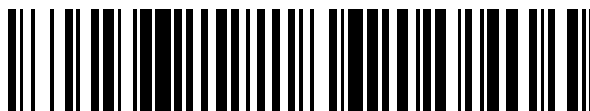


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 702 662**

51 Int. Cl.:

**C07D 209/08** (2006.01)

**C09B 23/02** (2006.01)

**G01N 33/533** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.05.2015** **PCT/GB2015/051337**

87 Fecha y número de publicación internacional: **12.11.2015** **WO15170102**

96 Fecha de presentación y número de la solicitud europea: **07.05.2015** **E 15723996 (3)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018** **EP 3140285**

54 Título: **Compuestos de polimetina y su uso como marcadores fluorescentes**

30 Prioridad:

**07.05.2014 GB 201408077**

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

**04.03.2019**

73 Titular/es:

**ILLUMINA CAMBRIDGE LIMITED (100.0%)**  
**Chesterford Research Park Little Chesterford**  
**Saffron Walden**  
**Essex CB10 1XL, GB**

72 Inventor/es:

**ROMANOV, NIKOLAI NIKOLAEVICH**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

ES 2 702 662 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

Compuestos de polimetina y su uso como marcadores fluorescentes

La presente divulgación se refiere a nuevos compuestos de polimetina y su uso como marcadores fluorescentes. En particular, los compuestos se pueden utilizar como marcadores fluorescentes para nucleótidos en aplicaciones de secuenciación de ácidos nucleicos.

## Antecedentes

La detección no radiactiva de ácidos nucleicos que utilizan marcadores fluorescentes es una tecnología importante en biología molecular. Muchos procedimientos empleados en la tecnología de ADN recombinante se basaban en gran medida en el uso de nucleótidos o polinucleótidos marcados radiactivamente con, por ejemplo, <sup>32</sup>P. Los compuestos radioactivos permiten la detección sensible de ácidos nucleicos y otras moléculas de interés. Sin embargo, existen serias limitaciones en el uso de isótopos radiactivos, como su coste, vida útil limitada y, lo que es más importante, consideraciones de seguridad. Eliminar la necesidad de marcadores radiactivos mejora la seguridad al tiempo que reduce el impacto ambiental y los costes asociados, por ejemplo, con la eliminación de reactivos. Los métodos susceptibles de detección fluorescente no radiactiva incluyen, a modo de ejemplo no limitante, secuenciación automatizada de ADN, métodos de hibridación, detección en tiempo real de productos de reacción en cadena de la polimerasa e inmunoensayos. Para muchas aplicaciones es deseable emplear múltiples marcadores fluorescentes distinguibles espectralmente para lograr la detección independiente de una pluralidad de analitos que se superponen espacialmente. En dichos métodos multiplex, se puede reducir el número de recipientes de reacción, lo que simplifica los protocolos experimentales y facilita la producción de kits de reactivos específicos a aplicación. En la secuenciación de ADN automatizada en múltiples colores, por ejemplo, la detección fluorescente múltiple permite el análisis de múltiples bases de nucleótidos en una única línea de electroforesis, lo que aumenta el rendimiento sobre los métodos de un solo color y reduce las incertidumbres asociadas con las variaciones de movilidad electroforética entre carriles.

Sin embargo, la detección de fluorescencia múltiple puede ser problemática y existe una serie de factores importantes que restringen la selección de marcadores fluorescentes. Primero, puede ser difícil encontrar compuestos de tinte cuyos espectros de emisión se resuelvan de manera espectral adecuadamente en una aplicación dada. Adicionalmente, cuando se utilizan diversos tintes fluorescentes juntos, para generar señales de fluorescencia en regiones espectrales distinguibles por excitación simultánea puede ser difícil porque las bandas de absorción de los tintes que podrían ser utilizadas para esto usualmente están ampliamente separadas, por lo que es difícil lograr más o menos eficiencia de excitación de fluorescencia igual incluso para dos tintes. Muchos métodos de excitación utilizan fuentes de luz de alta potencia como los láseres y, por lo tanto, el tinte debe tener suficiente estabilidad fotográfica para soportar dicha excitación. Una consideración final de particular importancia en los métodos de biología molecular es el grado en que los tintes fluorescentes deben ser compatibles con los reactivos químicos utilizados, tales como, por ejemplo, solventes y reactivos de síntesis de ADN, tampones, enzimas polimerasas y enzimas ligasa.

A medida que la tecnología de secuenciación avanza, se ha desarrollado una necesidad de compuestos de tintes fluorescentes adicionales, sus conjugados de ácido nucleico y conjuntos de tintes que satisfagan todas las restricciones anteriores y que sean particularmente adecuados para métodos moleculares de alto rendimiento tales como la secuenciación en fase sólida y similares.

Las moléculas de tinte fluorescente con propiedades de fluorescencia mejoradas tales como intensidad de fluorescencia, forma y máximo de longitud de onda de la banda de fluorescencia pueden mejorar la velocidad y la precisión de la secuenciación de ácidos nucleicos. La fuerte señal de fluorescencia es especialmente importante cuando las mediciones se realizan en tampones biológicos con base en agua y a una temperatura más alta, ya que la intensidad de fluorescencia de la mayoría de los tintes es significativamente menor en dichas condiciones. Más aún, la naturaleza de la base a la que se une un tinte también afecta el máximo de fluorescencia, la intensidad de la fluorescencia y otras propiedades del tinte espectral. Las interacciones específicas de secuencia entre las nucleobases y los tintes fluorescentes se pueden adaptar mediante el diseño específico de los tintes fluorescentes. La optimización de la estructura de los tintes fluorescentes puede mejorar la eficiencia de la incorporación de nucleótidos, reducir el nivel de errores de secuenciación y disminuir el uso de reactivos y, por lo tanto, los costes de la secuenciación de ácidos nucleicos.

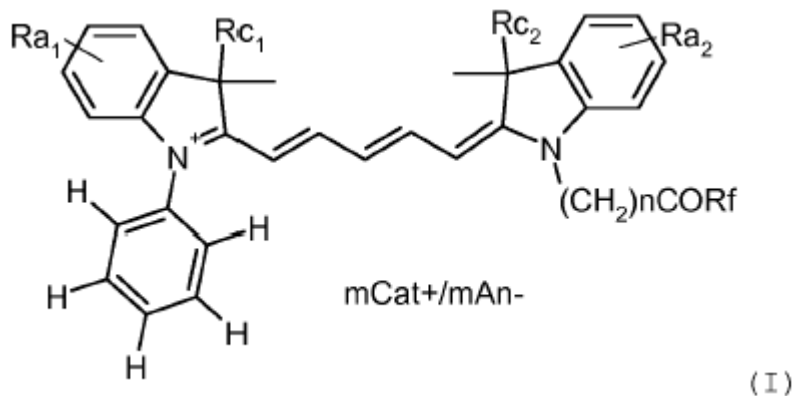
Los compuestos de polimetina se conocen de publicaciones anteriores. Ciertos compuestos de polimetina se describen en publicaciones que incluyen los documentos WO2013041117, WO2005014723, EP1792949 y WO2014135221.

En el presente documento se describen construcciones de polimetina mejoradas y su uso como marcadores de biomoléculas, particularmente como marcadores para nucleótidos utilizados en la secuenciación de ácidos nucleicos. Se pueden observar mejoras particulares en la eficiencia de la incorporación de nucleótidos marcados y la longitud de la secuencia de lectura que se puede obtener utilizando las nuevas construcciones fluorescentes.

## Resumen

La presente invención se define por las reivindicaciones.

De acuerdo con un primer aspecto esta divulgación proporciona compuestos de la fórmula (I) o formas mesoméricas del mismo:



5 en la que mCat+ o mAn- es un contraión orgánico o inorgánico cargado positivamente/negativamente y

m es un entero 0-3;

n es un entero 1-5;

Rf es OH u O-;

10 cada uno de Ra<sub>1</sub> y Ra<sub>2</sub> es independientemente H, SO<sub>3</sub>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente; en la que cualquiera de Ra<sub>1</sub> o Ra<sub>2</sub> es sulfonamida o SO<sub>3</sub>-; y cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido.

En ciertos ejemplos en los que Ra<sub>1</sub> o Ra<sub>2</sub> es SO<sub>3</sub>-; cualquiera de Rc<sub>1</sub> o Rc<sub>2</sub> puede ser un grupo ácido alquil sulfónico.

15 En otra realización los compuestos de la presente divulgación se pueden conjugar con una variedad de fracciones de sustrato tales como, por ejemplo, nucleósidos, nucleótidos, polinucleótidos, polipéptidos, carbohidratos, ligandos, partículas, células, superficies semisólidas (por ejemplo geles) y superficies sólidas. La conjugación se puede llevar a cabo a través del grupo carboxilo CORf, que se puede convertir en una amida o éster.

Por lo tanto, de acuerdo con un aspecto adicional de la divulgación, se proporcionan compuestos de tinte que comprenden grupos ligadores para permitir, por ejemplo, adhesión covalente a dichas fracciones de sustrato.

20 De acuerdo con un aspecto adicional la divulgación proporciona un compuesto de nucleósido o nucleótido definido por la fórmula: N-L-Dye, en la que N es un nucleótido, L es una fracción de ligador opcional y Dye es un compuesto fluorescente de acuerdo con la presente divulgación.

En un aspecto adicional la divulgación proporciona métodos de secuenciación que utilizan los compuestos de tinte de la presente divulgación.

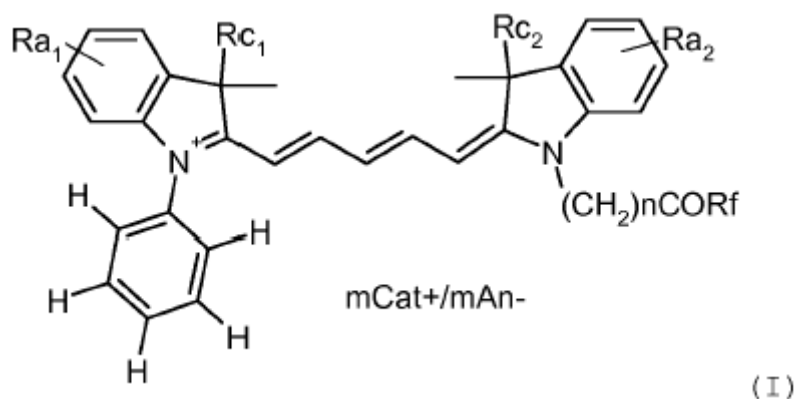
25 De acuerdo con un aspecto adicional la divulgación también proporciona kits que comprenden compuestos de tinte (libres o en forma conjugado) que se pueden utilizar en diversos ensayos inmunológicos, marcador de oligonucleótido y ácido nucleico y para secuenciación de ADN mediante síntesis. En aún otro aspecto la divulgación proporciona kits que comprenden 'conjuntos' de tinte particularmente adecuados para ciclos de secuenciación mediante síntesis sobre una plataforma de instrumentos automatizada.

Un aspecto adicional de la divulgación es la preparación química de compuestos de la divulgación.

30 Descripción detallada

Esta divulgación proporciona compuestos novedosos particularmente adecuados para métodos de detección de fluorescencia y secuenciación mediante síntesis. Los compuestos que tienen una fracción de indol N-fenilo son ventajosos en intensidad de fluorescencia, fotoestabilidad en comparación con análogos de N-alquilo y por lo tanto mejoran ciertas aplicaciones de secuenciación de ácidos nucleicos.

35 De acuerdo con un primer aspecto la divulgación proporciona compuestos de la fórmula (I) o formas mesoméricas del mismo:



en la que mCat+ o mAn- es un contraión orgánico o inorgánico cargado positivamente/negativamente y

m es un entero 0-3;

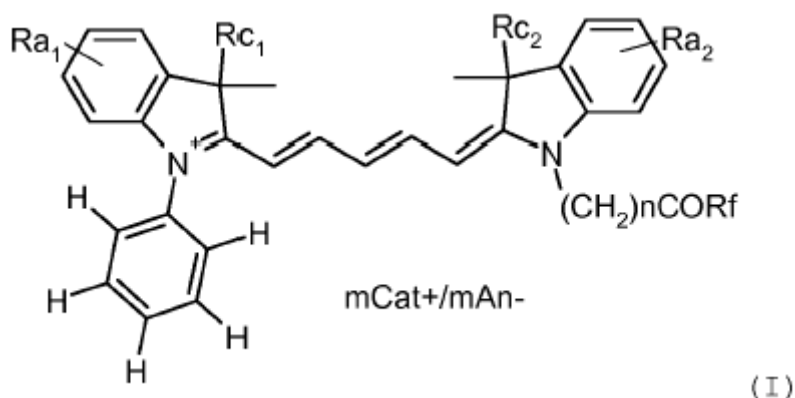
n es un entero 1-5;

5 Rf es OH u O-;

cada uno de Ra<sub>1</sub> y Ra<sub>2</sub> es independientemente H, SO<sub>3</sub><sup>-</sup>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente; en la que cualquiera de Ra<sub>1</sub> o Ra<sub>2</sub> es sulfonamida o SO<sub>3</sub><sup>-</sup>; y

cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido; en el que cuando Ra<sub>1</sub> o Ra<sub>2</sub> es SO<sub>3</sub><sup>-</sup>, cualquiera de i) Rc<sub>1</sub> o Rc<sub>2</sub> es un grupo ácido alquil sulfónico, o ii) Ra<sub>1</sub> o Ra<sub>2</sub> es una sulfonamida.

10 Un aspecto alternativo de la divulgación proporciona compuestos de la fórmula (I) o formas mesoméricas del mismo:



en las que mCat+ o mAn- es un contraión orgánico o inorgánico cargado positivamente/negativamente y

m es un entero 0-3;

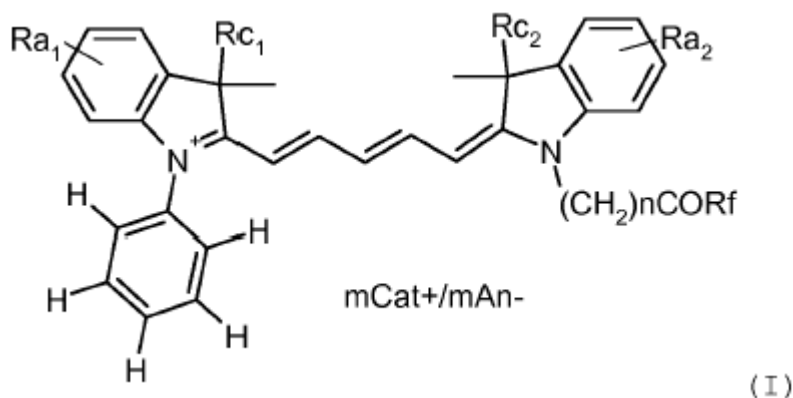
n es un entero 1-5;

15 Rf es OH u O-;

cada uno de Ra<sub>1</sub> y Ra<sub>2</sub> es independientemente H, SO<sub>3</sub><sup>-</sup>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente; en la que cualquiera de Ra<sub>1</sub> o Ra<sub>2</sub> es sulfonamida o SO<sub>3</sub><sup>-</sup>; y

cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido; en la que cuando Ra<sub>1</sub> o Ra<sub>2</sub> es SO<sub>3</sub><sup>-</sup>, cualquiera de Rc<sub>1</sub> o Rc<sub>2</sub> es un grupo ácido alquil sulfónico.

20 Un aspecto alternativo la divulgación proporciona compuestos de la fórmula (I) o formas mesoméricas del mismo:



en la que mCat+ o mAn- es un contraión orgánico o inorgánico cargado positivamente/negativamente y

m es un entero 0-3;

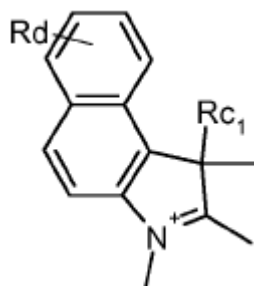
n es un entero 1-5;

5 Rf es OH u O- o un éster o conjugado de amina del mismo;

cada uno de Ra<sub>1</sub> y Ra<sub>2</sub> es independientemente H, SO<sub>3</sub><sup>-</sup>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente; en la que cualquiera de Ra<sub>1</sub> o Ra<sub>2</sub> es sulfonamida; y cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido.

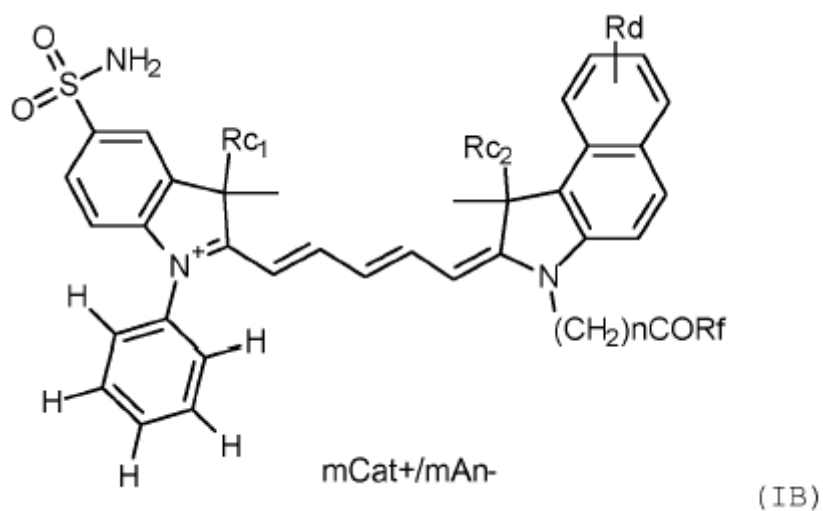
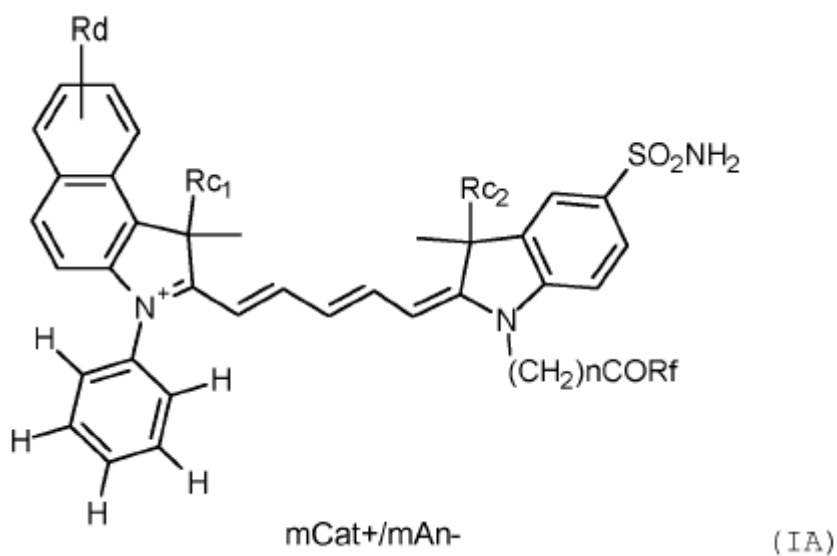
10 Las moléculas pueden contener una o más fracciones de sulfonamida o SO<sub>3</sub><sup>-</sup> en la posición Ra. Ra<sub>1</sub> y/o Ra<sub>2</sub> pueden ser sulfonamida. El otro Ra (Ra<sub>1</sub> o Ra<sub>2</sub>) puede ser independientemente H, SO<sub>3</sub><sup>-</sup>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente. Ra<sub>1</sub> o Ra<sub>2</sub> puede ser H. Ra<sub>1</sub> o Ra<sub>2</sub> pueden ser SO<sub>3</sub><sup>-</sup>. Ra<sub>1</sub> puede ser diferente a Ra<sub>2</sub>, por ejemplo la estructura puede tener un grupo de sulfonamida único a Ra<sub>1</sub>, y H como Ra<sub>2</sub>. Ra<sub>1</sub> y Ra<sub>2</sub> pueden ambos ser sulfonamida. La sulfonamida puede ser SO<sub>2</sub>NH<sub>2</sub> o SO<sub>2</sub>NHR, en la que R es un alquilo, alquilo sustituido, grupo arilo o arilo sustituido.

15 Ra<sub>1</sub> o Ra<sub>2</sub> pueden ser un anillo alifático, aromático o heterocíclico adicional fusionado a los carbonos adyacentes del anillo de indol. Por ejemplo, en dichos casos cuando se fusiona un anillo aromático, el grupo terminal de tintes puede representar una estructura de tipo



20 en la que Rd puede ser H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico.

De esta manera los tintes de la divulgación se pueden describir por la Fórmula (1A) o (1B):



en la que

m es un entero 0-3;

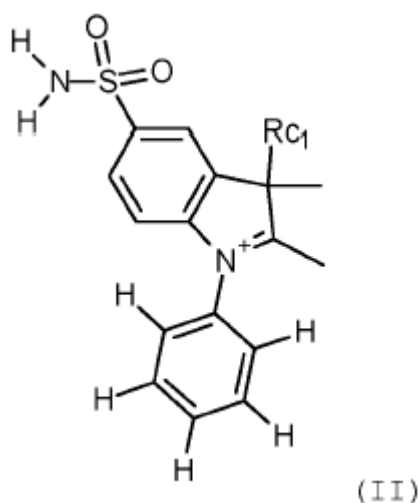
5 n es un entero 1-5;

Rf es OH u O-;

Rd es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi sulfonamida, o ácido sulfónico; y cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido; en la que cualquiera de Rc<sub>1</sub> o Rc<sub>2</sub> es un grupo ácido alquil sulfónico.

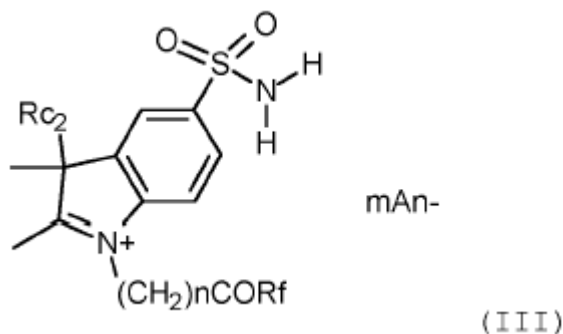
10 En la fórmula (IA) o (IB) los anillos adicionales fusionados a átomos de carbono adyacentes del anillo indol se pueden sustituir opcionalmente, por ejemplo con ácido sulfónico o sulfonamida.

El compuesto puede ser en el que la sulfonamida es parte de una estructura de fórmula (II):



en la que  $R_{C1}$  es alquilo o alquilo sustituido.

El compuesto puede ser en el que la sulfonamida es parte de una estructura de fórmula (III):



5 en la que

$m$  es un entero 0-3,  $n$  es un entero 1-5,  $R_f$  es OH u O-; y  $R_{C2}$  es alquilo o alquilo sustituido.

El grupo carboxilo CORf o sus derivados se une al átomo de nitrógeno del indol por una cadena de alquilo de longitud  $n$ , en la que  $n$  tiene 1-5 carbono o heteroátomos. La cadena puede ser  $(CH_2)_n$  en la que  $n$  es 1-5. El grupo puede ser  $(CH_2)_5COOH$ .

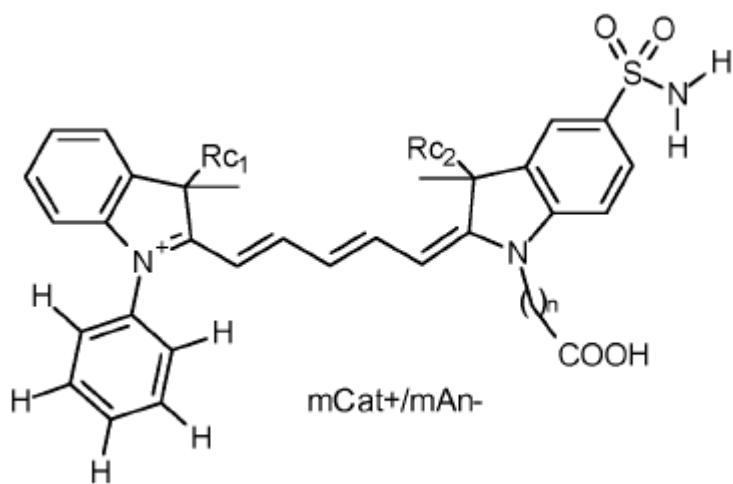
10 Las moléculas pueden contener una o más fracciones de alquil-sulfonato en la posición Rc. Cualquiera de  $R_{C1}$  y/o  $R_{C2}$  puede ser alquil- $SO_3^-$ . El otro Rc ( $R_{C1}$  o  $R_{C2}$ ) puede ser independientemente alquilo o alquilo sustituido.  $R_{C1}$  y  $R_{C2}$  pueden ser metilo, etilo, propilo, butilo, pentilo, hexilo o  $(CH_2)_qSO_3H$ , en el que  $q$  es 1-6.  $q$  puede ser 1-4.  $q$  puede ser 4.  $R_{C1}$  y  $R_{C2}$  puede ser un grupo alquilo sustituido.  $R_{C1}$  y  $R_{C2}$  puede contener una fracción de COOH o - $SO_3H$  o sus derivados de éster o amida.

15 En ciertas realizaciones, cuando uno de  $R_{A1}$  o  $R_{A2}$  es  $SO_3^-$ , y el otro de  $R_{A1}$  o  $R_{A2}$  es H o  $SO_3^-$ , cualquiera de  $R_{C1}$  o  $R_{C2}$  debe ser un grupo ácido alquil sulfónico. Los compuestos en los que uno de  $R_{A1}$  o  $R_{A2}$  es  $SO_3^-$ , y el otro de  $R_{A1}$  o  $R_{A2}$  es H o  $SO_3^-$  y  $R_{C1}$  y  $R_{C2}$  son metilo se rechazan de ciertas reivindicaciones si se requiere. La publicación WO2013041117 divulga ciertos compuestos en los que  $R_{A1}$  es H o  $SO_3^-$ ,  $R_{A2}$  es  $SO_3^-$  y  $R_{C1}$  y  $R_{C2}$  son metilo. Esta publicación no describe moléculas en las que Rc sea un grupo de ácido alquil sulfónico. Esta publicación no divulga

20 moléculas en las que Ra es sulfonamida.

El grupo COOH puede actuar como una fracción de enlace para una unión adicional o se liga a una molécula adicional. Una vez que ha ocurrido la conjugación, el COOH o el COO- se pueden convertir en una amida o un éster.

Ejemplos de los compuestos incluyen estructuras de acuerdo con fórmula (IV):



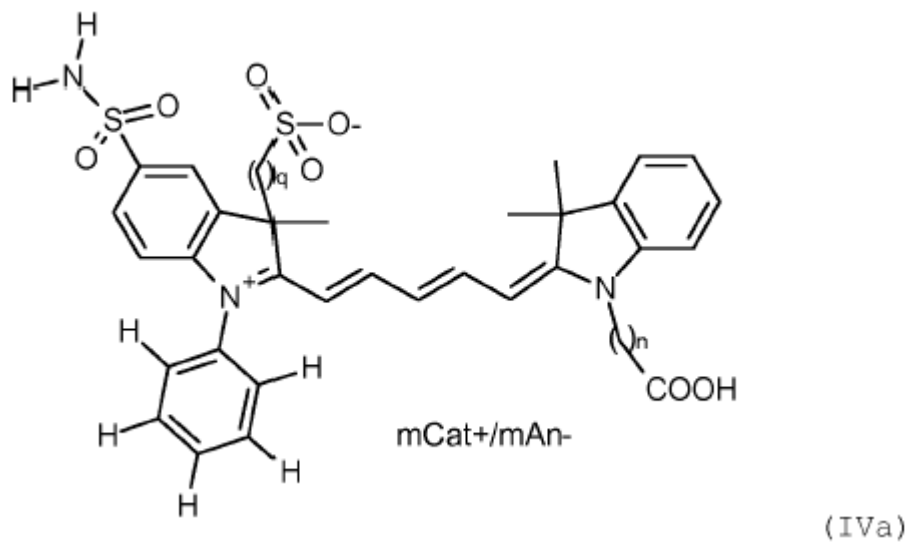
o una sal del mismo en la que mCat<sup>+</sup> o mAn<sup>-</sup> es un contraión orgánico o inorgánico cargado positivamente/negativamente y

m es un entero 0-3;

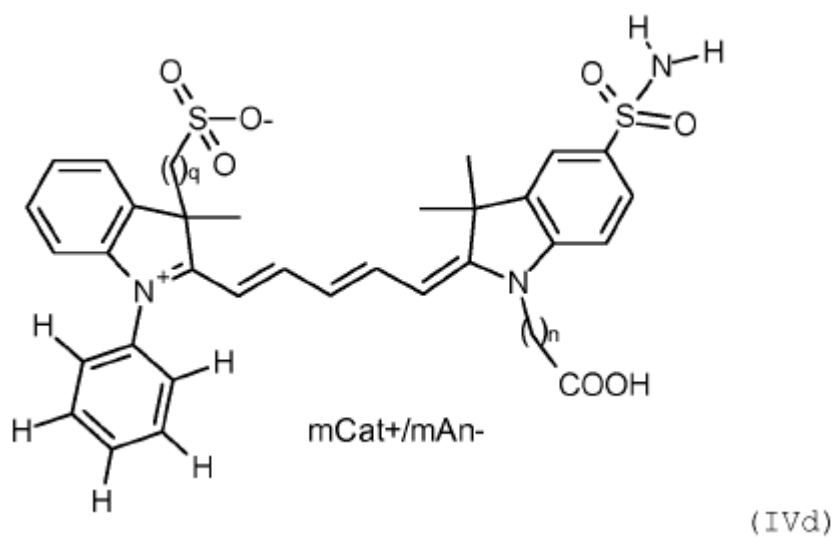
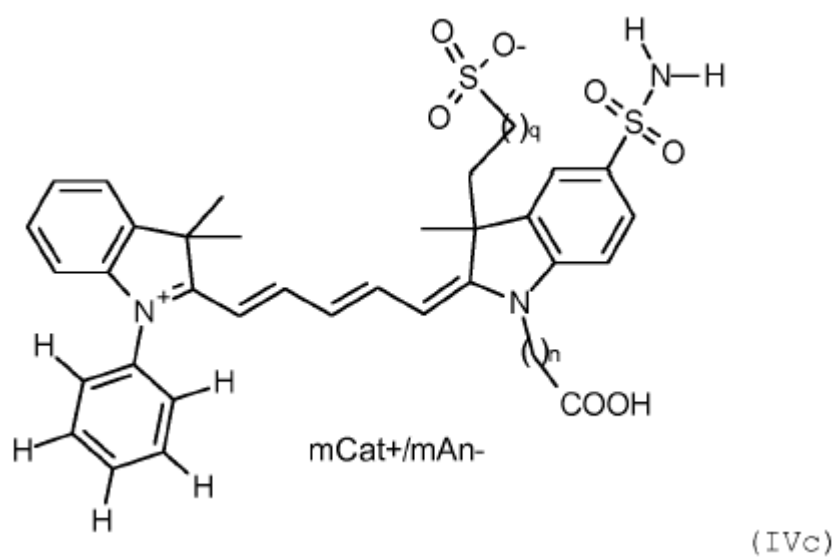
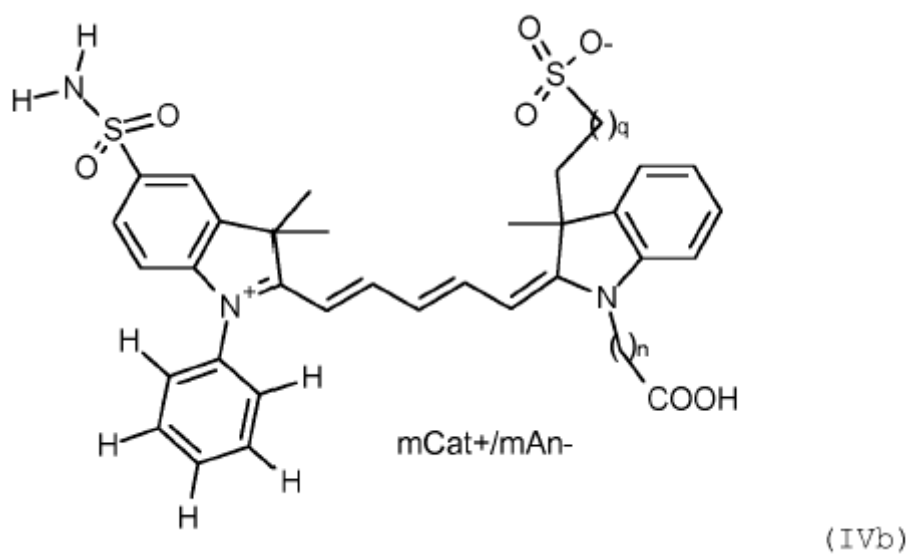
5 n es un entero 1-5; y

cada uno de Rc<sub>1</sub> y Rc<sub>2</sub> es independientemente alquilo o alquilo sustituido.

Ejemplos adicionales de compuestos incluyen estructuras de acuerdo con la fórmula (IVa) a (IVd):







5 o sales de los mismos, en los que mCat<sup>+</sup> o mAn<sup>-</sup> es un contraión orgánico o inorgánico cargado positivamente/negativamente y  
m es un entero 0-3;

n es un entero 1-5; y

q es un entero desde 1-6.

Un compuesto particularmente útil es un nucleótido u oligonucleótido marcado con un tinte como se describe en el presente documento. El nucleótido u oligonucleótido marcado puede tener el marcador unido al átomo de nitrógeno del indol a través de un grupo alquil-carboxi para formar una amida. El nucleótido u oligonucleótido marcado puede tener el marcador unido a la posición C5 de una base de pirimidina o la posición C7 de una base de purina 7-deaza a través de una fracción de ligador.

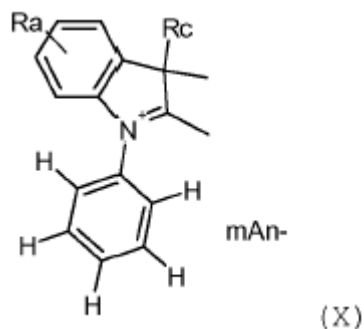
El nucleótido u oligonucleótido marcado también puede tener un grupo bloqueador unido covalentemente al azúcar de ribosa o desoxirribosa del nucleótido. El grupo de bloqueo puede estar unido en cualquier posición sobre el azúcar de ribosa o desoxirribosa. En realizaciones particulares, el grupo de bloqueo está en la posición 3' OH del azúcar de ribosa o desoxirribosa del nucleótido.

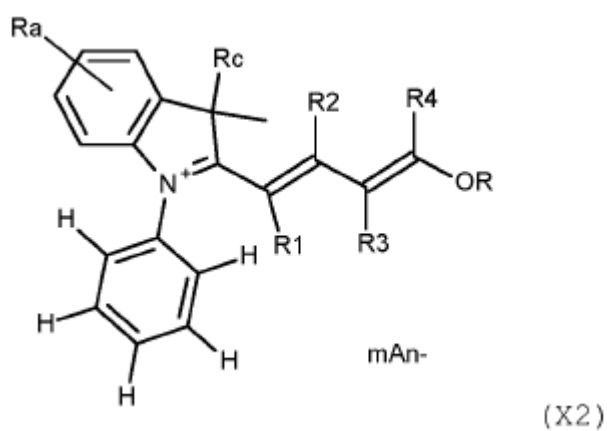
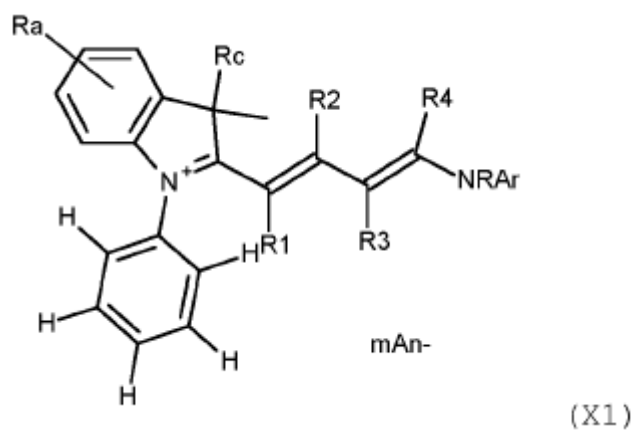
En este documento se proporcionan kits que incluyen dos o más nucleótidos en los que por lo menos un nucleótido es un nucleótido marcado con un compuesto de la presente divulgación. El kit puede incluir dos o más nucleótidos marcados. Los nucleótidos se pueden marcar con dos o más marcadores fluorescentes. Dos o más de los marcadores se pueden excitar utilizando una única fuente de excitación, que puede ser un láser. Por ejemplo, las bandas de excitación para los dos o más marcadores pueden estar solapadas, por lo menos parcialmente, de tal manera que la excitación en la región de solapamiento del espectro provoque que ambos marcadores emitan fluorescencia. En realizaciones particulares, la emisión de las dos o más marcadores se producirá en diferentes regiones del espectro, de tal manera que la presencia de por lo menos uno de los marcadores se puede determinar distinguiendo ópticamente la emisión.

El kit puede contener cuatro nucleótidos marcados, en los que el primero de los cuatro nucleótidos está marcado con un compuesto como se describe en este documento. En dicho kit, los nucleótidos segundo, tercero y cuarto se pueden marcar cada uno con un compuesto que es opcionalmente diferente del marcador en el primer nucleótido y opcionalmente diferente de los marcadores entre sí. Por lo tanto, uno o más de los compuestos pueden tener un máximo de absorbancia y/o de emisión distinto, de tal manera que los compuestos son distinguibles de otros compuestos. Por ejemplo, cada compuesto puede tener un máximo de absorbancia y/o emisión distinto, de tal manera que cada uno de los compuestos se distinga de los otros tres compuestos. Se entenderá que partes del espectro de absorbancia y/o espectro de emisión que no sean los máximos pueden diferir y estas diferencias se pueden aprovechar para distinguir los compuestos. El kit puede ser tal que dos o más de los compuestos tengan un máximo de absorbancia distinto a 600 nm. Los compuestos de la invención normalmente absorben luz en la región por encima de 640 nm.

Los compuestos, nucleótidos o kits que se exponen en el presente documento se pueden utilizar para detectar, medir o identificar un sistema biológico (que incluye, por ejemplo, procesos o componentes de los mismos). Las técnicas de ejemplo que pueden emplear los compuestos, nucleótidos o kits incluyen secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN, ensayo celular (por ejemplo, unión celular o análisis de la función celular), o ensayo de proteínas (por ejemplo, ensayo de unión de proteínas o ensayo de actividad de proteínas). El uso puede ser sobre un instrumento automatizado para llevar a cabo una técnica particular, tal como un instrumento de secuenciación automatizado. El instrumento de secuenciación puede contener dos láseres que operan en diferentes longitudes de onda.

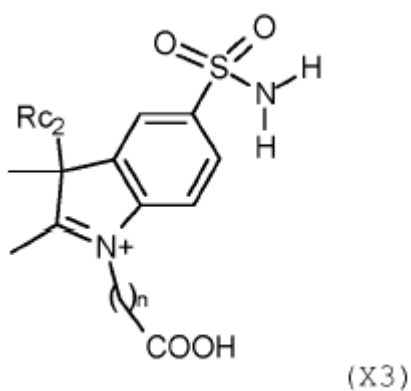
En este documento se divulga un método para sintetizar compuestos de la divulgación. Un compuesto de fórmula (X) y/o (X1), (X2) o una sal del mismo se pueden utilizar como un material de partida para la síntesis de tintes de polimetina simétricos o asimétricos:

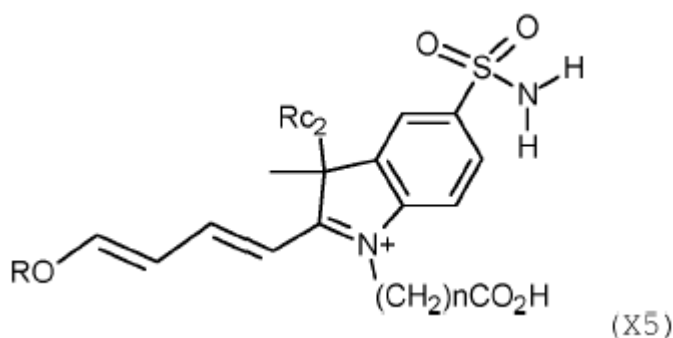
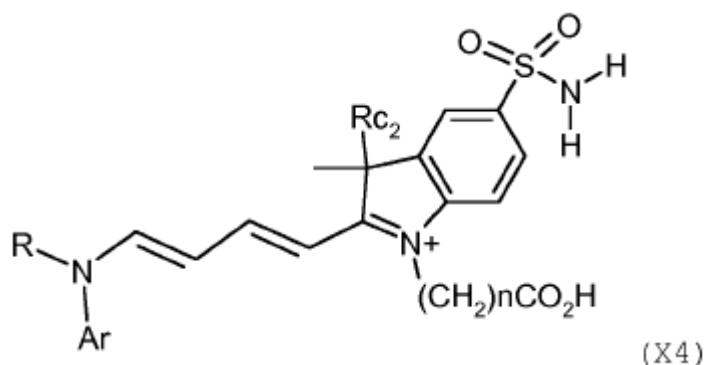




5 o una sal del mismo en la que Ra es H, SO<sub>3</sub><sup>-</sup>, sulfonamida, halógeno, o un anillo adicional fusionado a un átomo de carbono adyacente; R1, R2, R3 y R4 son independientemente H, alquilo o arilo; Rc<sub>2</sub> es alquilo o alquilo sustituido; Ar es un grupo aromático y R es un grupo alquilo.

En este documento se divulga un método para sintetizar compuestos de la divulgación. Un compuesto de fórmula (X3), (X4) o (X5) o una sal del mismo se puede utilizar como un material de partida para la síntesis de tintes de polimetina simétricos o asimétricos:





o una sal del mismo en la que n es un entero 1-5; Rc<sub>2</sub> es alquilo o alquilo sustituido; Ar es un grupo aromático y R es un grupo alquilo.

- 5 Como se utiliza en el presente documento, el término "alquilo" se refiere a hidrocarburo C<sub>1</sub>-C<sub>20</sub> y puede incluir anillos carbocíclicos no aromáticos C<sub>3</sub>-C<sub>10</sub>. En realizaciones particulares, los grupos alquilo son alquilo C<sub>1</sub>-C<sub>6</sub> que se refiere a radicales hidrocarburos saturados de cadena lineal o ramificada que contienen entre uno y seis átomos de carbono, respectivamente. Los grupos alquilo pueden incluir uno o más grupos insaturados, y por lo tanto incluyen alqueno y alquino.
- 10 El término "halógeno", como se utiliza en el presente documento, se refiere a fluoro (en lo sucesivo denominado F), cloro (en lo sucesivo denominado Cl), bromo (en lo sucesivo denominado Br) o yodo (en lo sucesivo denominado I), y usualmente se relaciona con la sustitución de un átomo de hidrógeno en un compuesto orgánico, esta sustitución es opcionalmente una sustitución completa del hidrógeno.
- 15 El término "alquilo sustituido", se refiere a grupos alquilo, alqueno o alquino como se definió anteriormente en los que pueden sustituir opcionalmente adicionalmente con, pero sin limitación, halo, ciano, SO<sub>3</sub><sup>-</sup>, SRa, ORa, NRbRc, oxo, CONRbRc, COOH y COORb. Ra, Rb y Rc se pueden seleccionar cada uno independientemente de H, alquilo, alquilo sustituido, alqueno, alqueno sustituido, alquino, alquino sustituido, arilo y arilo sustituido. Adicionalmente, dicho alquilo sustituido, alqueno sustituido y alquino sustituido puede estar opcionalmente interrumpido por al menos un heteroátomo o grupo seleccionado de O, NRb, S(O)<sub>t</sub> en el que t es 0 a 2, y similares. El alquilo sustituido
- 20 también cubre un grupo tal como bencilo en el que los grupos alquilo comprenden una fracción arilo o arilo sustituido.

Los tintes de acuerdo con la presente divulgación se pueden sintetizar a partir de una variedad de diferentes materiales de partida, que incluyen indoles de N-fenilo. Los tintes se pueden hacer simétricamente, de tal manera que el mismo indol esté en ambos extremos de la cadena de trimetina, o asimétricamente, de tal manera que

25 diferentes indoles estén en cualquier extremo del cromóforo. Los métodos para preparar tintes de polimetina son bien conocidos en la técnica.

De acuerdo con un aspecto de la divulgación, se proporcionan compuestos tintes adecuados para la unión a fracciones de sustrato, que comprenden particularmente grupos ligadores para permitir la unión a fracciones de sustrato. Las fracciones de sustrato pueden ser virtualmente cualquier molécula o sustancia a la que se puedan

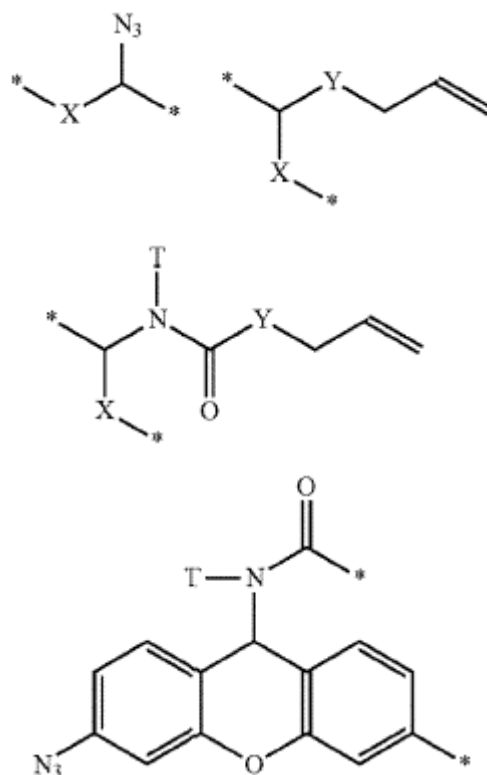
30 conjugar los tintes de la divulgación y, a modo de ejemplo no limitativo, pueden incluir nucleósidos, nucleótidos, polinucleótidos, carbohidratos, ligandos, partículas, superficies sólidas, polímeros orgánicos e inorgánicos, cromosomas, núcleos, células vivas y combinaciones o ensamblajes de los mismos. Los tintes pueden conjugarse mediante un ligador opcional mediante una variedad de medios que incluyen atracción hidrofóbica, atracción iónica y unión covalente. En particular, los tintes se conjugan con el sustrato por unión covalente. Más particularmente, la

35 unión covalente es por medio de un grupo ligador.

La conjugación del compuesto de tinte al sustrato se puede llevar a cabo a través del grupo carboxilo CORf, que puede convertirse en una amida o éster.

Los tintes de acuerdo con la presente divulgación pueden incluir un grupo ligador reactivo en una de las posiciones de los sustituyentes para la unión covalente del tinte a otra molécula. Los grupos de enlace reactivo son fracciones capaces de formar un enlace (por ejemplo, un enlace covalente o no covalente). En una realización particular, el ligador puede ser un ligador de división. El uso del término "ligador de división" no significa que se deba eliminar todo el ligador. El sitio de división se puede ubicar en una posición en el ligador que da como resultado que parte del ligador permanezca unido al tinte y/o a la fracción de sustrato después de la división. Los ligadores de división pueden ser, a modo de ejemplo no limitante, ligadores de división por electrófilos, ligadores de división enzimáticamente, ligadores de división nucleófilos, ligadores que se pueden escindir por división, de división bajo condiciones reductoras (por ejemplo, enlaces que contienen disulfuro o azida), condiciones oxidativas, de división mediante el uso de ligadores de captura de seguridad y de división por mecanismos de eliminación. El uso de un ligador de división para unir el compuesto de tinte a una fracción de sustrato brinda la opción de quitar el marcador, por ejemplo, después de detección, evitando de esta manera cualquier señal de interferencia en las etapas posteriores.

Se pueden encontrar grupos ligadores útiles en la publicación PCT número WO2004/018493, ejemplos de los cuales incluyen ligadores que se pueden dividir utilizando fosfinas solubles en agua o catalizadores de metales de transición solubles en agua formados a partir de un metal de transición y ligandos por lo menos parcialmente solubles en agua. En solución acuosa, estos últimos forman complejos de metales de transición por lo menos parcialmente solubles en agua. Dichos ligadores de división se pueden utilizar para conectar bases de nucleótidos a marcadores tales como los tintes expuestos en este documento. Se pueden encontrar ligadores particulares en la publicación PCT número WO2004/018493, tales como aquellos que incluyen fracciones de la fórmula:



(en la que X se selecciona del grupo que comprende O, S, NH y NQ en la que Q es un grupo alquilo sustituido o no sustituido C1-10, Y se selecciona del grupo que comprende O, S, NH y N(alil), T es hidrógeno o un grupo alquilo sustituido o no sustituido C1-C10 y \* indica en donde se conecta la fracción al resto del nucleótido o nucleósido).

En realizaciones particulares, se puede alterar la longitud del conector entre un tinte fluorescente (fluoróforo) y una base de guanina, por ejemplo, al introducir un grupo separador de polietilenglicol, aumentando de esta manera la intensidad de fluorescencia en comparación con el mismo fluoróforo unido a la base de guanina a través de otros enlaces conocidos en la técnica. Los ligadores de ejemplo y sus propiedades se establecen en la solicitud de patente GB número 0517097.2, publicada como WO07020457. El diseño de los ligadores, y especialmente su mayor longitud, puede permitir mejoras en el brillo de los fluoróforos unidos a las bases de guanina de los nucleótidos de guanosina cuando se incorporan a polinucleótidos tales como el ADN. Por lo tanto, cuando el tinte se utiliza en cualquier método de análisis que emplee la detección de un marcador de tinte fluorescente adherido a un nucleótido

que contiene guanina, puede ser ventajoso utilizar un ligador que tenga un grupo separador de fórmula  $-((CH_2)_2O)_n-$  en el que n es un número entero entre 2 y 50, por ejemplo, como se describe en el documento WO07020457.

La presente divulgación proporciona adicionalmente conjugados de nucleósidos y nucleótidos marcados con uno o más de los tintes establecidos en este documento (nucleótidos modificados). Los nucleósidos y nucleótidos marcados son útiles para marcar polinucleótidos formados por síntesis enzimática, tales como, a modo de ejemplo no limitante, en amplificación por PCR, amplificación isotérmica, amplificación en fase sólida, secuenciación de polinucleótidos (por ejemplo, secuenciación en fase sólida), reacciones de traducción de mella y similares.

Los nucleósidos y nucleótidos pueden marcarse en los sitios en el azúcar o nucleobase. Como se conoce en la técnica, un "nucleótido" consiste en una base nitrogenada, un azúcar y uno o más grupos fosfato. En el ARN, el azúcar es 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 nitrogenada es un derivado de purina o pirimidina. Las purinas pueden ser adenina (A) o guanina (G), y las pirimidinas pueden ser citosina (C), timina (T) o en el contexto de ARN, uracilo (U). El átomo de desoxirribosa C-1 está unido a N-1 de una pirimidina o N-9 de una purina. Un nucleótido es también un éster de fosfato de un nucleósido, con esterificación que ocurre en el grupo hidroxilo unido al C-3 o C-5 del azúcar. Los nucleótidos usualmente son mono, di o trifosfatos.

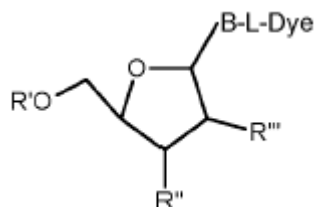
Un "nucleósido" es estructuralmente similar a un nucleótido pero falta las fracciones de fosfato. Un ejemplo de un análogo de nucleósido sería uno en el que el marcador esté unido a la base y no haya un grupo fosfato unido a la molécula de azúcar.

Aunque la base usualmente se denomina purina o pirimidina, el experto apreciará que hay derivados y análogos disponibles que no alteran la capacidad del nucleótido o nucleósido para experimentar el emparejamiento de bases de Watson-Crick. "Derivado" o "análogo" significa un compuesto o molécula cuya estructura central es la misma, o se parece mucho a la de un compuesto original, pero que tiene una modificación química o física, tal como por ejemplo, un grupo lateral diferente o adicional, lo que permite que el nucleótido o nucleósido derivado se una a otra molécula. Por ejemplo, la base puede ser una deazapurina. En realizaciones particulares, los derivados son capaces de experimentar el emparejamiento de Watson-Crick. "Derivado" y "análogo" también incluyen, por ejemplo, un nucleótido sintético o derivado de nucleósido que tiene fracciones de bases modificadas y/o fracciones de azúcar modificadas. Dichos derivados y análogos se analizan, por ejemplo, en Scheit, Nucleotide analogs (John Wiley & Son, 1980) y Uhlman et al., Chemical Reviews 90:543-584, 1990. Los análogos de nucleótidos también pueden tener enlaces fosfodiéster modificados que incluyen fosforotioato, fosforoditioato, alquilfosfonato, fosforanilidato, enlaces de fosforamidato y similares.

Un tinte se puede unir a cualquier posición en una base de nucleótido, por ejemplo, a través de un ligador. En realizaciones particulares, el emparejamiento de bases de Watson-Crick todavía se puede llevar a cabo para el análogo resultante. Los sitios de marcado de nucleobases particulares incluyen la posición C5 de una base de pirimidina o la posición C7 de una base de purina 7-deaza. Como se describió anteriormente, se puede utilizar un grupo ligador para unir covalentemente un tinte al nucleósido o nucleótido.

En realizaciones particulares, el nucleósido o nucleótido marcado puede ser incorporable enzimáticamente y extensible enzimáticamente. De acuerdo con lo anterior, una fracción de ligador puede tener una longitud suficiente para conectar el nucleótido al compuesto de tal manera que el compuesto no interfiera significativamente con la unión y el reconocimiento globales del nucleótido por una enzima de replicación de ácido nucleico. Por lo tanto, el ligador también puede comprender una unidad separadora. Las distancias del separador, por ejemplo, la base de nucleótidos de un sitio de división o marcador.

Los nucleósidos o nucleótidos marcados con tintes de la divulgación pueden tener la fórmula:



en la que el Dye es un compuesto de tinte de acuerdo con la presente divulgación, B es una nucleobase, tal como, por ejemplo, uracilo, timina, citosina, adenina, guanina y similares y L es un grupo ligador opcional que puede o no estar presente.

R' puede ser H, monofosfato, difosfato, trifosfato, tiofosfato, un análogo de éster de fosfato, -O- unido a un grupo que contiene fósforo reactivo o -O- protegido por un grupo bloqueador. R'' puede ser H, OH, una fosforamidita o un grupo bloqueador 3'OH y R''' es H u OH.

Cuando R" es fosforamidita, R' es un grupo protector de hidroxilo de división en ácido que permite el posterior acoplamiento de monómeros bajo condiciones de síntesis automatizadas.

En una realización particular, el grupo de bloqueo es separado e independiente del compuesto de tinte, es decir, no está directamente unido a él. En una realización alternativa, el tinte puede comprender todo o parte del grupo bloqueador 3'OH. Por lo tanto, R" puede ser un grupo bloqueador 3'OH que puede o no comprender un compuesto de tinte descrito en este documento.

En otra realización alternativa más, no hay ningún grupo de bloqueo en el carbono 3' del azúcar pentosa y el tinte (o el tinte y el constructor ligador) unidos a la base, por ejemplo, pueden ser de un tamaño o estructura suficiente para actuar como un bloque para la incorporación de un nucleótido adicional. Por lo tanto, el bloqueo puede deberse a un impedimento estérico o puede deberse a una combinación de tamaño, carga y estructura, ya sea que el tinte esté o no unido a la posición 3' del azúcar.

En otra realización alternativa más, el grupo de bloqueo está presente en el carbono 2' o 4' de la azúcar pentosa y puede ser de un tamaño o estructura suficiente para actuar como un bloque para la incorporación de un nucleótido adicional. El uso de un grupo de bloqueo permite controlar la polimerización, tal como al detener la extensión cuando se incorpora un nucleótido modificado. Si el efecto de bloqueo es reversible, por ejemplo a modo de ejemplo no limitativo al cambiar las condiciones químicas o al eliminar un bloque químico, la extensión se puede detener en ciertos puntos y luego permitir que continúe.

En otra realización particular, un grupo de bloqueo 3'OH comprenderá las fracciones divulgadas en el documento WO2004/018497. Por ejemplo, el grupo de bloqueo puede ser azidometilo ( $\text{CH}_2\text{N}_3$ ) o alilo.

En una realización particular, un ligador (entre el tinte y el nucleótido) y un grupo de bloqueo están presentes y son fracciones separadas. En realizaciones particulares, tanto el ligador como el grupo de bloqueo son de división bajo condiciones sustancialmente similares. Por lo tanto, los procesos de desprotección y desbloqueo pueden ser más eficientes, ya que solo se requerirá un único tratamiento para eliminar tanto el compuesto de tinte como el bloque. Sin embargo, en algunas realizaciones, un ligador y un grupo de bloqueo no necesitan ser de división bajo condiciones similares, en lugar de estar individualmente de división bajo condiciones distintas.

Esta divulgación también abarca polinucleótidos que incorporan compuestos de tintes. Dichos polinucleótidos pueden ser ADN o ARN que comprenden respectivamente desoxirribonucleótidos o ribonucleótidos unidos en el enlace fosfodiéster. Los polinucleótidos de acuerdo con la divulgación pueden comprender nucleótidos naturales, nucleótidos no naturales (o modificados) distintos de los nucleótidos modificados de la divulgación o cualquier combinación de los mismos, en combinación con por lo menos un conjunto de nucleótidos modificado (por ejemplo, marcado con un compuesto de tinte) establecido en este documento. Los polinucleótidos de acuerdo con la divulgación también pueden incluir enlaces de estructura principal no naturales y/o modificaciones químicas no nucleotídicas. También se contemplan estructuras químicas que comprenden mezclas de ribonucleótidos y desoxirribonucleótidos que comprenden por lo menos un nucleótido modificado de acuerdo con la divulgación.

Los nucleótidos modificados (o nucleósidos) que comprenden un compuesto de tinte de acuerdo con la presente divulgación se pueden utilizar en cualquier método de análisis, tales como los métodos que incluyen la detección de un marcador fluorescente unido a un nucleótido o nucleósido, ya sea por sí solo o incorporado a o asociado con una mayor estructura molecular o conjugado. En este contexto, el término "incorporado en un polinucleótido" puede significar que el fosfato 5' está unido en el enlace fosfodiéster al grupo hidroxilo 3' de un segundo nucleótido (modificado o no modificado), que puede formar parte de una cadena polinucleotídica más larga. El extremo 3' de un nucleótido modificado establecido en este documento puede o no unirse en el enlace fosfodiéster al fosfato 5' de un nucleótido adicional (modificado o no modificado). Por lo tanto, en una realización no limitativa, la divulgación proporciona un método para detectar un nucleótido modificado incorporado en un polinucleótido que comprende: (a) incorporar por lo menos un nucleótido modificado de la divulgación en un polinucleótido y (b) detectar los nucleótidos modificados incorporados en el polinucleótido al detectar la señal fluorescente del compuesto de tinte unido a dicho nucleótido modificado.

Este método puede incluir: una etapa sintética (a) en la que uno o más nucleótidos modificados de acuerdo con la divulgación se incorporan en un polinucleótido y una etapa de detección (b) en la que se detectan uno o más nucleótidos modificados incorporados en los polinucleótidos al detectar o medir cuantitativamente su fluorescencia.

En una realización de la presente divulgación, por lo menos un nucleótido modificado se incorpora a un polinucleótido en una etapa sintética por la acción de una enzima polimerasa. Sin embargo, se pueden utilizar otros métodos para unir nucleótidos modificados a polinucleótidos, tales como por ejemplo la síntesis química de oligonucleótidos o la unión de oligonucleótidos marcados a oligonucleótidos no marcados. Por lo tanto, el término "incorporar", cuando se utiliza en referencia a un nucleótido y polinucleótido, puede abarcar la síntesis de polinucleótidos por métodos químicos, de esta manera como métodos enzimáticos.

En una realización específica, se lleva a cabo una etapa sintética y puede comprender opcionalmente incubar una cadena de polinucleótido de plantilla con una mezcla de reacción que comprende nucleótidos modificados marcados de forma fluorescente de la divulgación. También se puede proporcionar una polimerasa bajo condiciones que

permitan la formación de un enlace fosfodiéster entre un grupo hidroxilo 3' libre en una cadena de polinucleótido híbrida a la cadena de polinucleótido de plantilla y un grupo fosfato 5' en el nucleótido modificado. Por lo tanto, una etapa sintética puede incluir la formación de una cadena de polinucleótidos como se indica mediante el emparejamiento de bases complementarias de nucleótidos a una cadena de plantilla.

- 5 En todas las realizaciones del método, la etapa de detección se puede llevar a cabo mientras que la cadena de polinucleótidos en la que se incorporan los nucleótidos modificados se hibrida a una cadena de plantilla, o después de una etapa de desnaturalización en la que se separan las dos cadenas. Etapas adicionales, por ejemplo las etapas de reacción química o enzimática o etapas de purificación, se pueden incluir entre una etapa sintética y una etapa de detección. En particular, la cadena objetivo que incorpora el nucleótido modificado se puede aislar o purificar y luego procesarse o utilizarse en un análisis posterior. A modo de ejemplo, los polinucleótidos objetivo marcados con nucleótidos modificados en una etapa sintética se pueden utilizar posteriormente como sondas o cebadores marcados. En otras realizaciones, el producto de una etapa sintética establecida en el presente documento se puede someter a etapas de reacción adicionales y, si se desea, el producto de estas etapas posteriores se puede purificar o aislar.
- 10 Las condiciones adecuadas para una etapa sintética serán bien conocidas por aquellos familiarizados con las técnicas estándar de biología molecular. En una realización, una etapa sintética puede ser análoga a una reacción de extensión de cebador estándar utilizando precursores de nucleótidos, que incluyen los nucleótidos modificados establecidos en este documento, para formar una cadena objetivo extendida complementaria de la cadena de plantilla en presencia de una enzima polimerasa adecuada. En otras realizaciones, una etapa sintética puede formar parte de una reacción de amplificación que produce un producto de amplificación de cadena doble marcado que comprende cadenas complementarias híbridas derivadas de la copia de las cadenas de polinucleótidos objetivo y de plantilla. Otras etapas sintéticas de ejemplo incluyen la traducción de mella, polimerización por desplazamiento de cadena, marcado de ADN cebado al azar, etc. Una enzima polimerasa particularmente útil para una etapa sintética es una que es capaz de catalizar la incorporación de uno o más de los nucleótidos modificados establecidos en el presente documento. Se puede utilizar una variedad de polimerasas naturales o modificadas. A modo de ejemplo, se puede utilizar una polimerasa termoestable para una reacción sintética que se lleva a cabo utilizando condiciones de termociclado, mientras que una polimerasa termoestable puede no ser deseada para reacciones de extensión de cebadores isotérmicos. Las polimerasas termoestables adecuadas que son capaces de incorporar los nucleótidos modificados de acuerdo con la divulgación incluyen aquellas descritas en los documentos WO 2005/024010 o WO06120433. En reacciones sintéticas que se llevan a cabo a temperaturas más bajas, tales como 37°C, las enzimas polimerasas no necesitan ser polimerasas termoestables, por lo tanto, la elección de la polimerasa dependerá de una serie de factores tales como la temperatura de reacción, el pH, la actividad de desplazamiento de la cadena y similares.
- 15 En realizaciones no limitantes específicas, la divulgación abarca métodos de secuenciación de ácidos nucleicos, resecuenciación, secuenciación del genoma completo, puntuación de polimorfismo de nucleótido único, o cualquier otra aplicación que implique la detección del nucleótido modificado o nucleósido marcado con tintes establecidos en este documento cuando se incorporan en un polinucleótido. Cualquiera de una variedad de otras aplicaciones que se benefician del uso de polinucleótidos marcados con los nucleótidos modificados que comprenden tintes fluorescentes puede utilizar nucleótidos modificados o nucleósidos marcados con tintes establecidos en el presente documento.
- 20 En una realización particular, la divulgación proporciona el uso de nucleótidos modificados que comprenden compuestos de tintes de acuerdo con la divulgación en una reacción de secuenciación por síntesis de polinucleótidos. La secuenciación por síntesis generalmente implica la adición secuencial de uno o más nucleótidos u oligonucleótidos a una cadena de polinucleótidos en crecimiento en la dirección 5' a 3' utilizando una polimerasa o ligasa para formar una cadena de polinucleótidos extendida complementaria al ácido nucleico de la plantilla que se va a secuenciar. La identidad de la base presente en uno o más de los nucleótidos agregados se puede determinar en una etapa de detección o "formación de imágenes". La identidad de la base agregada se puede determinar después de cada etapa de incorporación de nucleótidos. La secuencia de la plantilla se puede inferir utilizando las reglas convencionales de emparejamiento de bases de Watson-Crick. El uso de los nucleótidos modificados marcados con tintes en este documento establecidos para la determinación de la identidad de una sola base puede ser útil, por ejemplo, en la calificación de polimorfismos de un solo nucleótido, y dichas reacciones de extensión de una sola base están dentro del alcance de esta divulgación.
- 25 En una realización de la presente divulgación, la secuencia de un polinucleótido de plantilla se determina al detectar la incorporación de uno o más nucleótidos en una cadena naciente complementaria al polinucleótido de plantilla que se va a secuenciar a través de la detección de marcadores fluorescentes unidos al nucleótido incorporado. La secuenciación del polinucleótido plantilla se puede cebar con un cebador adecuado (o se puede preparar como una construcción de horquilla que contendrá el cebador como parte de la horquilla), y la cadena naciente se extiende de manera gradual mediante la adición de nucleótidos al extremo 3' del cebador en una reacción catalizada por polimerasa.
- 30 En realizaciones particulares, cada uno de los diferentes trifosfatos de nucleótidos (A, T, G y C) se puede marcar con un fluoróforo único y también comprende un grupo bloqueador en la posición 3' para prevenir la polimerización no



controlada. Alternativamente, uno de los cuatro nucleótidos puede estar sin marcar (oscuro). La enzima polimerasa incorpora un nucleótido en la cadena naciente complementaria al polinucleótido de plantilla, y el grupo de bloqueo evita la incorporación adicional de nucleótidos. Cualquier nucleótido no incorporado se puede lavar y la señal fluorescente de cada nucleótido incorporado se puede "leer" ópticamente por medios adecuados, tales como un dispositivo de carga acoplada que utiliza excitación láser y filtros de emisión adecuados. El grupo bloqueador 3' y los compuestos de tinte fluorescentes luego se pueden eliminar (desproteger), (simultánea o secuencialmente) para establecer la cadena naciente para la incorporación de nucleótidos adicionales. Normalmente, la identidad del nucleótido incorporado se determinará después de cada etapa de incorporación, pero esto no es estrictamente esencial. Del mismo modo, la patente de Estados Unidos 5,302,509 divulga un método para secuenciar polinucleótidos inmovilizados sobre un soporte sólido.

El método, como se ejemplificó anteriormente, utiliza la incorporación de nucleótidos bloqueados en 3', A, G, C y T marcados con fluorescencia en una cadena en crecimiento complementaria al polinucleótido inmovilizado, en presencia de ADN polimerasa. La polimerasa incorpora una base complementaria al polinucleótido objetivo, pero se evita que se agregue más por el grupo bloqueador 3'. El marcador del nucleótido incorporado luego se puede determinar y el grupo de bloqueo puede eliminarse por división química para permitir que se produzca una polimerización adicional. La plantilla de ácido nucleico que debe secuenciarse en una reacción de síntesis por secuenciación puede ser cualquier polinucleótido que se desee secuenciar. La plantilla de ácido nucleico para una reacción de secuenciación normalmente comprenderá una región de doble cadena que tiene un grupo hidroxilo 3' libre que sirve como cebador o punto de inicio para la adición de nucleótidos adicionales en la reacción de secuenciación. La región de la plantilla que se secuenciará sobresale de este grupo hidroxilo 3' libre en la cadena complementaria. La región que sobresale de la plantilla a secuenciar puede ser de cadena sencilla pero puede ser de doble cadena, siempre que haya una "mella" presente sobre la cadena complementaria de la cadena de plantilla que se va a secuenciar para proporcionar un grupo 3'OH libre para el inicio de la reacción de secuenciación. En dichas realizaciones, la secuenciación puede proceder por desplazamiento de cadena. En ciertas realizaciones, un cebador que contiene el grupo hidroxilo 3' libre se puede agregar como un componente separado (por ejemplo, un oligonucleótido corto) que se hibrida a una región de cadena única de la plantilla que se va a secuenciar. Alternativamente, el cebador y la cadena de plantilla que se va a secuenciar pueden formar cada uno parte de una cadena de ácido nucleico parcialmente auto-complementaria capaz de formar un dúplex intramolecular, tal como, por ejemplo, una estructura de bucle de horquilla. Los polinucleótidos de horquilla y los métodos por los cuales se pueden unir a los soportes sólidos se divulgan en las publicaciones de solicitudes internacionales Nos. WO0157248 y WO2005/047301. Los nucleótidos se pueden agregar sucesivamente a un cebador en crecimiento, dando como resultado la síntesis de una cadena de polinucleótido en la dirección 5' a 3'. La naturaleza de la base que se ha agregado se puede determinar, en particular, pero no necesariamente después de cada adición de nucleótidos, proporcionando de esta manera información de secuencia para la plantilla de ácido nucleico. Por lo tanto, un nucleótido se incorpora a una cadena de ácido nucleico (o polinucleótido) al unir el nucleótido al grupo hidroxilo 3' libre de la cadena de ácido nucleico a través de la formación de un enlace fosfodiéster con el grupo fosfato 5' del nucleótido.

La plantilla de ácido nucleico que se va que se va a secuenciar puede ser ADN o ARN, o incluso una molécula híbrida compuesta de desoxinucleótidos y ribonucleótidos. La plantilla de ácido nucleico puede comprender nucleótidos de origen natural y/o no natural y enlaces de estructura principal naturales o no naturales, siempre que estos no eviten la copia de la plantilla en la reacción de secuenciación.

En ciertas realizaciones, la plantilla de ácido nucleico que se va a secuenciar se puede unir a un soporte sólido mediante cualquier método de enlace adecuado conocido en la técnica, por ejemplo, mediante unión covalente. En ciertas realizaciones, los polinucleótidos de plantilla se pueden unir directamente a un soporte sólido (por ejemplo, un soporte a base de sílice). Sin embargo, en otras realizaciones de la divulgación, la superficie del soporte sólido se puede modificar de alguna manera para permitir la unión covalente directa de polinucleótidos de plantilla, o inmovilizar los polinucleótidos de plantilla a través de un hidrogel o multicapa de polielectrolito, que se puede unir de forma no covalente al soporte sólido.

Las disposiciones en las que los polinucleótidos se han unido directamente a soportes basados en sílice son aquellas que se divulgan, por ejemplo, en el documento WO00006770, en el que los polinucleótidos se inmovilizan en un soporte de vidrio por reacción entre un grupo epóxido colgante sobre el vidrio con un grupo amino interno sobre el polinucleótido. Adicionalmente, los polinucleótidos se pueden unir a un soporte sólido mediante la reacción de un nucleófilo a base de azufre con el soporte sólido, por ejemplo, como se describe en el documento WO2005/047301. Aún un ejemplo adicional de polinucleótidos de plantilla soportada en sólidos es en la que los polinucleótidos de plantilla están unidos a hidrogel soportado sobre soportes basados en sílice u otros sólidos, por ejemplo, como se describe en los documentos WO00/31148, WO01/01143, WO02/12566, WO03/014392. Patente de Estados Unidos No. 6,465,178 y el documento WO00/53812.

Una superficie particular en la que se pueden inmovilizar los polinucleótidos de la plantilla es un hidrogel de poli(acrilamida). Los hidrogeles de poli(acrilamida) se describen en las referencias citadas anteriormente y en el documento WO2005/065814.

Las moléculas de la plantilla de ADN se pueden unir a perlas o micropartículas, por ejemplo, como se describe en la Patente de Estados Unidos No. 6,172,218. La unión a perlas o micropartículas puede ser útil para aplicaciones de secuenciación. Las colecciones de perlas se pueden preparar en las que cada perla contiene diferentes secuencias de ADN. Colecciones y métodos de ejemplo para su creación se describen en Nature. 437, 376-380 (2005); Science. 309, 5741, 1728-1732 (2005). La secuenciación de disposiciones de dichas perlas utilizando nucleótidos establecidos en el presente documento está dentro del alcance de la divulgación.

Las plantillas que se deben secuenciar pueden formar parte de una "disposición" sobre un soporte sólido, en cuyo caso la disposición puede tomar cualquier forma conveniente. Por lo tanto, el método de la divulgación es aplicable a todos los tipos de disposiciones de alta densidad, que incluyen disposiciones de molécula única, disposiciones agrupadas y disposiciones de perlas. Los nucleótidos modificados marcados con compuestos de tintes de la presente divulgación se pueden utilizar para secuenciar las plantillas en esencialmente cualquier tipo de disposición, que incluyen, pero no se limitan a, aquellas formadas por inmovilización de moléculas de ácido nucleico sobre un soporte sólido.

Sin embargo, los nucleótidos modificados marcados con compuestos de tintes de la divulgación son particularmente ventajosos en el contexto de la secuenciación de disposiciones agrupadas. En disposiciones agrupadas, las distintas regiones en la disposición (a menudo referidas como sitios o características) comprenden múltiples moléculas de plantilla de polinucleótidos. En general, las múltiples moléculas polinucleotídicas no se pueden resolver individualmente por medios ópticos y, en cambio, se detectan como un conjunto. Dependiendo de cómo se forme la disposición, cada sitio sobre la disposición puede comprender múltiples copias de una molécula de polinucleótido individual (por ejemplo, el sitio es homogéneo para una especie particular de ácido nucleico de cadena simple o doble) o incluso múltiples copias de un pequeño número de diferentes moléculas de polinucleótidos (por ejemplo, copias múltiples de dos especies diferentes de ácidos nucleicos). Las disposiciones agrupadas de las moléculas de ácido nucleico se pueden producir utilizando técnicas generalmente conocidas en el arte. A modo de ejemplo, los documentos WO 98/44151 y WO00/18957, cada uno de los cuales se incorpora en el presente documento, describen métodos de amplificación de ácidos nucleicos en los que tanto la plantilla como los productos de amplificación permanecen inmovilizados sobre un soporte sólido para formar disposiciones formadas por agrupaciones o "colonias" de moléculas de ácido nucleico inmovilizadas. Las moléculas de ácido nucleico presentes en las disposiciones agrupadas preparadas de acuerdo con estos métodos son plantillas adecuadas para la secuenciación utilizando los nucleótidos modificados marcados con compuestos de tintes de la divulgación.

Los nucleótidos modificados marcados con compuestos de tintes de la presente divulgación también son útiles en la secuenciación de plantillas en disposiciones de una sola molécula. El término "disposición de una sola molécula" o "SMA", como se utiliza en el presente documento, se refiere a una población de moléculas polinucleotídicas, distribuidas (o dispuestas) sobre un soporte sólido, en el que el espaciado de cualquier polinucleótido individual de todos los demás de la población es tal que es posible resolver individualmente las moléculas de polinucleótido individuales. Las moléculas de ácido nucleico objetivo inmovilizadas sobre la superficie del soporte sólido pueden de esta manera ser capaces de resolverse por medios ópticos en algunas realizaciones. Esto significa que una o más señales distintas, cada una representando un polinucleótido, ocurrirán dentro del área resoluble del dispositivo de formación de imágenes en particular utilizado.

En razón a que la detección de una sola molécula se puede lograr cuando el espaciado entre las moléculas de polinucleótidos adyacentes en una disposición es por lo menos 100 nm, más particularmente por lo menos 250 nm, aún más particularmente por lo menos 300 nm, incluso más particularmente por lo menos 350 nm. Por lo tanto, cada molécula se puede resolver y detectar individualmente como un solo punto de fluorescencia de molécula, y la fluorescencia de dicho solo punto de fluorescencia de molécula también presenta una sola etapa de fotodecoloración.

Los términos "resuelto individualmente" y "resolución individual" se utilizan en este documento para especificar que, cuando se visualiza, es posible distinguir una molécula en la disposición de sus moléculas vecinas. La separación entre moléculas individuales sobre la disposición se determinará, en parte, por la técnica particular utilizada para resolver las moléculas individuales. Las características generales de las disposiciones de una sola molécula se entenderán por referencia a las solicitudes publicadas WO00/06770 y WO 01/57248. Aunque un uso de los nucleótidos modificados de la divulgación está en las reacciones de secuenciación por síntesis, la utilidad de los nucleótidos modificados no está limitada a dichos métodos. De hecho, los nucleótidos se pueden utilizar ventajosamente en cualquier metodología de secuenciación que requiera la detección de marcadores fluorescentes unidos a nucleótidos incorporados en un polinucleótido.

En particular, los nucleótidos modificados marcados con compuestos de tintes de la divulgación se pueden utilizar en protocolos automatizados de secuenciación fluorescente, particularmente secuenciación del ciclo terminador de tinte fluorescente basada en el método de secuenciación de terminación de cadena de Sanger y colaboradores. Dichos métodos generalmente utilizan enzimas y ciclos de secuenciación para incorporar dideoxinucleótidos marcados con fluorescencia en una reacción de secuenciación de la extensión del cebador. Los llamados métodos de secuenciación de Sanger, y los protocolos relacionados (tipo Sanger), utilizan una terminación de cadena aleatoria con dideoxinucleótidos marcados.

Por lo tanto, la presente descripción también abarca nucleótidos modificados marcados con compuestos de tintes que son didesoxinucleótidos que carecen de grupos hidroxilo en las posiciones 3' y 2', siendo adecuados dichos didesoxinucleótidos modificados para uso en métodos de secuenciación de tipo Sanger y similares.

Los nucleótidos modificados marcados con compuestos de tintes de la presente divulgación que incorporan grupos bloqueadores 3', se reconocerá, también pueden ser de utilidad en los métodos de Sanger y protocolos relacionados, ya que el mismo efecto logrado utilizando dideoxi nucleótidos modificados se puede lograr al utilizar nucleótidos modificados que tienen grupos bloqueadores 3'-OH: ambos impiden la incorporación de nucleótidos posteriores. Cuando los nucleótidos de acuerdo con la presente divulgación, y que tienen un grupo bloqueador 3' se utilicen en los métodos de secuenciación de tipo Sanger, se apreciará que los compuestos de tinte o marcadores detectables unidos a los nucleótidos no necesitan estar conectados a través de ligadores de división, ya que en cada instancia en la que se incorpora un nucleótido marcado de la divulgación; no es necesario incorporar posteriormente los nucleótidos y, por lo tanto, no es necesario eliminar el marcador del nucleótido.

La presente divulgación también proporciona kits que incluyen nucleósidos modificados y/o nucleótidos marcados con tintes. Dichos kits generalmente incluirán por lo menos un nucleótido modificado o nucleósido marcado con un tinte establecido en este documento junto con por lo menos un componente adicional. El o los componentes adicionales pueden ser uno o más de los componentes identificados en un método establecido anteriormente o en la sección de ejemplos a continuación. A continuación se establecen algunos ejemplos no limitantes de componentes que se pueden combinar en un kit de la presente divulgación.

En una realización particular, un kit puede incluir por lo menos un nucleótido modificado o nucleósido marcado con un tinte establecido en este documento junto con nucleótidos o nucleósidos modificados o no modificados. Por ejemplo, los nucleótidos modificados marcados con tintes de acuerdo con la divulgación se pueden suministrar en combinación con nucleótidos nativos o no marcados, y/o con nucleótidos marcados con fluorescencia o cualquier combinación de los mismos. De acuerdo con lo anterior, los kits pueden comprender nucleótidos modificados marcados con tintes de acuerdo con la divulgación y nucleótidos modificados marcados con otros, por ejemplo, compuestos de tintes de la técnica anterior. Las combinaciones de nucleótidos se pueden proporcionar como componentes individuales separados (por ejemplo, un tipo de nucleótido por recipiente o tubo) o como mezclas de nucleótidos (por ejemplo, dos o más nucleótidos mezclados en el mismo recipiente o tubo).

Cuando los kits comprenden una pluralidad, particularmente dos, más particularmente cuatro, nucleótidos modificados marcados con un compuesto de tinte, los diferentes nucleótidos pueden marcarse con diferentes compuestos de tintes, o uno puede ser oscuro, sin compuestos de tintes. Cuando los diferentes nucleótidos están marcados con diferentes compuestos de tintes, una de las características de los kits es que dichos compuestos de tintes son tintes fluorescentes distinguibles espectralmente. Como se utiliza en este documento, el término "tintes fluorescentes distinguibles espectralmente" se refiere a tintes fluorescentes que emiten energía fluorescente en longitudes de onda que se pueden distinguir por un equipo de detección fluorescente (por ejemplo, una plataforma de secuenciación de ADN comercial basada en capilares) cuando dos o más de dichos tintes están presentes en una muestra. Cuando dos nucleótidos modificados marcados con compuestos de tintes fluorescentes se suministran en forma de kit, es una característica de algunas realizaciones que los tintes fluorescentes distinguibles espectralmente se puedan excitar a la misma longitud de onda, como por ejemplo con el mismo láser. Cuando cuatro nucleótidos modificados marcados con compuestos de tintes fluorescentes se suministran en forma de kit, es una característica de algunas realizaciones que dos de los tintes fluorescentes distinguibles espectralmente pueden ser excitados en una longitud de onda y los otros dos tintes distinguibles espectralmente pueden ser excitados en otra longitud de onda. Las longitudes de onda de excitación particulares son 532 nm, 630 nm a 700 nm, particularmente 660 nm.

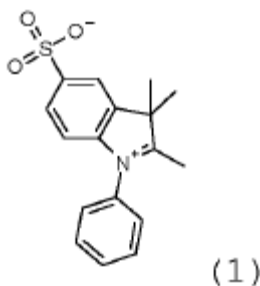
En una realización, un kit incluye un nucleótido modificado marcado con un compuesto de la presente divulgación y un segundo nucleótido modificado marcado con un segundo tinte en el que los tintes tienen una diferencia de absorbancia máxima de por lo menos 10 nm, particularmente de 20 nm a 50 nm. Más particularmente, los dos compuestos de tinte tienen cambios de Stokes de entre 15 y 40 nm, donde el "cambio de Stokes" es la distancia entre la absorción máxima y las longitudes de onda de emisión máxima.

En una realización adicional, un kit puede incluir adicionalmente otros dos nucleótidos modificados marcados con tintes fluorescentes en los que los tintes son excitados por el mismo láser a 488 nm a 550 nm, particularmente a 532 nm. Los tintes pueden tener una diferencia de absorbancia máxima de por lo menos 10 nm, particularmente de 20 nm a 50 nm. Más particularmente, los dos compuestos de tinte pueden tener cambios de Stokes de entre 20 y 40 nm. Aún más particularmente, los dos compuestos de tintes pueden tener un máximo de absorbancia diferente por debajo de 640 nm, particularmente por debajo de 600 nm. Los tintes particulares que se distinguen espectralmente de los tintes de polimetina de la presente divulgación y que cumplen los criterios anteriores son análogos de polimetina como se describe en la Patente de Estados Unidos No. 5,268,486 (por ejemplo, Cy3) o WO 0226891 (Alexa 532; Molecular Probes A20106) o polimetinas asimétricas como se divulga en la Patente de Estados Unidos No. 6,924,372. Los tintes alternativos incluyen análogos de rodamina, por ejemplo, tetrametil-rodamina y análogos de la misma.

- En una realización alternativa, los kits de la divulgación pueden contener nucleótidos en los que la misma base está marcada con dos compuestos diferentes. Un primer nucleótido se puede marcar con un compuesto de la divulgación. Un segundo nucleótido se puede marcar con un compuesto espectralmente distinto, por ejemplo, un tinte "verde" que absorbe a menos de 600 nm. Un tercer nucleótido se puede marcar como una mezcla del compuesto de la divulgación y el compuesto espectralmente distinto, y el cuarto nucleótido puede ser "oscuro" y no contener marcador. En términos simples, por lo tanto, los nucleótidos 1-4 se pueden marcar como "verde", "rojo", "rojo/verde" y oscuro. Para simplificar aún más la instrumentación, se pueden marcar cuatro nucleótidos con dos tintes excitados con un solo láser, y por lo tanto el etiquetado de los nucleótidos 1-4 puede ser 'rojo 1', 'rojo 2' 'rojo 1/rojo 2', y oscuro.
- Los nucleótidos pueden contener dos tintes de la presente divulgación. Los tintes en los que  $Ra_1$  o  $Ra_2$  son un anillo aromático adicional fusionado a los carbonos adyacentes del anillo de indol absorben a una longitud de onda más larga que cuando los tintes no tienen la conjugación aromática adicional. Un kit puede contener dos o más nucleótidos marcados con tintes de la divulgación. Un kit puede contener un nucleótido marcado con un compuesto de la divulgación, en el que cada uno de  $Ra_1$  y  $Ra_2$  es independientemente H,  $SO_3^-$ , sulfonamida o halógeno, y un nucleótido marcado con un compuesto de la divulgación en el que uno o ambos  $Ra_1$  y  $Ra_2$  son un anillo fusionado a un átomo de carbono adyacente. Los kits pueden contener un nucleótido adicional en el que el nucleótido está marcado con un tinte que absorbe en la región de 520 nm a 560 nm. Los kits pueden contener además un nucleótido sin marcar.
- Aunque los kits se ejemplificaron anteriormente con respecto a las configuraciones que tienen diferentes nucleótidos que están marcados con diferentes compuestos de tinte, se entenderá que los kits pueden incluir 2, 3, 4 o más nucleótidos diferentes que tienen el mismo compuesto de tinte.
- En realizaciones particulares, un kit puede incluir una enzima polimerasa capaz de catalizar la incorporación de los nucleótidos modificados en un polinucleótido. Otros componentes que se incluirán en dichos kits pueden incluir tampones y similares. Los nucleótidos modificados marcados con tintes de acuerdo con la divulgación, y otros componentes de cualquier nucleótido que incluyen mezclas de diferentes nucleótidos, se pueden proporcionar en el kit en una forma concentrada para ser diluidos antes de uso. En dichas realizaciones, también se puede incluir un tampón de dilución adecuado. De nuevo, uno o más de los componentes identificados en un método establecido en el presente documento pueden incluirse en un kit de la presente divulgación.
- Se observa que, tal como se utiliza en esta especificación y en las reivindicaciones adjuntas, las formas singulares "un", "una" y "el" incluyen referentes plurales, a menos que estén limitados de forma expresa e inequívoca a un referente.

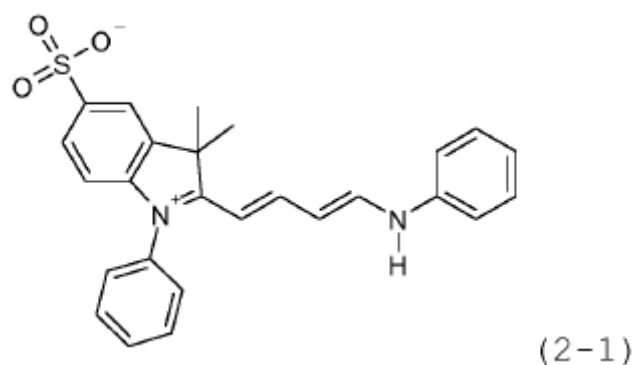
#### Detalles Experimentales

##### 2,3,3-Trimetil-1-fenil-3H-indolio-5-sulfonato (1)

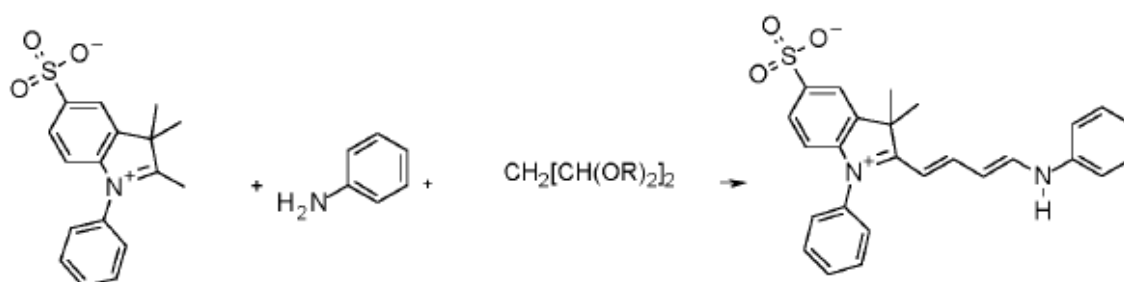


- 2-Metileno-3,3-trimetil-1-fenil-2,3-dihidro-1H-indol (1 g, 4.25 mmol) se disolvió en 1 ml de ácido sulfúrico a temperatura < 5°C y 1 ml de ácido sulfúrico fumante (20 %) se agregó con agitación. La solución se agitó a temperatura ambiente 1 h luego se calienta a 60°C durante 3 h. El producto precipitado con éter de dietilo lavado con acetona y etanol. Rendimiento 0.7 g (52 %). La estructura se confirmó mediante RMN.

##### 2-(4-Anilino-1,3-butadienil)-3,3-trimetil-1-fenil-3H-indolio-5-sulfonato (2-1)

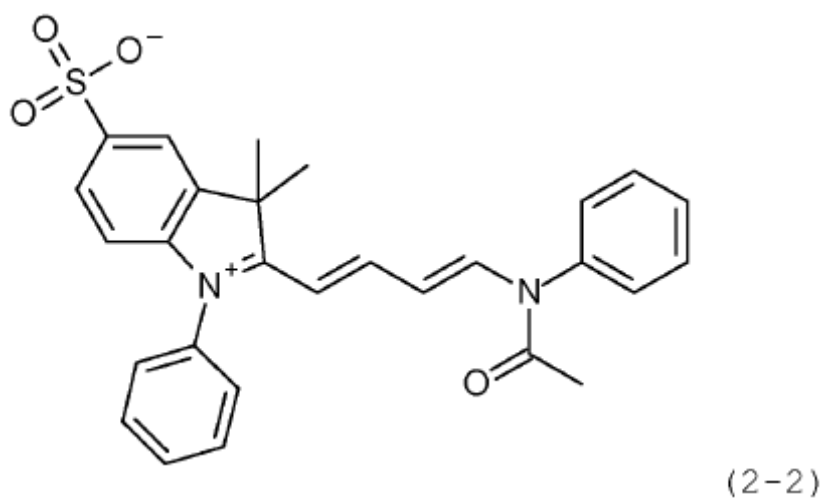


Esquema de reacción

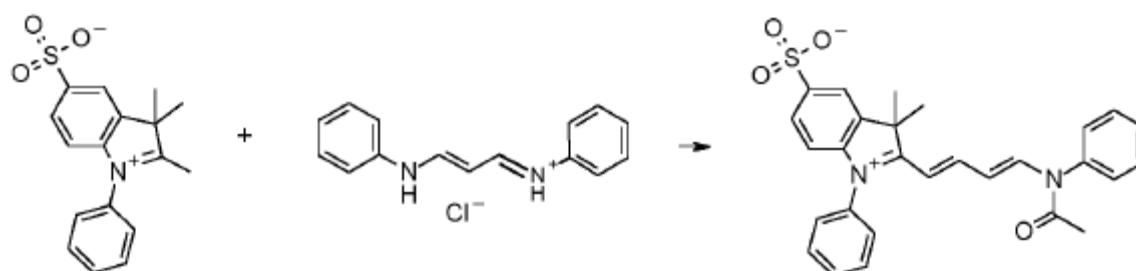


5 Una mezcla de 2,3,3-trimetil-1-fenil-3H-indolio-5-sulfonato (0.63 g), anilina (0.2 g) y 1,1,3,3-tetraetoxipropano (0.35 ml) se calentó a 70°C durante 90 min. Se formó un fundido rojo. El producto se trituró con éter de dietilo y se filtró. Rendimiento 0.6 g (68 %).

2-(4-Acetanilidobutadienil-1)-3,3-trimetil-1-fenil-3H-indolio-5-sulfonato (2-2)



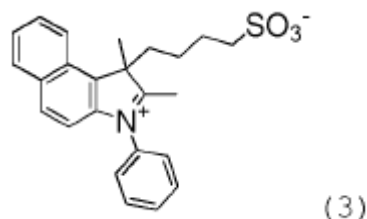
Esquema de reacción:



10

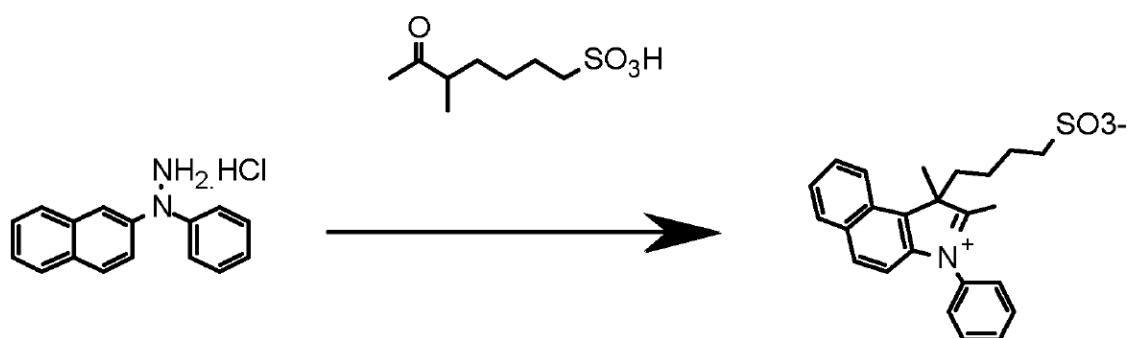
Una mezcla de 2,3,3-trimetil-1-fenil-3H-indolio-5-sulfonato (0.315 g), clorhidrato de dialdehidianilo de ácido malónico (0.25 g), ácido acético (1 ml) y anhídrido acético (2 ml) se calentó a 60°C durante 3 horas y luego a 50°C durante la noche. Se formó una solución naranja. El producto se filtró y se lavó con éter de dietilo. Rendimiento 0.24 g (49%).

1,2-dimetil-1-(4-sulfonatobutil)-3-fenil-1H-benzo[e]indolio (3)



5

Esquema de reacción:



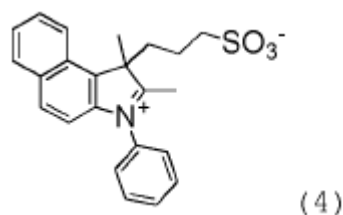
El clorhidrato de N-(2-naftilo), N-fenilhidracina (19.51 mmol, 5.28 g), ácido 5-metil-6-oxoheptanosulfónico (17.18 mmol, 3.70 g) y ZnCl<sub>2</sub> anhidro (17.18 mmol, 2.34 g) en etanol absoluto (30 ml) se agitó a temperatura ambiente durante 30 minutos, luego a 80°C durante 2 horas. El progreso de la reacción se verificó mediante TLC (10% de H<sub>2</sub>O en CH<sub>3</sub>CN). Después de terminación, la reacción se enfrió y el solvente se eliminó bajo vacío. El residuo se disolvió en DCM y se purificó por columna flash sobre gel de sílice. Rendimiento: 3.06 g, 42%.

10

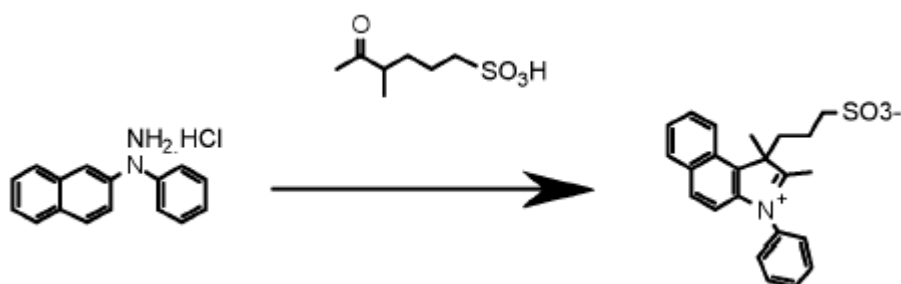
RMN de protones: (MeOH-D<sub>4</sub>): 8.28 (0.5H, d, J = 8Hz); 8.05-8.02 (1H, m); 7.89 (0.5H, d, J = 8Hz); 7.75-7.66 (3H, m); 7.65-7.60 (1H, m); 1.49-1.43 (1.5H, m); 7.31-7.25 (2H, m); 7.16 (.5H, d, J = 9Hz); 7.07 (.5H, appt, J = 7.4Hz); 6.61 (0.5H, d, J = 8Hz); 2.85-2.35 (4H, m); 1.88 (3H, appd, J = 9Hz); 1.75-1.4 (5H, m); 1.35-1.25 (0.5H, m); 1.1-0.95 (0.5H, m); 0.8-0.65 (0.5H, m); 0.58-0.45 (0.5H, m).

15

1,2-Dimetil-1-(3-sulfonatopropil)-3-fenil-1H-benzo[e]indolio (4)



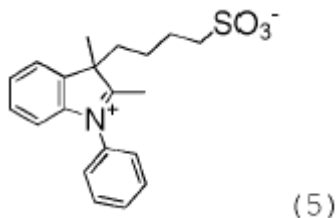
Esquema de reacción:



20

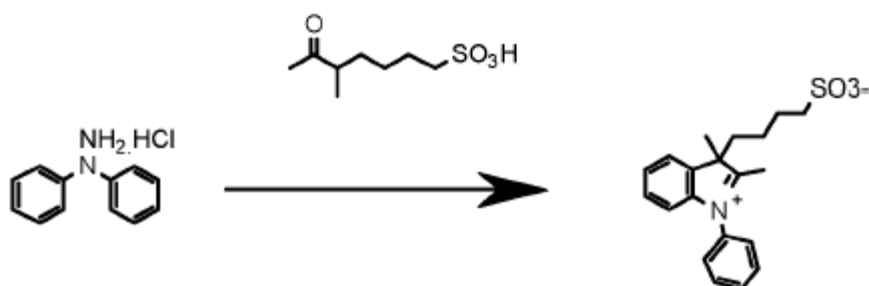
El compuesto del título se preparó como el compuesto anterior a partir de clorhidrato de N-(2-naftil)-N-fenilhidrazina y ácido 4-metil-5-oxopentanosulfónico. El producto se purificó por columna flash sobre gel de sílice. Rendimiento: 40%. Estructura confirmada por espectro de RMN.

2,3-dimetil-3-(4-sulfonatobutil)-1-fenil-3H-indolio (5)



5

Esquema de reacción:

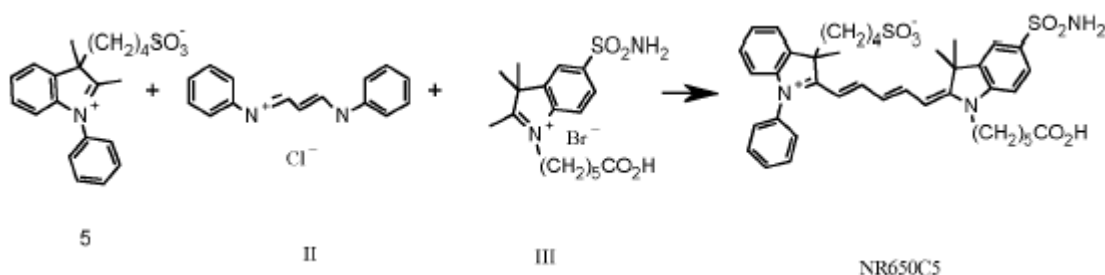


Clorhidrato de N,N-difenilhidrazina (0.01 mol, 2.2 g), ácido 5-metil-6-oxoheptanosulfónico (0.017 mol, 3.0 g) en ácido acético glacial (20 ml) se agitaron a temperatura ambiente (~20 °C) durante una hora y luego a 100 °C durante 3 horas (verificación por TLC). La mezcla de reacción se enfrió y el solvente se eliminó bajo vacío. El residuo se lavó con éter de dietilo y se purificó por columna flash sobre gel de sílice. Rendimiento: 2 g (56%). Estructura confirmada por espectro de RMN.

10

Nombre del producto: NR650C5

Esquema de síntesis de producto



15

Procedimiento de síntesis de tinte:

Se agitaron 2,3-dimetil-3-(4-sulfonatobutil)-1-fenil-3H-indolio (5) (70 mg) y clorhidrato de malondialdehído dianilo (II, 51 mg) en una mezcla de anhídrido acético y ácido acético (4 ml/1 ml) a 90 °C durante 2 h (reacción de control por TLC 15% de H<sub>2</sub>O en MeCN). La mezcla de reacción se dejó durante la noche a temperatura ambiente. Los solventes se eliminaron al vacío, el residuo amarillo se lavó con éter de dietilo y se disolvió en ácido acético (5 ml). La sal de indolio (III, 85 mg) se agregó a la solución del compuesto intermedio preparado en la etapa anterior, seguido de piridina (0.5 ml). Esta mezcla de reacción se agitó a 80 °C durante 3 h y se dejó durante la noche a temperatura ambiente. Los solventes se eliminaron bajo vacío. El residuo azul oscuro se lavó con éter de dietilo, se disolvió en una mezcla de agua y acetonitrilo (~ 5%), se filtró y se purificó por HPLC preparativo. Se recolectaron fracciones de color azul (absorción máxima ~ 650 nm) y los solventes se eliminaron al vacío.

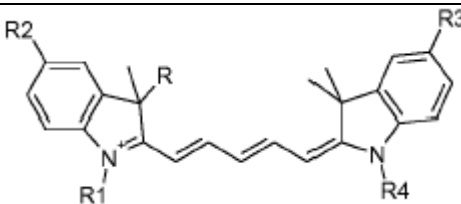
20

25

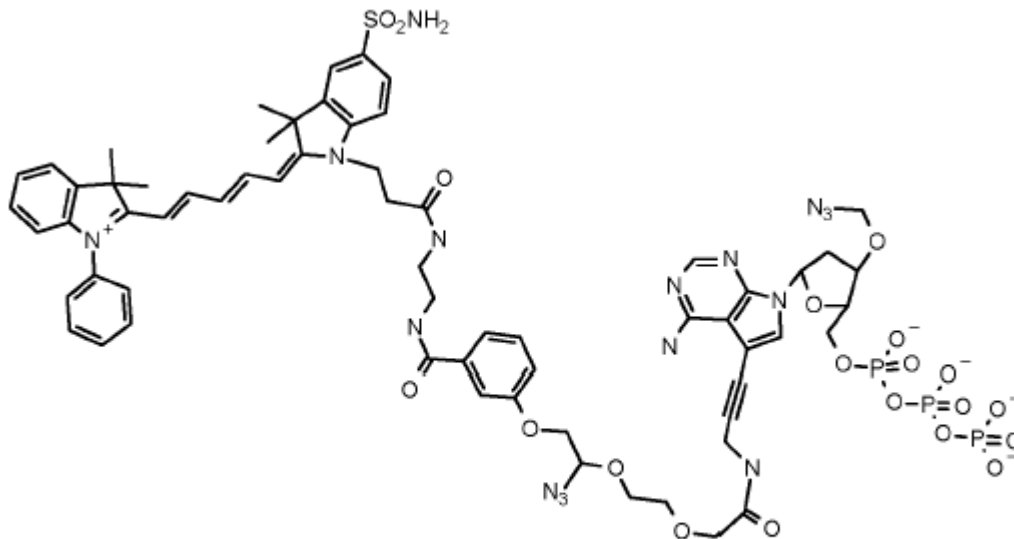
Los tintes D1-D7 se prepararon de manera similar utilizando materiales de partida apropiados.

En la Tabla 1, las propiedades espectrales de algunos tintes preparados de esta manera (soluciones en agua) en comparación con parámetros similares del tinte análogo estructural conocido Std (EP1810998A2).

Tabla 1

|  |                                                            |                              |                                 |                                                   |                                                            |            |              |
|------------------------------------------------------------------------------------|------------------------------------------------------------|------------------------------|---------------------------------|---------------------------------------------------|------------------------------------------------------------|------------|--------------|
| Tinte                                                                              | R1                                                         | R2                           | R3                              | R4                                                | R                                                          | Abs Max nm | Fluor Max nm |
| D1                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | H                            | SO <sub>3</sub> <sup>-</sup>    | (CH <sub>2</sub> ) <sub>5</sub> CO <sub>2</sub> H | C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> | 650        | 670          |
| D2                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | Benzo                        | SO <sub>3</sub> <sup>-</sup>    | (CH <sub>2</sub> ) <sub>5</sub> CO <sub>2</sub> H | C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> | 671        | 693          |
| D4                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | H                            | SO <sub>2</sub> NH <sub>2</sub> | (CH <sub>2</sub> ) <sub>2</sub> CO <sub>2</sub> H | CH <sub>3</sub>                                            | 649        | 652          |
| D5                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | SO <sub>3</sub> <sup>-</sup> | SO <sub>2</sub> NH <sub>2</sub> | (CH <sub>2</sub> ) <sub>2</sub> CO <sub>2</sub> H | CH <sub>3</sub>                                            | 651        | 671          |
| D6                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | H                            | SO <sub>2</sub> NH <sub>2</sub> | (CH <sub>2</sub> ) <sub>2</sub> CO <sub>2</sub> H | C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> | 652        | 674          |
| D7                                                                                 | C <sub>6</sub> H <sub>5</sub>                              | H                            | SO <sub>2</sub> NH <sub>2</sub> | (CH <sub>2</sub> ) <sub>5</sub> CO <sub>2</sub> H | C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> | 650        | 675          |
| std                                                                                | C <sub>4</sub> H <sub>8</sub> SO <sub>3</sub> <sup>-</sup> | H                            | SO <sub>2</sub> NH <sub>2</sub> | (CH <sub>2</sub> ) <sub>2</sub> CO <sub>2</sub> H | CH <sub>3</sub>                                            | 643        | 665          |

Tinte de conjugado de nucleótidos D4 (FFA4)



**ff-dATP- D4**

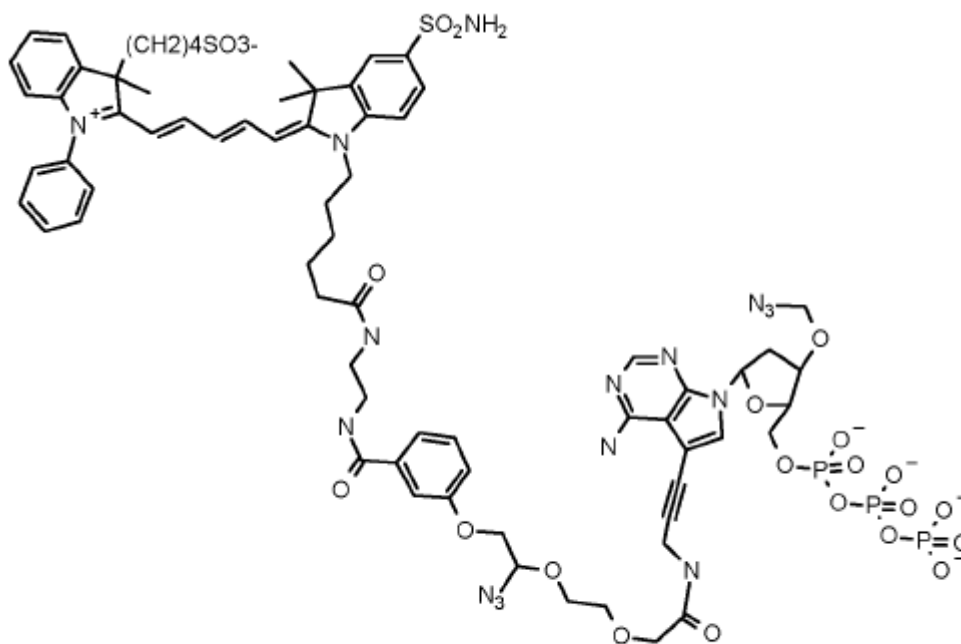
##### 5 Preparación:

Se agregaron DMA anhidro (5 ml) y base de Hunig (0.06 mL) a la muestra seca del tinte D4 (88 mg). Luego se agregó una solución de TSTU (25 mg) en 5 mL de DMA seco. Se desarrolló el color azul del éster activado. La mezcla de reacción se agitó a temperatura ambiente durante 1 h. De acuerdo con la TLC (20% de H<sub>2</sub>O en CH<sub>3</sub>CN) se completó la activación. Una vez completada la activación, esta solución se agregó a la solución de pppA-LN3 como una sal de trietilamonio (47 mg) en agua (7 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. El progreso del acoplamiento se verificó mediante TLC (20% de H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~4°C con un baño de hielo, luego se agregó una solución de TEAB 0.1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó



a una columna con ~50 g de suspensión de resina DEAE sephadex en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Las fracciones coloreadas se recolectaron y se evaporaron, luego se evaporaron de nuevo con agua para eliminar más TEAB y se sometieron a vacío hasta secado. El residuo luego se volvió a disolver en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa con un tamaño de poro de 0.2 nm en un matraz y se almacenó en el congelador. El producto se purificó por HPLC utilizando una columna de fase inversa C18 con acetonitrilo- TEAB 0.1 M. Rendimiento 76% (con base en la densidad óptica de la solución utilizando el coeficiente de extinción estimado).

Tinte de conjugado de nucleótidos D7 (FFA7)



**ff-dATP-D7**

Preparación:

Se agregaron DMA anhidra (5 ml) y base de Hunig (0.06 ml) a la muestra seca del tinte D7 (100 mg). Luego se agregó una solución de TSTU (25 mg) en 5 mL de DMA seco. Se desarrolló el color azul del éster activado. La mezcla de reacción se agitó a temperatura ambiente durante 1 h. De acuerdo con la TLC (20% de H<sub>2</sub>O en CH<sub>3</sub>CN) se completó la activación. Una vez completada la activación, esta solución se agregó a la solución de pppA-LN3 como una sal de trietilamonio (47 mg) en agua (7 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. El progreso del acoplamiento se verificó mediante TLC (20% de H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~4°C con un baño de hielo, luego se agregó una solución de TEAB 0.1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó a una columna con ~50 g de suspensión de resina DEAE sephadex en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Las fracciones coloreadas se recolectaron y se evaporaron, luego se evaporaron de nuevo con agua para eliminar más TEAB y se sometieron a vacío hasta secado. El residuo luego se volvió a disolver en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa con un tamaño de poro de 0.2 nm en un matraz y se almacenó en el congelador. El producto se purificó por HPLC utilizando una columna de fase inversa C18 con acetonitrilo- TEAB 0.1 M. Rendimiento 57% (con base en la densidad óptica de la solución utilizando el coeficiente de extinción estimado).

Los conjugados con tintes D1-D7 se prepararon de manera similar utilizando materiales de partida apropiados.

En la Tabla 2, se presentan las propiedades fluorescentes de algunos conjugados marcados con nuevos tintes preparados de esta manera en comparación con parámetros similares de análogo estructural conocido con base en el tinte conocido Std (US7109314B2, EP1810998A2).

Tabla 2

| Nucleótido | Tinte | Intensidad de fluorescencia 20°C (%) | Intensidad de fluorescencia 60°C (%) |
|------------|-------|--------------------------------------|--------------------------------------|
|------------|-------|--------------------------------------|--------------------------------------|

|        |     |           |           |
|--------|-----|-----------|-----------|
| FFA1   | D1  | 493 (175) | 294 (167) |
| FFA4   | D4  | 365 (130) | 225 (128) |
| FFA5   | D5  | 517 (184) | 330 (188) |
| FFAStd | std | 281 (100) | 176 (100) |

A partir de estos datos, se puede ver que las intensidades de fluorescencia de los nucleótidos marcados con nuevos tintes donde R1 es fenilo es significativamente mayor en comparación con los nucleótidos marcados con análogos estructurales anteriores conocidos (std) en los que R1 no es un grupo fenilo.

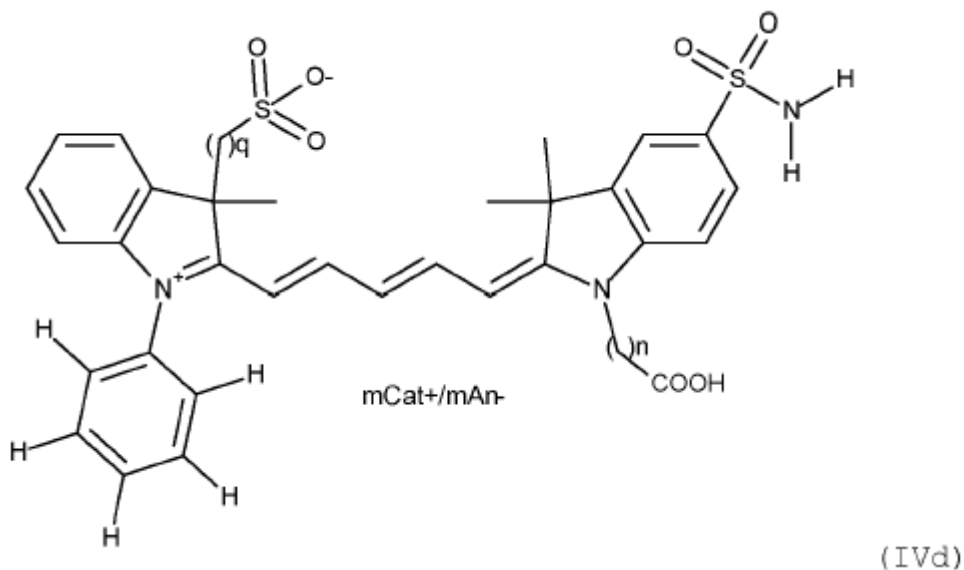
- 5 Se han analizado algunos de los nucleótidos marcados con nuevos tintes en el sistema de secuenciación de Illumina y los datos (mostrados en la tabla 3 a continuación) revelan que el uso de nuevos compuestos proporciona una tasa de error más baja para las aplicaciones de secuenciación de ácidos nucleicos que se comparan con un análogo estructural conocido en el que R1 no es fenilo.

Tabla 3

| Número de serie | Tinte | Densidad | %PF  | Ajuste de fase | Preajuste de fase | Error prm. 250 cyc R1 |
|-----------------|-------|----------|------|----------------|-------------------|-----------------------|
| M395_225650C5   | D6    | 289      | 78.8 | 0.22           | 0.13              | 1.26                  |
| M395_196650C4A  | D7    | 257      | 83.6 | 0.18           | 0.10              | 1.25                  |
| M395_131,V,4    | Std   | 310      | 79.6 | 0.15           | 0.18              | 1.69                  |

# REIVINDICACIONES

1. Un compuesto de fórmula (IVd) (IV) o formas mesoméricas del mismo:



en la que mCat+ o mAn- es un contraión orgánico o inorgánico cargado positivamente/negativamente y

5 m es un entero 0-3;

n es un entero 1-5; y

q es un entero desde 1-6.

2. Un compuesto de acuerdo con la reivindicación 1 en el que q es 4.

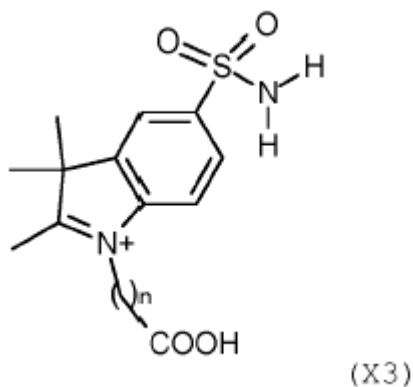
3. Un nucleótido u oligonucleótido marcado con un compuesto de acuerdo con las reivindicaciones 1-2.

10 4. Un nucleótido u oligonucleótido de acuerdo con la reivindicación 3 en el que el nucleótido u oligonucleótido se une mediante conversión del grupo COOH a una amida o éster.

5. Un kit que comprende dos o más nucleótidos en el que por lo menos un nucleótido es un nucleótido marcado de acuerdo con la reivindicación 3 o reivindicación 4.

15 6. Uso de un nucleótido de acuerdo con la reivindicación 3 o reivindicación 4, un oligonucleótido de acuerdo con la reivindicación 3 o reivindicación 4 o un kit de acuerdo con la reivindicación 5 en secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN o ensayos de unión de proteína.

7. Un método para sintetizar un compuesto de acuerdo con las reivindicaciones 1-2 utilizando el siguiente material de partida:



20 o una sal del mismo en la que n es un entero 1-5.