272	82,23	78,52	1,44	3,48
ffNs + ffA-BocLys-NR550S0	211	82,97	0,126	0,146	0,226	0,249	82,54	80,45	1,17	1,99

Fig. 12A

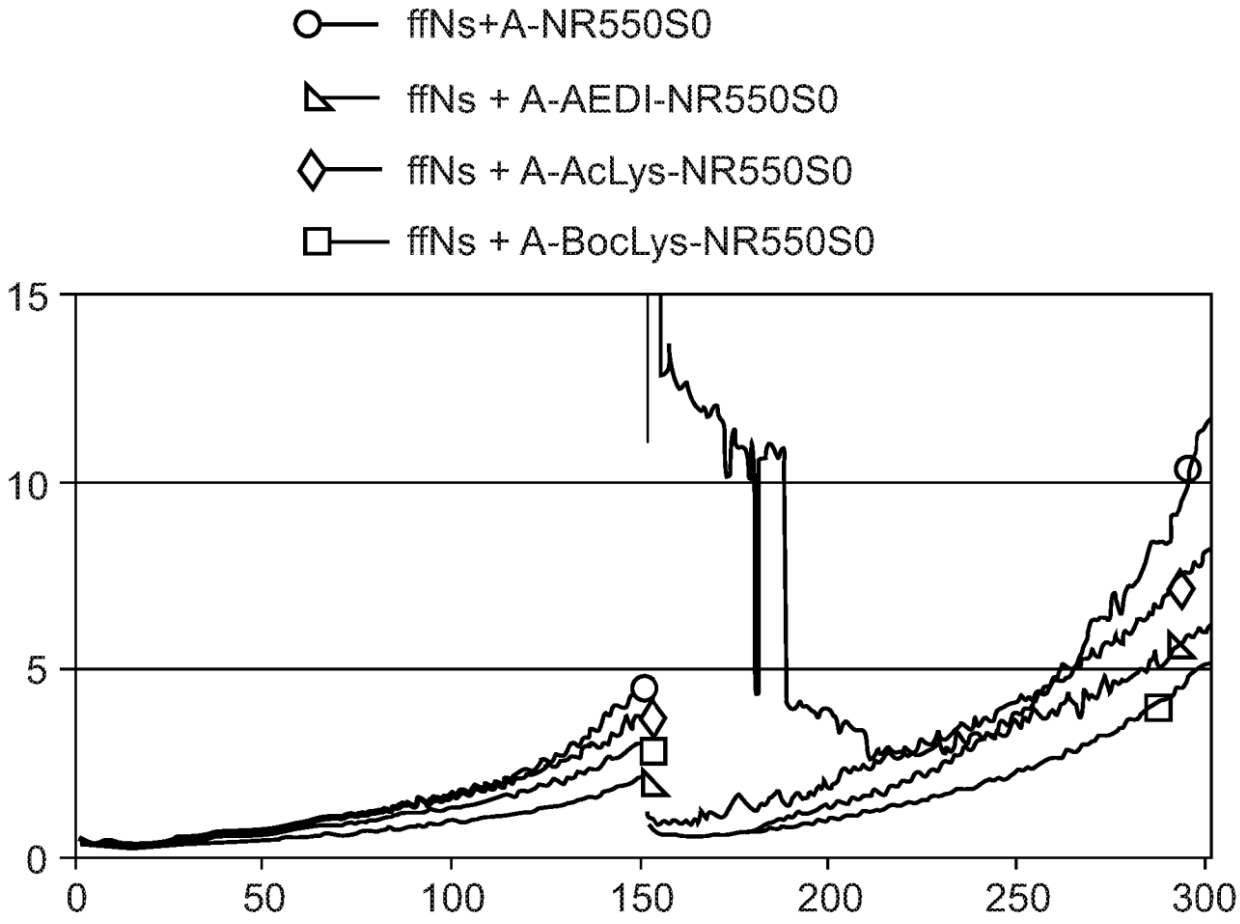

ffNs+A-NR550S0
ffNs + A-AEDI-NR550S0
ffNs + A-AcLys-NR550S0
ffNs + A-BocLys-NR550S0
15
10
5
0
0
50
100
150
200
250
300

Fig. 12B