

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 783 851**

51 Int. Cl.:

C07D 403/06 (2006.01)

C07D 209/08 (2006.01)

C09B 23/02 (2006.01)

C09B 23/06 (2006.01)

G01N 33/00 (2006.01)

G01N 33/533 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.09.2016 PCT/GB2016/052986**

87 Fecha y número de publicación internacional: **30.03.2017 WO17051201**

96 Fecha de presentación y número de la solicitud europea: **26.09.2016 E 16775846 (5)**

97 Fecha y número de publicación de la concesión europea: **19.02.2020 EP 3353166**

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

30 Prioridad:
25.09.2015 GB 201516987

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
18.09.2020

73 Titular/es:
**ILLUMINA CAMBRIDGE LIMITED (100.0%)
19 Granta Park, Great Abington
Cambridge CB21 6DF, GB**

72 Inventor/es:
**FRANCAIS, ANTOINE y
LIU, XIAOHAI**

74 Agente/Representante:
SÁEZ MAESO, Ana

ES 2 783 851 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

5 La presente descripción se refiere a los nuevos compuestos de polimetina y su uso como marcadores fluorescentes. En particular, los compuestos pueden usarse como marcadores fluorescentes para nucleótidos en solicitudes de secuenciación de ácido nucleico.

Antecedentes

10

Se hace referencia a varias publicaciones y documentos de patente en esta solicitud con el fin de describir más completamente el estado de la técnica a la cual pertenece esta descripción.

15

La detección no radiactiva de ácidos nucleicos utilizando marcadores fluorescentes es una tecnología importante en biología molecular. Muchos procesos empleados en la tecnología de ADN recombinante dependían anteriormente en gran medida del uso de nucleótidos o polinucleótidos marcados radiactivamente con, por ejemplo, P³². Los compuestos radiactivos permiten la detección sensible de ácidos nucleicos y otras moléculas de interés. Sin embargo, existen serias limitaciones en el uso de isótopos radiactivos, tal como su gasto, vida útil limitada y, lo más importante consideraciones de seguridad. Al eliminar la necesidad de marcadores radiactivos se mejora la seguridad mientras que se reduce el impacto ambiental y los costos asociados con, por ejemplo, la eliminación de reactivos. Los métodos susceptibles de detección fluorescente no radiactiva incluyen, por medio de ejemplo no limitante, secuenciación automática de ADN, métodos de hibridación, detección en tiempo real de productos de la reacción en cadena de la polimerasa e inmunoensayos.

25

Para muchas solicitudes, es conveniente emplear múltiples marcadores fluorescentes espectralmente distinguibles con el fin de lograr la detección independiente de una pluralidad de analitos superpuestos espacialmente. En tales métodos múltiples, el número de recipientes de reacción puede reducirse, al simplificar los protocolos experimentales y facilitar la producción de kits de reactivos específicos de la aplicación. En la secuenciación de ADN automatizada de múltiples colores, por ejemplo, la detección fluorescente múltiple permite el análisis de múltiples bases de nucleótidos en un solo carril de electroforesis, de esta manera aumenta el rendimiento sobre los métodos de un solo color y reduce las incertidumbres asociadas con las variaciones de movilidad electroforética entre carriles.

30

Sin embargo, la detección fluorescente múltiple puede ser problemática y existen una serie de factores importantes que restringen la selección de marcadores fluorescentes. Primero, puede ser difícil encontrar compuestos colorantes cuyos espectros de emisión se resuelvan espectralmente de forma adecuada en una solicitud dada. Además, cuando se usan varios colorantes fluorescentes juntos, para generar señales de fluorescencia en regiones espectrales distinguibles por excitación simultánea puede ser difícil, porque las bandas de absorción de los colorantes que pueden usarse para esto, generalmente están ampliamente separadas, por lo que es difícil lograr más o menos la eficiencia de excitación de fluorescencia igual incluso para dos colorantes. Muchos métodos de excitación usan fuentes de luz de alta energía como láseres y, por lo tanto, el colorante debe tener suficiente fotoestabilidad para resistir dicha excitación. Una consideración final de particular importancia en los métodos de biología molecular es la medida en la cual los colorantes fluorescentes deben ser compatibles con la química de los reactivos usados, tal como por ejemplo disolventes y reactivos de síntesis de ADN, tampones, enzimas polimerasas y enzimas ligasas.

35

40

45

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

50

Las moléculas de colorante fluorescente con propiedades de fluorescencia mejoradas, tales como intensidad de fluorescencia, la forma y la longitud de onda máxima de la banda de fluorescencia, pueden mejorar la velocidad y precisión de la secuenciación del ácido nucleico. La fuerte señal de fluorescencia es especialmente importante cuando las mediciones se realizan en tampones biológicos a base de agua y a altas temperaturas, ya que la intensidad de fluorescencia de la mayoría de los colorantes es significativamente menor en tales condiciones. Además, la naturaleza de la base a la cual se une un colorante también afecta el máximo de fluorescencia, la intensidad de la fluorescencia y otras propiedades del colorante espectral. Las interacciones específicas de secuencia entre las nucleobases y los colorantes fluorescentes pueden adaptarse mediante un diseño específico de los colorantes fluorescentes. Las optimizaciones de la estructura de los colorantes fluorescentes pueden mejorar la eficiencia de la incorporación de nucleótidos, reducir el nivel de errores de secuenciación y disminuir el uso de reactivos en, y por lo tanto, los costos de la secuenciación del ácido nucleico.

55

60

65

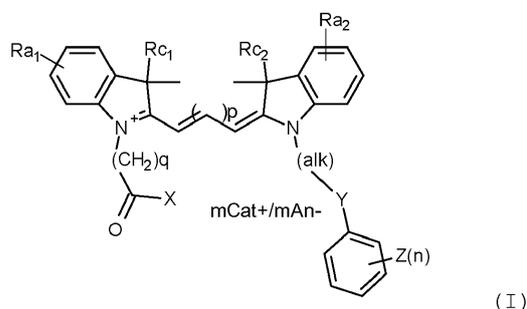
En la presente descripción se describen construcciones de polimetina mejoradas y su uso como marcadores de biomoléculas, particularmente como marcadores para nucleótidos usados en la secuenciación del ácido nucleico. Se pueden ver mejoras particulares en la eficiencia de la incorporación de nucleótidos marcados y la longitud y precisión de la secuencia de lectura obtenible mediante el uso de las nuevas construcciones fluorescentes. Las moléculas descritas a continuación son particularmente ventajosas en situaciones donde se usan fuentes de excitación de alta energía. La

excitación de alta energía puede dar como resultado un grado de saturación de la energía, donde el número de fotones emitidos por los marcadores deja de ser lineal con un aumento en la energía de excitación. La saturación de la energía puede ser causada por efectos de enfriamiento molecular en relación con la estructura de los marcadores, o puede ser causada por marcadores que tienen una vida útil de fluorescencia más larga que la deseada, de manera que los marcadores pasan un período más largo de lo deseado en el estado excitado antes de emitir un fotón. Donde múltiples marcadores diferentes se usan en paralelo, tener uno de los marcadores que alcanza la saturación de la energía cuando otros marcadores continúan siendo más brillantes, hace que la diferenciación de los marcadores sea más difícil. Los marcadores descritos en la presente descripción son particularmente ventajosos para alcanzar la saturación de la energía a un nivel más alto que los correspondientes compuestos marcados de la técnica anterior. Los documentos WO2014/135221, WO2015/170102, US5486616, US4839265 y WO2007/020457 describen compuestos de la técnica.

Resumen

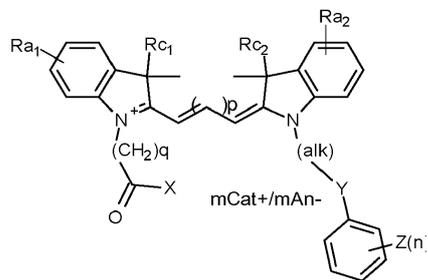
La invención se define por las reivindicaciones.

De acuerdo con un primer aspecto, esta descripción proporciona compuestos de la fórmula (I) o formas mesoméricas de los mismos:



en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3;
 p es un entero de 1-2;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 Y es S, O o CH2;
 Z es OH;
 n es un entero 0-3;
 X es OH u O⁻ o un conjugado de éster o amida del mismo;
 cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;
 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

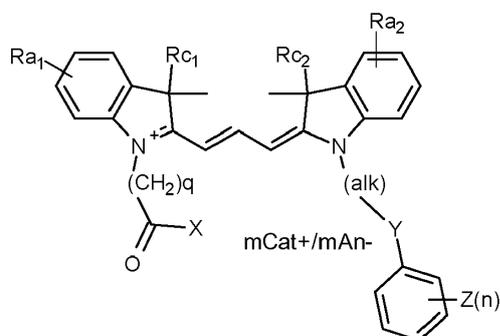
La invención proporciona compuestos de la fórmula (I) o formas mesoméricas de los mismos:



en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3;
 p es un entero de 1-2;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 Y es O;
 Z es OH;
 n es 1;
 X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y
 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO₃⁻, o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico.

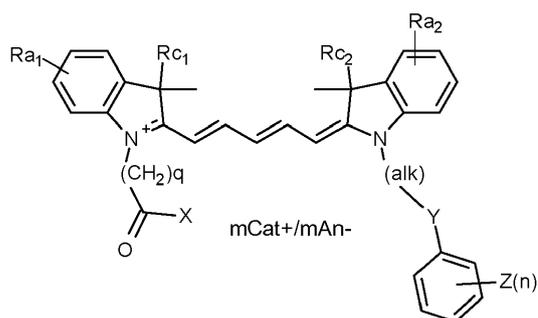
P puede ser uno; en cuyo caso los compuestos son de la fórmula (Ia) o formas mesoméricas de los mismos:



(Ia)

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 Y es S, O o CH₂;
 Z es OH;
 n es un entero 0-3;
 X es OH u O⁻ o un conjugado de éster o amida del mismo;
 cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;
 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

P puede ser dos; en cuyo caso los compuestos son de la fórmula (Ib) o formas mesoméricas de los mismos:



(Ib)

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 Y es S, O o CH₂;
 Z es OH;
 n es un entero 0-3;
 X es OH u O⁻ o un conjugado de éster o amida del mismo;
 cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;
 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

5 En ciertos ejemplos donde R_{a1} o R_{a2} es H; ya sea R_{c1} o R_{c2} puede ser un grupo del ácido alquilsulfónico. Los compuestos requieren uno o más sustituyentes de ácido sulfónico, de modo que al menos uno de R_{a1} o R_{a2} es SO_3^- o R_{a1} o R_{a2} es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO_3^- o R_{c1} o R_{c2} es un grupo del ácido alquilsulfónico.

En ciertos ejemplos donde p es 2, R_{c1} o R_{c2} es un grupo del ácido alquilsulfónico.

10 En ciertos ejemplos, r es igual a 3.

En ciertos ejemplos, n es 1. El grupo OH de Z puede estar en la posición 4 en el anillo. En ciertos ejemplos, n es 0 y el anillo de fenilo no está sustituido. En ciertos ejemplos donde n es 0, Y es S u O (y no CH_2).

15 En otra modalidad, los compuestos de la presente descripción 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 puede llevarse a cabo a través del grupo carboxilo $C(=O)-X$, que se puede convertir en una amida o un éster.

20 De acuerdo con un aspecto adicional de la descripción, por lo tanto, se proporcionan compuestos colorantes que comprenden grupos de enlace para permitir, por ejemplo, la unión covalente a tales fracciones de sustrato.

25 De acuerdo con un aspecto adicional, la descripción proporciona un compuesto de nucleósido o de nucleótido definido por la fórmula: N-L-Dye, en donde N es un nucleótido, L es una fracción enlazadora opcional y Dye es un compuesto fluorescente de acuerdo con la presente descripción.

En un aspecto adicional, la descripción proporciona métodos de secuenciación mediante el uso de los compuestos colorantes de la presente descripción.

30 De acuerdo con un aspecto adicional, la descripción también proporciona kits que comprenden compuestos colorantes (libres o en forma conjugada) que pueden usarse en varios ensayos inmunológicos, oligonucleótidos y ácidos nucleicos marcados y para la secuenciación de ADN por síntesis. En otro aspecto más, la descripción proporciona kits que comprenden 'conjuntos' de colorantes particularmente adecuados para ciclos de secuenciación por síntesis en una plataforma de instrumentos automatizada.

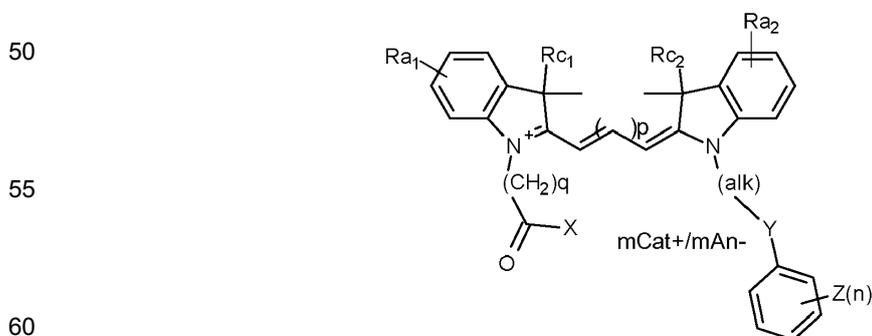
35 Un aspecto adicional de la descripción es la preparación química de los compuestos de la descripción.

Descripción detallada

40 Esta descripción proporciona nuevos compuestos particularmente adecuados para métodos de detección de fluorescencia y secuenciación por síntesis. Los compuestos que tienen una fracción 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 del ácido nucleico.

45 La invención se define por las reivindicaciones.

De acuerdo con un primer aspecto la descripción proporciona compuestos de la fórmula (I) o formas mesoméricas de los mismos:



(I)

65 en donde $mCat^+$ o mAn^- es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

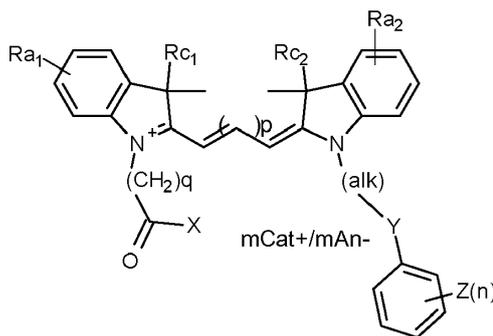
n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

La descripción proporciona un compuesto de la fórmula (I) o formas mesoméricas del mismo:



(I)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

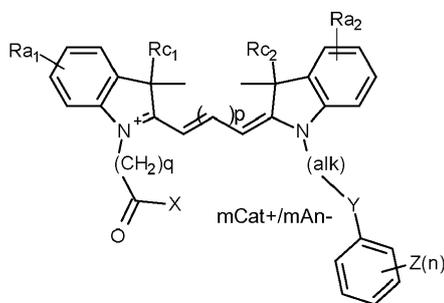
n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO₃⁻, o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico.

La descripción proporciona un compuesto de la fórmula (I) o formas mesoméricas del mismo:



(I)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

5 n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

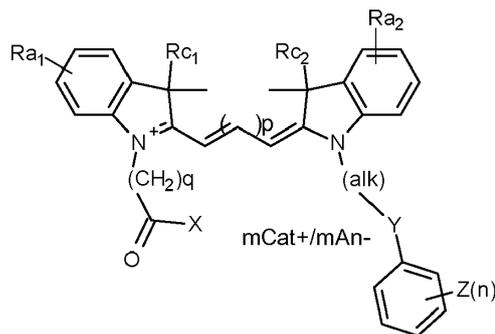
cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

10 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde cuando n es 0, Y es S u O.

10

La descripción proporciona un compuesto de la fórmula (I) o formas mesoméricas del mismo:

15



20

25

(I)

en donde mCat⁺ o mAn⁻ es un contraión cargado positivamente/negativamente orgánico o inorgánico y

30

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

35

Z es OH;

n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

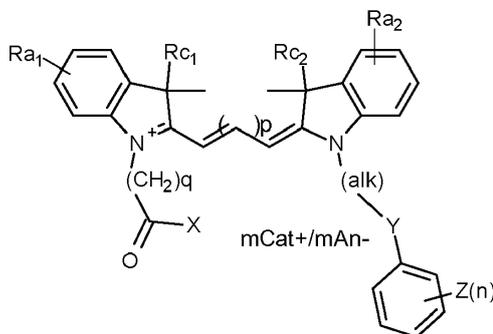
cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

40

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO₃⁻ o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico y en el que cuando n es 0, Y es S u O.

La invención proporciona compuestos de la fórmula (I) o formas mesoméricas de los mismos:

45



50

55

(I)

en donde mCat⁺ o mAn⁻ es un contraión cargado positivamente/negativamente orgánico o inorgánico y

60

m es un entero de 0-3;

p es un entero de 1-2;

65

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es O;
Z es OH;
n es 1;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

5 cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

10 cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO₃⁻, o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico.

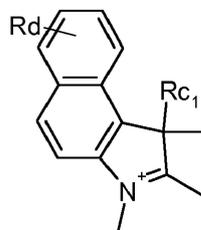
10

15 Las moléculas pueden contener una o más sulfonamidas o fracciones SO₃⁻ en la posición Ra. Ra₁ y/o Ra₂ pueden ser SO₃⁻ o sulfonamida. El otro Ra (Ra₁ o Ra₂) puede ser independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente. Ra₁ o Ra₂ puede ser H. Ra₁ o Ra₂ puede ser SO₃⁻. Ra₁ puede ser diferente a Ra₂, por ejemplo, la estructura puede tener un solo grupo sulfonamida en Ra₁ y H como Ra₂. Ra₁ y Ra₂ ambos pueden ser sulfonamida. La sulfonamida puede ser SO₂NH₂ o SO₂NHR, donde R es un grupo alquilo, alquilo sustituido, arilo o arilo sustituido. Donde ni Ra₁ o Ra₂ es un SO₃⁻ o un anillo adicional fusionado a un átomo de carbono adyacente, luego Rc₁ o Rc₂ debe ser un grupo de ácido alquilsulfónico.

15

20 Ra₁ o Ra₂ puede ser otro anillo alifático, aromático o heterocíclico fusionado a carbonos adyacentes del anillo de indol. Por ejemplo, en tales casos cuando se fusiona un anillo aromático, el grupo final de colorantes puede representar una estructura de tipo

20



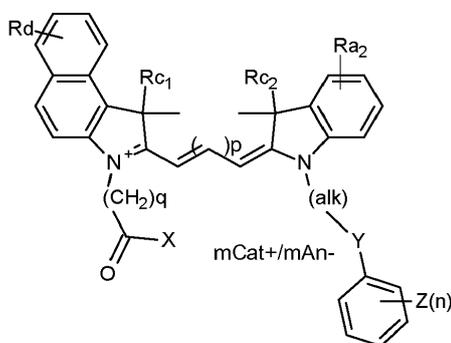
25

30

donde Rd puede ser H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico.

35 Por lo tanto, los colorantes de la descripción se pueden describir mediante la Fórmula (IC) o (ID) o formas mesoméricas de los mismos:

35

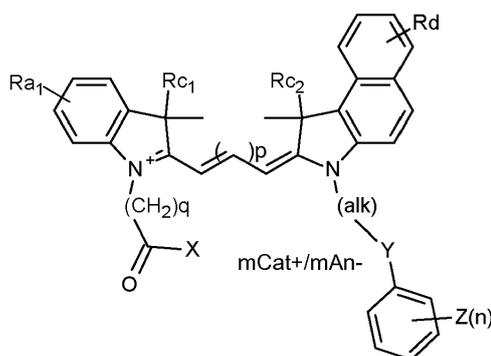


40

45

50

(IC)



55

60

65

(ID)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

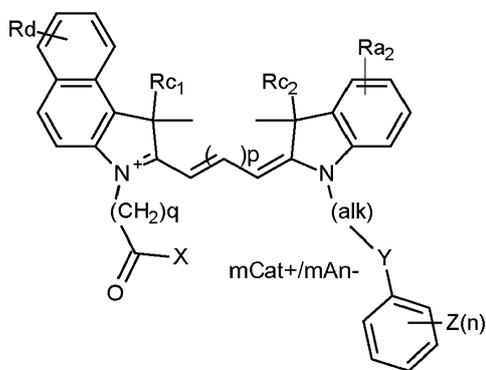
X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

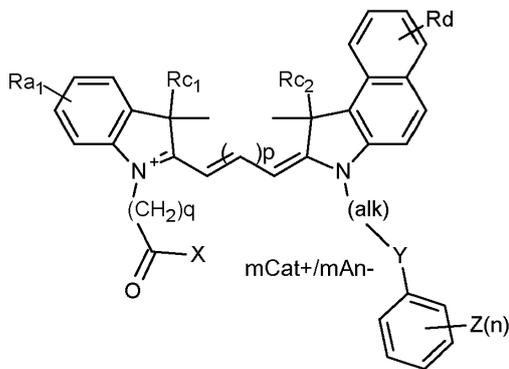
cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido; y

Rd es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico.

Por lo tanto, los colorantes de la descripción se pueden describir mediante la Fórmula (IC) o (ID) o formas mesoméricas de los mismos:



(IC)



(ID)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

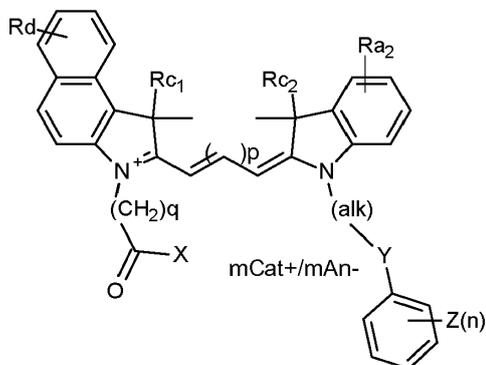
X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

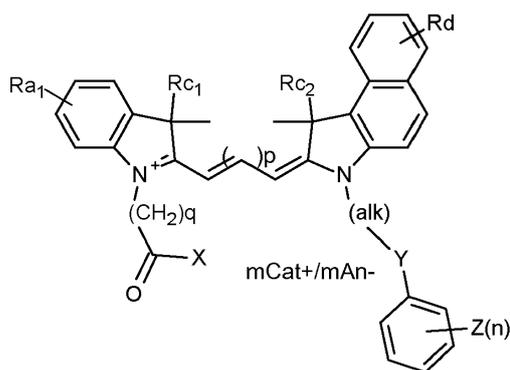
cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido; y

Rd es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Rd es SO₃⁻ o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico.

Por lo tanto, los colorantes de la descripción se pueden describir mediante la Fórmula (IC) o (ID) o formas mesoméricas de los mismos:



(IC)



(ID)

en donde mCat⁺ o mAn⁻ es un contraión cargado positivamente/negativamente orgánico o inorgánico y

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido; y

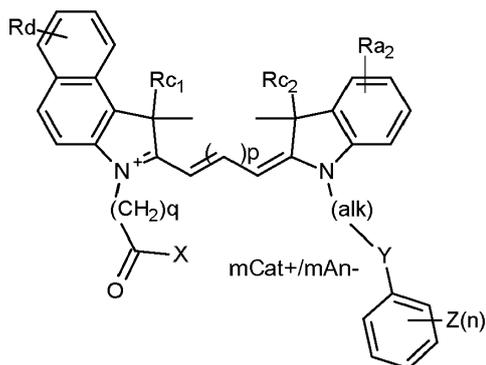
Rd es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico, en donde cuando n es 0, Y es S u O.

Por lo tanto, los colorantes de la descripción se pueden describir mediante la Fórmula (IC) o (ID) o formas mesoméricas de los mismos:

5

10

15

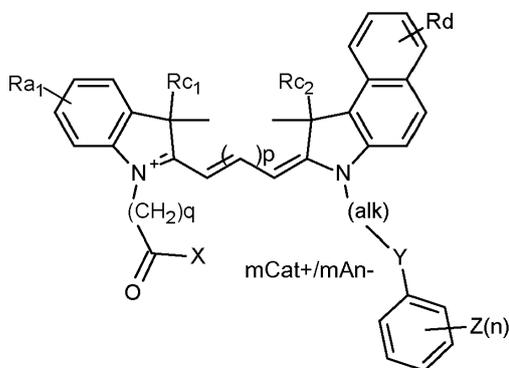


(IC)

20

25

30



(ID)

35

40

45

50

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de R_{a1} y R_{a2} es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

cada uno de R_{c1} y R_{c2} es independientemente alquilo o alquilo sustituido; y

R_d es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico, en donde al menos uno de R_{a1} o R_{a2} es SO₃⁻ o R_d es SO₃⁻ o R_{c1} o R_{c2} es un grupo de ácido alquilsulfónico y en donde cuando n es 0, Y es S u O.

55

En la fórmula (IC) o (ID), los anillos adicionales fusionados a átomos de carbono adyacentes del anillo de indol pueden estar opcionalmente sustituidos, por ejemplo, con ácido sulfónico o sulfonamida.

60

El grupo carboxi C(=O)-X o sus derivados está unido al átomo de nitrógeno indol mediante una cadena alquilo de longitud q, donde q es 1-5 carbonos o heteroátomos. La cadena puede ser (CH₂)_q donde q es 1-5. El grupo puede ser (CH₂)₅COOH.

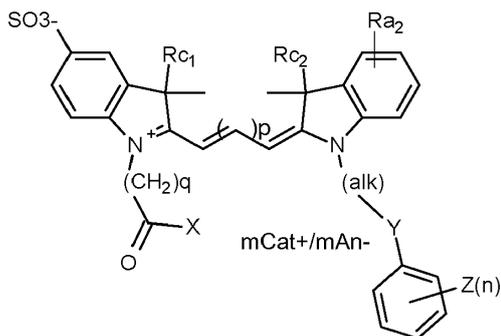
65

Las moléculas pueden contener uno o más fracciones alquilsulfonato en la posición R_c. Ya sea R_{c1} y/o R_{c2} puede ser alquil-SO₃⁻. El otro R_c (R_{c1} o R_{c2}) puede ser independientemente alquilo o alquilo sustituido. R_{c1} y R_{c2} puede ser independientemente metilo, etilo, propilo, butilo, pentilo, hexilo o (CH₂)_tSO₃H, donde t es 1-6. t puede ser 1-4. t puede ser 4. R_{c1} y R_{c2} puede ser un grupo alquilo sustituido. R_{c1} y R_{c2} puede contener una fracción COOH o -SO₃H o sus derivados éster o amida.

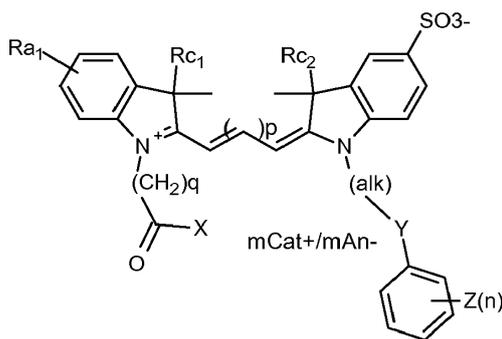
En ciertas modalidades, cuando uno de Ra₁ o Ra₂ es SO₃⁻ y el otro de Ra₁ o Ra₂ es H o SO₃⁻, ya sea Rc₁ o Rc₂ también puede ser un grupo de ácido alquilsulfónico.

El grupo COOH que se muestra como C(=O)-X puede actuar como una fracción enlazadora para una unión adicional o se une a una molécula adicional. Una vez que se produce la conjugación, el COOH o COO⁻ se convierte en una amida o éster.

Los ejemplos de compuestos incluyen las estructuras de acuerdo con la fórmula (II) o (IIa) o las formas mesoméricas de los mismos:



(II)



(IIa)

en donde mCat⁺ o mAn⁻ es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

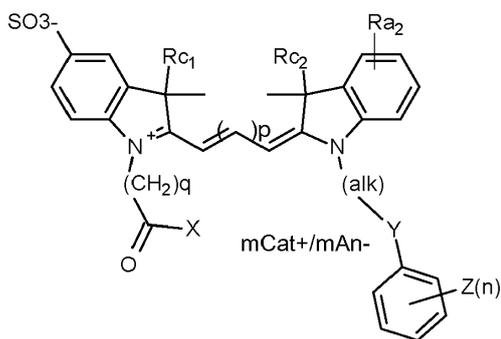
cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

Los ejemplos de compuestos incluyen las estructuras de acuerdo con la fórmula (II) o (IIa) o las formas mesoméricas de los mismos:

5

10

15

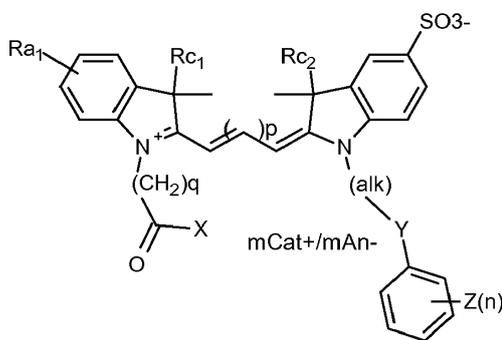


(II)

20

25

30



(IIa)

35

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

40

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

45

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde cuando n es 0, Y es S u O.

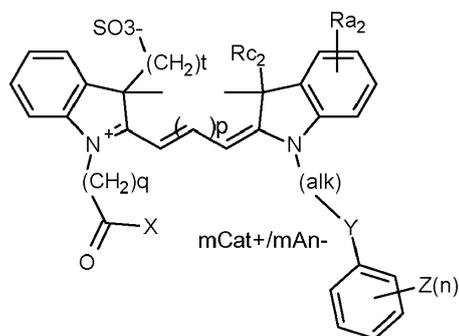
Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con la fórmula (IIIa) o (IIIb):

50

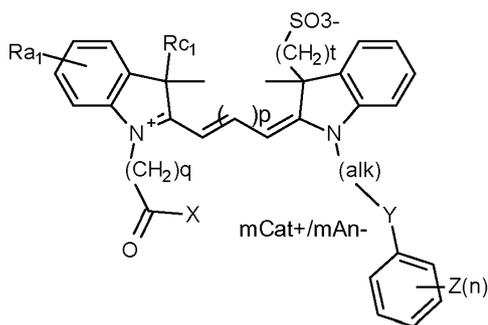
55

60

65



(IIIa)



(IIIb)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

t es un entero de 1-6;

Y es S, O o CH₂;

Z es OH;

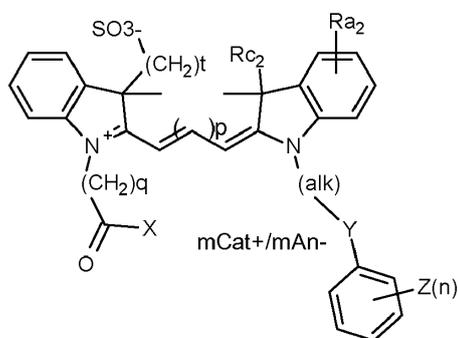
n es un entero 0-3;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

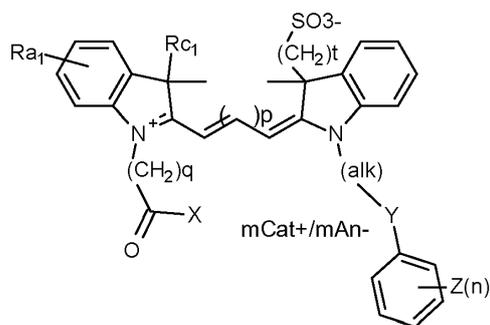
cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido.

Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con la fórmula (IIIa) o (IIIb):



(IIIa)



(IIIb)

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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

5 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

t es un entero de 1-6;

Y es S, O o CH₂;

Z es OH;

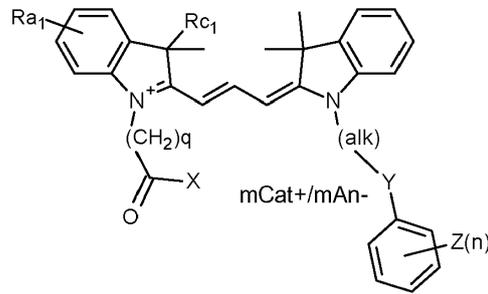
n es un entero 0-3;

10 X es OH u O⁻ o un conjugado de éster o amida del mismo;

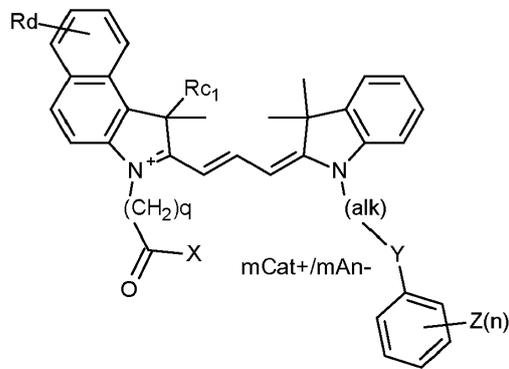
cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde cuando n es 0, Y es S u O.

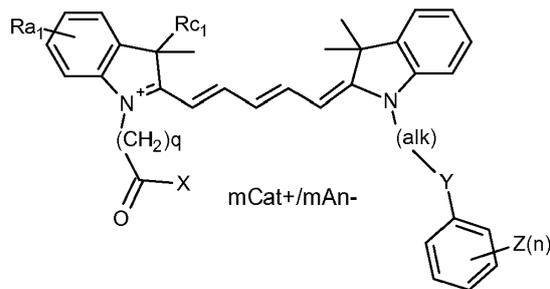
15 Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con la fórmula (IVa) a (IVd):



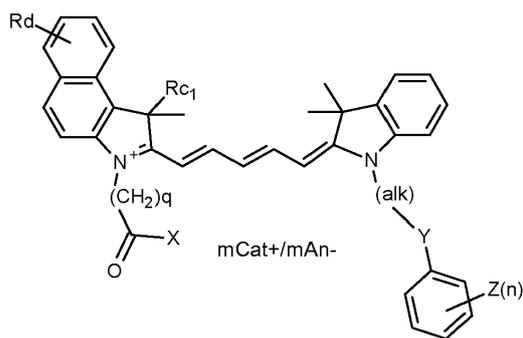
(IVa)



(IVb)



(IVc)



(IVd)

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

m es un entero de 0-3;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es S, O o CH₂;

Z es OH;

n es un entero 0-3;

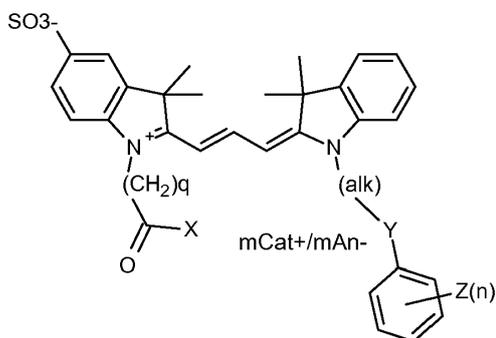
X es OH u O⁻ o un conjugado de éster o amida del mismo;

Ra₁ es H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente;

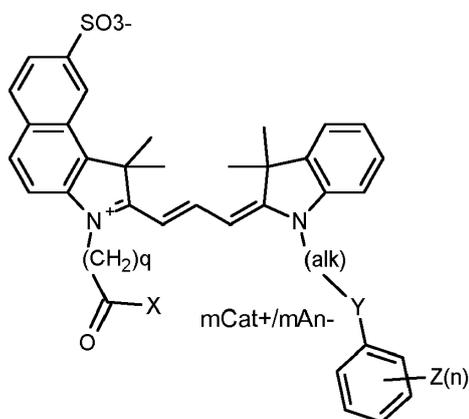
Rc₁ es alquilo o alquilo sustituido; y

Rd es H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico.

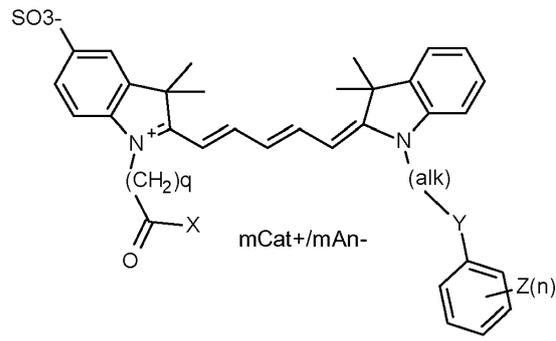
Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con la fórmula (Va) a (Vd):



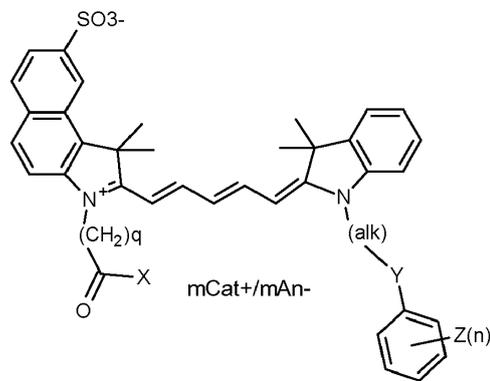
(Va)



(Vb)



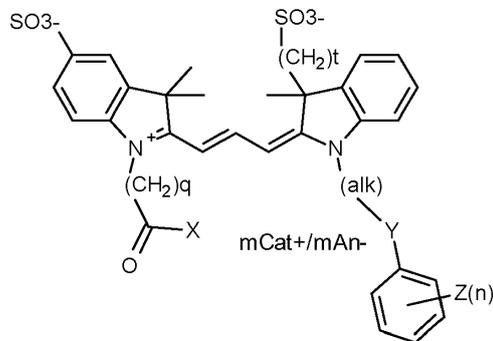
(Vc)



(Vd)

35 en donde mCat⁺ o mAn⁻ es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 Y es S, O o CH₂;
 40 Z es OH;
 n es un entero 0-3; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo.

45 Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con la fórmula (VIa) a (VId):

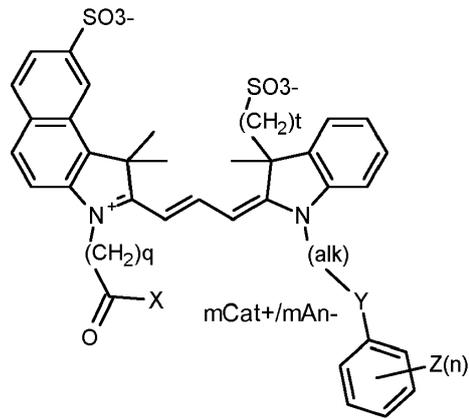


(VIa)

5

10

15



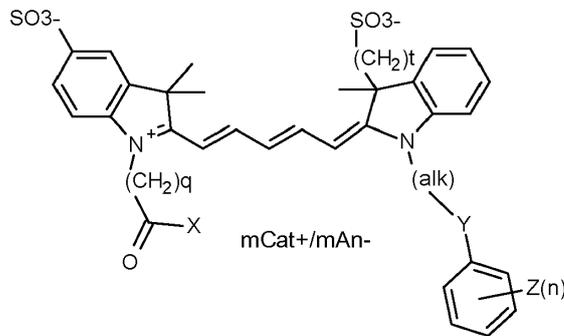
(VIb)

20

25

30

35



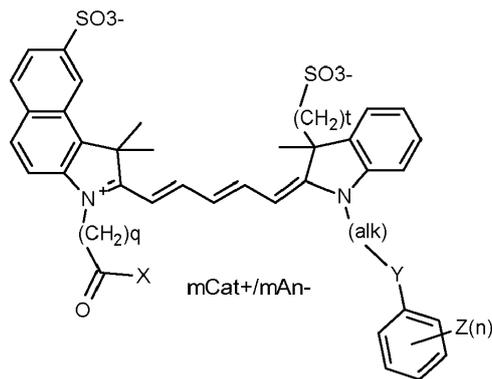
(VIc)

40

45

50

55



(VI d)

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;
 t es un entero de 1-6;
 Y es S, O o CH₂;
 Z es OH;
 n es un entero 0-3; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo.

5 alk es una cadena alquilo, alqueno o alquino de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces. Alk puede ser un grupo $(CH_2)_r$ donde r es 1-5. Alk puede ser $(CH_2)_3$. Alternativamente, la cadena de carbono puede contener uno o más dobles enlaces o triples enlaces. La cadena puede contener un enlace $-CH_2-CH=CH-CH_2-$, opcionalmente con grupos adicionales CH_2 . La cadena puede contener un enlace $-CH_2-C\equiv C-CH_2$, opcionalmente con grupos adicionales CH_2 .

En cualquiera de los ejemplos dados en las fórmulas VII a VIII; r puede ser igual a 3.

10 En cualquiera de los ejemplos dados en las fórmulas I a VIII; q puede ser igual a 5.

En cualquiera de los ejemplos dados en la fórmula III, fórmula VI o fórmula VII; t puede ser igual a 4.

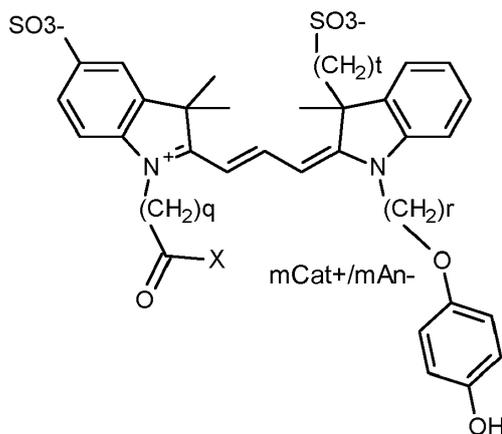
15 En cualquiera de los ejemplos dados en las fórmulas I a VI; n puede ser igual a 1-3. En cualquiera de los ejemplos dados en las fórmulas I a VI; n puede ser igual a 1. En cualquiera de los ejemplos dados en las fórmulas I a VI; n puede ser un entero 0-1. Donde n es 1, el grupo OH de Z puede estar en cualquier posición del anillo. El grupo OH de Z puede estar en la posición 4. Donde n es 2 o 3, los grupos OH de Z pueden estar en cualquier posición en el anillo de fenilo.

En cualquiera de los ejemplos dados en las fórmulas I a VI; cuando n es cero, Y puede ser igual a O o S y no CH_2 .

20 En cualquiera de los ejemplos dados en las fórmulas I a VI; Y puede ser igual a O.

En cualquiera de los ejemplos dados en las fórmulas I a VI; Y puede ser igual a O. Donde Y es O, n puede ser 0-3. Donde Y es CH_2 , n puede ser 1-3.

25 Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con las fórmulas (VIIa) a (VIId):



(VIIa)

50

55

60

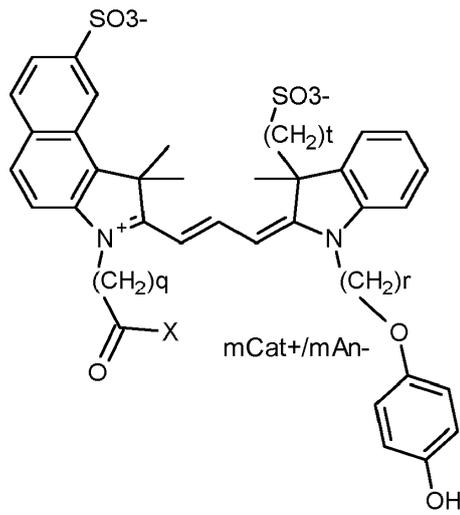
65

5

10

15

20



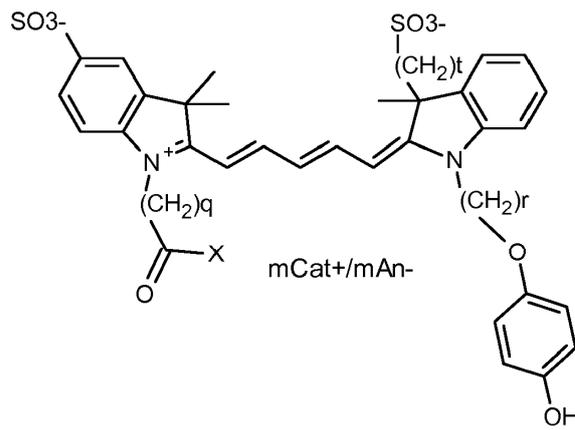
(VIIb)

25

30

35

40



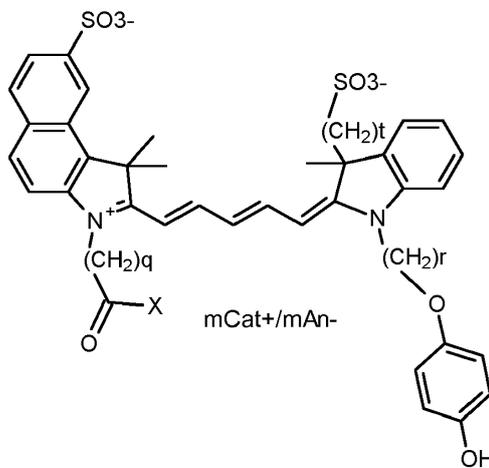
(VIIc)

45

50

55

60



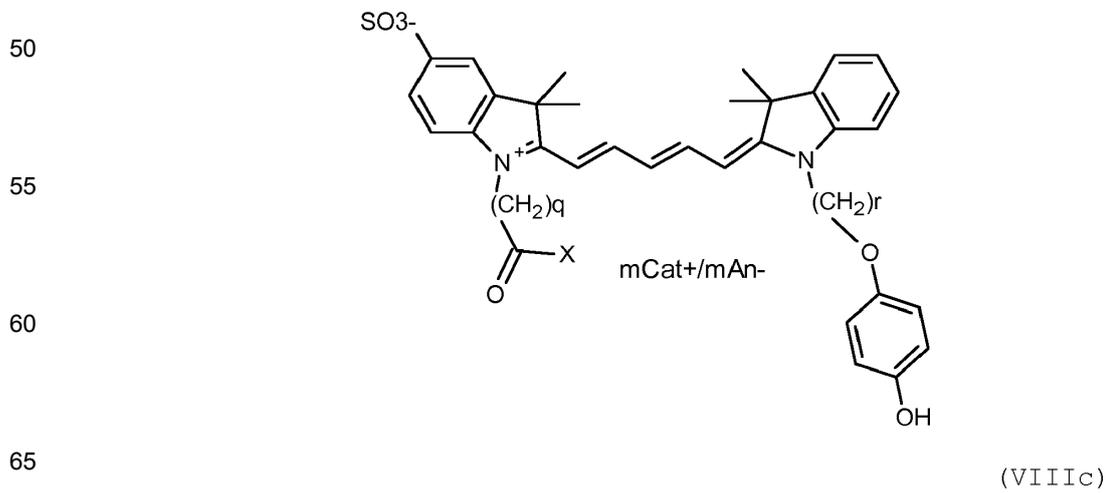
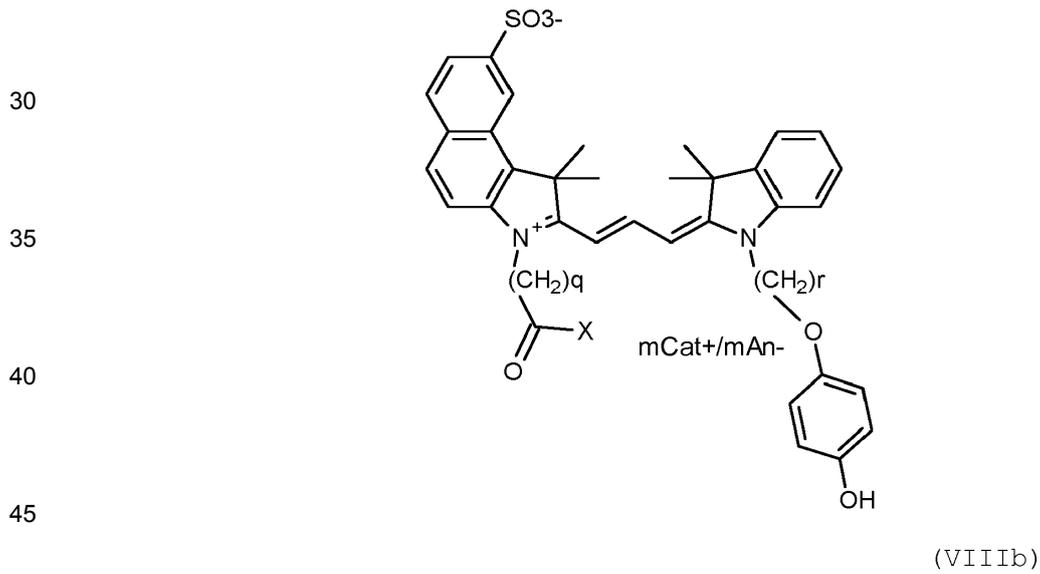
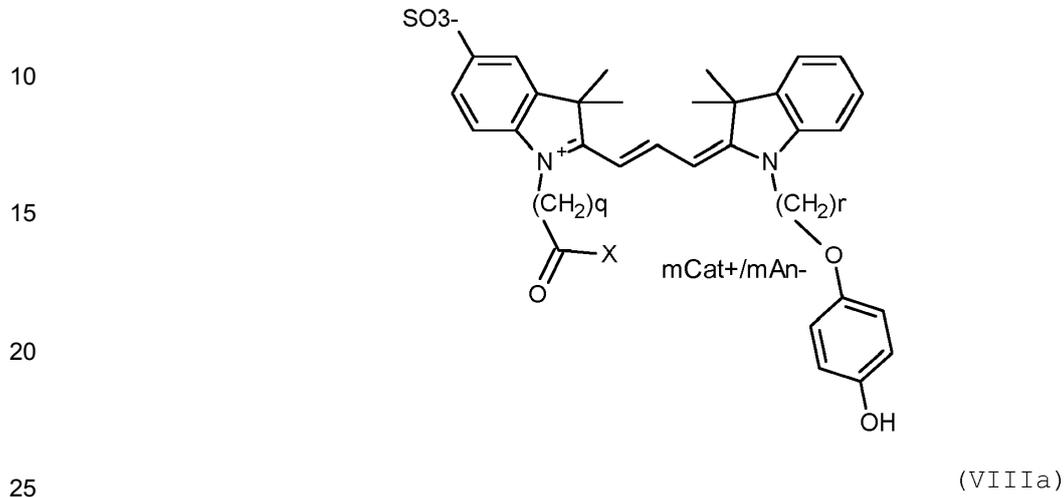
(VIIId)

65

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;

r es un entero de 1-5;
 t es un entero de 1-6; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo.

5 Los ejemplos adicionales de compuestos incluyen las estructuras de acuerdo con las fórmulas (VIIIa) a (VIIId):

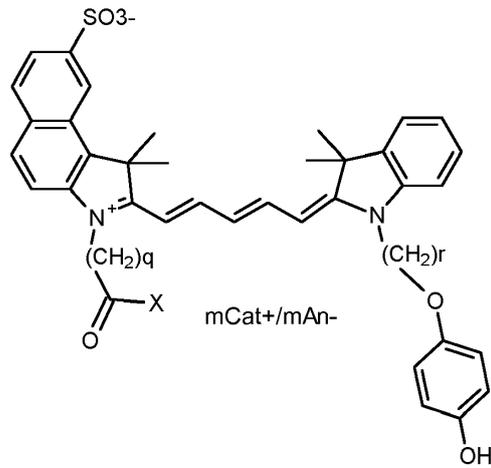


5

10

15

20



(VIIId)

25

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3; q es un entero de 1-5; r es un entero de 1-5; y X es OH u O⁻ o un conjugado de éster o amida del mismo.

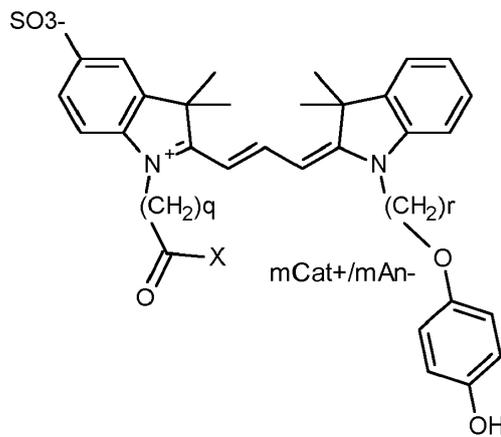
Los compuestos de la invención pueden estar representados por la fórmula (VIIIa) o la fórmula (VIIId):

30

35

40

45



(VIIIa)

50

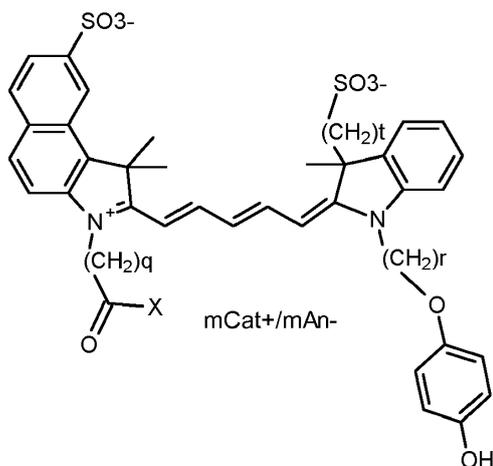
en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y m es un entero de 0-3; q es un entero de 1-5; r es un entero de 1-5; y X es OH u O⁻ o un conjugado de éster o amida del mismo;

55

60

65

5
10
15
20
25
30
35
40
45
50
55
60
65



(VIIId)

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 r es un entero de 1-5;
 t es un entero de 1-6; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo.

Un compuesto particularmente útil es un nucleótido u oligonucleótido marcado con un colorante como se describe en la presente descripción.

El nucleótido u oligonucleótido marcado puede tener el marcador unido al átomo de nitrógeno de 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 pirimidina o la posición C7 de una base 7-desaza purina a través de una fracción enlazadora.

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

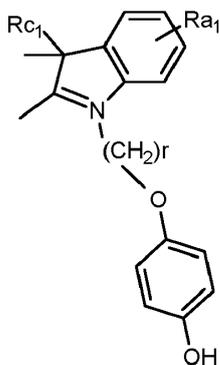
En la presente descripción se proporcionan kits que incluyen dos o más nucleótidos en donde al menos un nucleótido es un nucleótido marcado con un compuesto de la presente descripción. El kit puede incluir dos o más nucleótidos marcados. Los nucleótidos se pueden marcar con dos o más marcadores fluorescentes. Dos o más de los marcadores pueden excitarse mediante el uso de una sola fuente de excitación, que puede ser un láser. Por ejemplo, las bandas de excitación para los dos o más marcadores pueden solaparse al menos parcialmente, de modo que la excitación en la región de solapamiento del espectro hace que ambos marcadores emitan fluorescencia. En modalidades particulares, la emisión de los dos o más marcadores ocurrirá en diferentes regiones del espectro de manera que se pueda determinar la presencia de al menos uno de los marcadores mediante la distinción óptica de la emisión.

El kit puede contener cuatro nucleótidos marcados, donde el primero de cuatro nucleótidos se marca con un compuesto como se describe en la presente descripción. En dicho kit, el segundo, tercero y cuarto nucleótidos pueden marcarse 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 absorbanza y/o un máximo de emisión distintos de manera que los compuestos se pueden distinguir de otros compuestos. Por ejemplo, cada compuesto puede tener un máximo de absorbanza y/o un máximo de emisión distinto de manera que cada uno de los compuestos se puede distinguir de los otros tres compuestos. Se debe entender que las partes del espectro de absorbanza y/o espectro de emisión distintos de los máximos pueden diferir y estas diferencias pueden explotarse para distinguir los compuestos. El kit puede ser de manera que dos o más de los compuestos tienen un máximo de absorbanza por encima de 600 nm. Los compuestos de la invención típicamente absorben luz en la región por encima de 640 nm.

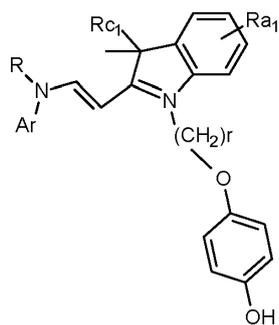
Los compuestos, nucleótidos o kits que se exponen en la presente descripción pueden usarse para detectar, medir o identificar un sistema biológico (que incluye, por ejemplo, procesos o componentes del mismo). Las técnicas ilustrativas que pueden emplear los compuestos, nucleótidos o kits incluyen secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN, ensayo celular (por ejemplo, análisis de unión celular o función celular) o ensayo de proteínas (por ejemplo, ensayo de unión a proteínas o ensayo de actividad de proteínas). El uso puede ser en

un instrumento automatizado para llevar a cabo una técnica particular, tal como un instrumento de secuenciación automatizada. El instrumento de secuenciación puede contener dos láseres que operan a diferentes longitudes de onda.

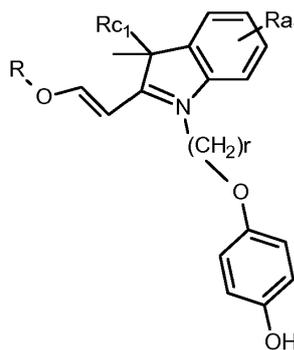
En la presente descripción se describe un método para sintetizar compuestos de la descripción. Un compuesto de fórmula (X) y/o (X1), (X2) (X3) o (X4) o una sal del mismo puede usarse como material de partida para la síntesis de colorantes de polimetina simétricos o asimétricos:



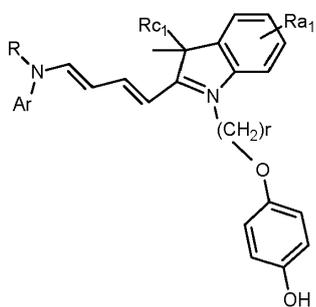
(X)



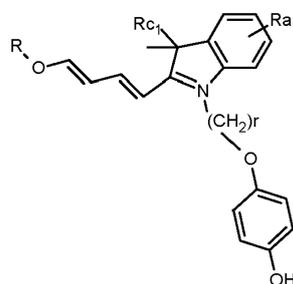
(X1)



(X2)



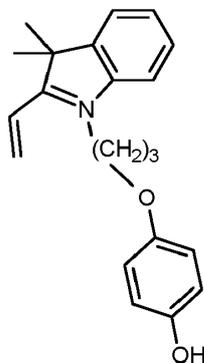
(X3)



(X4)

o una sal de los mismos en donde Ra_1 es H, SO_3^- , sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; Rc_1 es alquilo o alquilo sustituido; Ar es un grupo aromático y R es un grupo alquilo. Donde se muestran ejemplos específicos de 4-hidroxifenilo, otros grupos hidroxilo adicionales pueden estar sustituidos en el anillo en casos donde n es mayor que uno. r puede ser igual a 3.

En la presente descripción se describe un método para sintetizar compuestos de la descripción. Un compuesto de fórmula (X5) o una sal del mismo puede usarse como material de partida para la síntesis de colorantes de polimetina simétricos o asimétricos:



(X5)

Como se usa en la presente descripción, el término "alquilo" se refiere a hidrocarburo C_1-C_{20} y puede incluir anillos carbocíclicos no aromáticos C_3-C_{10} . En modalidades particulares, los grupos alquilo son alquilo C_1-C_6 que se refieren a radicales de hidrocarburo de cadena lineal o ramificada, saturados, que contienen entre uno y seis átomos de carbono, respectivamente. En casos particulares, los grupos alquilo pueden incluir uno o más grupos insaturados, y por lo tanto incluir alqueno y alquino.

El término "halógeno" como se usa en la presente descripción se refiere a fluoro- (en adelante designado como F), cloro- (en adelante designado como Cl), bromo- (en adelante designado como Br) o yodo- (en adelante designado como I), y usualmente se refiere a la sustitución de un átomo de hidrógeno en un compuesto orgánico, esta sustitución es opcionalmente una sustitución completa para el hidrógeno.

El término "alquilo sustituido" se refiere a grupos alquilo, alqueno o alquino como se definió anteriormente, donde pueden estar opcionalmente sustituidos adicionalmente con, pero no limitado a, halo, ciano, SO_3^- , SRA, ORa, NRbRc, oxo, CONRbRc, COOH y COORb. Ra, Rb y Rc pueden seleccionarse cada uno independientemente de H, alquilo, alquilo sustituido, alqueno, alqueno sustituido, alquino, alquino sustituido, arilo y arilo sustituido. Además, dicho alquilo sustituido, alqueno sustituido y alquino sustituido pueden interrumpirse opcionalmente por al menos un heteroátomo o grupo seleccionado de O, NRb, $S(O)_t$, donde t es de 0 a 2, y similares. El alquilo sustituido también cubre un grupo tal como bencilo donde los grupos alquilo comprenden una fracción adicional arilo o arilo sustituido.

Los colorantes de acuerdo con la presente descripción pueden sintetizarse a partir de una variedad de diferentes materiales de partida, que incluyen los indoles sustituidos con N-propil-4-hidroxifeniléter. Los colorantes pueden hacerse simétricamente, de modo que el mismo indol está en ambos extremos de la cadena de polimetina, o asimétricamente, de manera que diferentes indoles están en cada extremo del cromóforo. Los métodos para preparar colorantes de polimetina se conocen bien en la técnica.

De acuerdo con un aspecto de la descripción, se proporcionan compuestos colorantes adecuados para la unión a fracciones de sustrato, que comprenden particularmente grupos de enlace para permitir la unión a fracciones de sustrato. Las fracciones de sustrato pueden ser prácticamente cualquier molécula o sustancia a la cual se puedan conjugar los colorantes de la descripción y, por medio de ejemplo no limitante, pueden incluir nucleósidos, nucleótidos, polinucleótidos, carbohidratos, ligandos, partículas, superficies sólidas, polímeros orgánicos e inorgánicos, cromosomas, núcleos, células vivas y combinaciones o asociaciones de los mismos. Los colorantes pueden conjugarse mediante un enlazador opcional por una variedad de medios que incluyen atracción hidrófoba, atracción iónica y unión covalente. Particularmente, los colorantes se conjugan con el sustrato mediante unión covalente. Más particularmente, la unión covalente es por medio de un grupo de enlace.

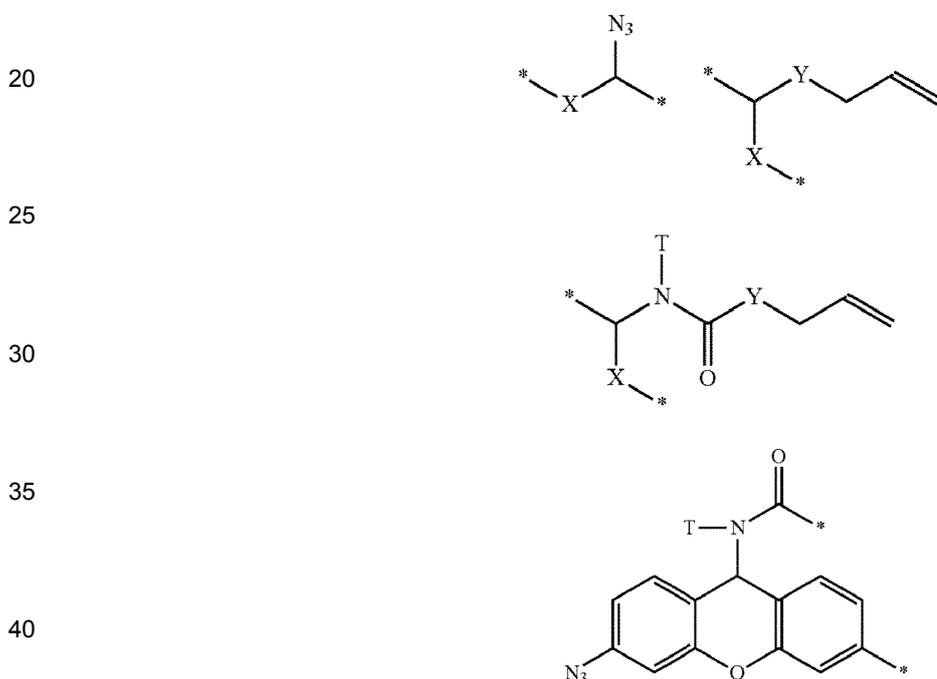
La conjugación del compuesto colorante con el sustrato puede llevarse a cabo a través del grupo carboxilo COX, que puede convertirse en una amida o un éster.

Los colorantes de acuerdo con la presente descripción pueden incluir un grupo de enlace reactivo en una de las posiciones sustituyentes para la unión covalente del colorante a otra molécula. Los grupos de enlace reactivos son fracciones capaces de formar un enlace (por ejemplo, un enlace covalente o no covalente). En una modalidad particular, el enlazador puede ser un enlazador escindible. El uso del término "enlazador escindible" no pretende implicar que se requiere el enlazador completo para eliminarse. El sitio de escisión puede situarse en una posición en el enlazador que hace que parte del enlazador permanezca adherido al colorante y/o a la fracción de sustrato después del clivaje. Los enlazadores escindibles

5 pueden ser, por medio de ejemplo no limitante, enlazadores escindibles electrofílicamente, enlazadores escindibles enzimáticamente, enlazadores escindibles nucleofílicamente, enlazadores fotoescindibles, escindibles bajo condiciones reductoras (por ejemplo, enlazadores que contienen disulfuro o azida), condiciones oxidativas, escindibles mediante el uso de enlaces de seguridad-captura y escindibles por mecanismos de eliminación. El uso de un enlazador escindible para unir el compuesto colorante a una fracción de sustrato proporciona la opción de eliminar el marcador, por ejemplo, después de la detección, lo que evita de ese modo cualquier señal interferente en las etapas posteriores.

10 Se pueden encontrar grupos de enlaces útiles en el número de la publicación PCT WO2004/018493 ejemplos de los cuales incluyen enlazadores que pueden escindirse mediante el uso de fosfinas solubles en agua o catalizadores de metales de transición solubles en agua formados a partir de un metal de transición y ligandos al menos parcialmente solubles en agua. En solución acuosa, los últimos forman complejos de metales de transición al menos parcialmente solubles en agua. Dichos enlazadores escindibles pueden usarse para conectar bases de nucleótidos a marcadores tales como los colorantes expuestos en la presente descripción.

15 Los enlazadores particulares se pueden encontrar en el número de la publicación PCT WO2004/018493 tales como los que incluyen las fracciones de la fórmula:



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

50 En modalidades particulares, la longitud del enlazador entre un colorante fluorescente (fluoróforo) y una base guanina puede alterarse, por ejemplo, mediante la introducción de un grupo espaciador de polietilenglicol, de esta manera aumenta la intensidad de esta fluorescencia en comparación con el mismo fluoróforo unido a la base de guanina a través de otros enlaces conocidos en la técnica. Los enlazadores ilustrativos y sus propiedades se exponen en el número de solicitud de la patente GB 0517097.2, publicado como WO07020457. El diseño de los enlazadores, y especialmente de su mayor longitud, puede permitir mejoras en el brillo de los fluoróforos unidos a las bases guanina de los nucleótidos de guanosina cuando se incorporan a polinucleótidos como el ADN. Por lo tanto, cuando el colorante es para uso en cualquier método de análisis que emplea detección de un marcador de colorante fluorescente unido a un nucleótido que contiene guanina, puede ser ventajoso usar un enlazador que tenga un grupo espaciador de fórmula - ((CH₂)₂O)_n- en donde n es un entero entre 2 y 50, por ejemplo, como se describe en el documento WO07020457.

60 La presente descripción proporciona además conjugados de nucleósidos y nucleótidos marcados con uno o más de los colorantes expuestos en la presente descripción (nucleótidos modificados). Los nucleósidos y nucleótidos marcados son útiles para marcar polinucleótidos formados por síntesis enzimática, tales como, por medio 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 muesca y similares.

65 Los nucleósidos y nucleótidos pueden marcarse en sitios en el azúcar o la 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 la 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 del ARN, uracilo (U). El átomo C-1 de la desoxirribosa está unido al N-1 de una pirimidina o al N-9 de una purina. Un nucleótido es además 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 son usualmente mono, di o trifosfatos.

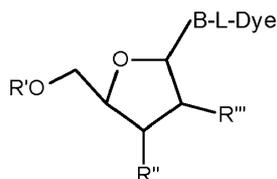
Un "nucleósido" es estructuralmente similar a un nucleótido, pero carece de las fracciones fosfato. Un ejemplo de un análogo de nucleósido sería uno en el que se une el marcador a la base y no hay ningún grupo fosfato unido a la molécula de azúcar.

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

Un colorante se puede unir a cualquier posición en una base de nucleótidos, por ejemplo, a través de un enlazador. En modalidades 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 7-desaza purina. Como se describió anteriormente, se puede usar un grupo de enlace para unir covalentemente un colorante al nucleósido o nucleótido.

En modalidades particulares, el nucleósido o nucleótido marcado puede ser incorporable enzimáticamente y extensible enzimáticamente. En consecuencia, una fracción enlazadora puede tener una longitud suficiente para conectar el nucleótido al compuesto de manera que el compuesto no interfiera significativamente con la unión y el reconocimiento general del nucleótido por una enzima de replicación de ácido nucleico. Por lo tanto, el enlazador puede comprender además una unidad espaciadora. El espaciador separa, por ejemplo, la base de nucleótidos de un sitio de escisión o marcador.

Los nucleósidos o nucleótidos marcados con colorantes de la descripción pueden tener la fórmula:



Donde Dye es un compuesto colorante de acuerdo con la presente descripción, B es una nucleobase, tal como, por ejemplo, uracilo, timina, citosina, adenina, guanina y similares, y L es un grupo de enlace opcional que puede o no estar presente. R' puede ser H, monofosfato, difosfato, trifosfato, tiosfosfato, un análogo de éster de fosfato, -O- unido a un grupo que contiene fósforo reactivo u -O- protegido por un grupo de bloqueo. R'' puede ser H, OH, una fosforamidita o un grupo de bloqueo 3'OH y R''' es H u OH.

Cuando R'' es fosforamidita, R' es un grupo protector de hidroxilo escindible con ácido que permite el subsecuente acoplamiento de monómeros bajo condiciones de síntesis automatizadas.

En una modalidad particular, el grupo de bloqueo está separado e independiente del compuesto colorante, es decir, no está directamente unido a él. En una modalidad alternativa, el colorante puede comprender todo o parte del grupo de bloqueo 3'OH. Por lo tanto, R'' puede ser un grupo de bloqueo 3'OH que puede comprender o no un compuesto colorante divulgado en la presente descripción.

En aún otra modalidad alternativa, no hay un grupo de bloqueo sobre el carbono 3' del azúcar pentosa y el colorante (o construcción de colorante y enlazador) unido a la base, por ejemplo, puede 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 colorante se una o no a la posición 3' del azúcar.

Todavía en otra modalidad alternativa más, el grupo de bloqueo está presente en el carbono 2' o 4' del 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 detener la extensión cuando se incorpora un nucleótido modificado. Si el efecto de bloqueo es reversible, por ejemplo, por medio de ejemplo no limitante, al cambiar las condiciones químicas o al eliminar un bloqueo químico, la extensión se puede detener en ciertos puntos y luego permitir que continúe.

En otra modalidad particular, un grupo de bloqueo 3'OH comprenderá fracciones descritas en el documento WO2004/018497. Por ejemplo, el grupo de bloqueo puede ser azidometilo (CH_2N_3) o alilo.

En una modalidad particular, un enlazador (entre colorante y nucleótido) y un grupo de bloqueo están presentes y son fracciones separadas. En modalidades particulares, el enlazador y el grupo de bloqueo son ambos escindibles bajo condiciones sustancialmente similares. Por lo tanto, los procesos de desprotección y desbloqueo pueden ser más eficiente, ya que sólo un único tratamiento se requerirá para eliminar el compuesto colorante y el bloqueo. Sin embargo, en algunas modalidades, un enlazador y un grupo de bloqueo no necesitan ser escindibles bajo condiciones similares, sino que son escindibles individualmente bajo condiciones distintas.

Esta descripción también abarca polinucleótidos que incorporan compuestos colorantes. Dichos polinucleótidos pueden ser ADN o ARN comprendidos respectivamente de desoxirribonucleótidos o ribonucleótidos unidos en un enlace fosfodiéster. Los polinucleótidos de acuerdo con la descripción pueden comprender nucleótidos de origen natural, nucleótidos de origen no natural (o modificados) distintos de los nucleótidos modificados de la descripción o cualquier combinación de los mismos, en combinación con al menos un nucleótido modificado (por ejemplo, marcado con un compuesto colorante) expuesto en la presente descripción. Los polinucleótidos de acuerdo con la descripción también pueden incluir enlaces de cadena principal no naturales y/o modificaciones químicas no nucleotídicas. También se contemplan estructuras quiméricas compuestas de mezclas de ribonucleótidos y desoxirribonucleótidos que comprenden al menos un nucleótido modificado de acuerdo con la descripción.

Los nucleótidos modificados (o nucleósidos) que comprenden un compuesto colorante de acuerdo con la presente descripción pueden usarse en cualquier método de análisis, tales como métodos que incluyen la detección de un marcador fluorescente unido a un nucleótido o nucleósido, ya sea solo o incorporado dentro o asociado con una estructura molecular más grande o conjugado. En este contexto, el término "incorporado en un polinucleótido" puede significar que el fosfato 5' está unido en un 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 expuesto en la presente descripción puede o no unirse en un enlace fosfodiéster al fosfato 5' de un nucleótido adicional (modificado o no modificado). Por lo tanto, en una modalidad no limitante, la descripción proporciona un método para detectar un nucleótido modificado incorporado en un polinucleótido que comprende: (a) incorporar al menos un nucleótido modificado de la descripción en un polinucleótido y (b) detectar el(los) nucleótido(s) modificado(s) incorporado(s) en el polinucleótido por la detección de la señal fluorescente del compuesto colorante unido a dicho(s) nucleótido(s) modificado(s).

Este método puede incluir: una etapa sintética (a) en la que uno o más nucleótidos modificados de acuerdo con la descripción se incorporan a un polinucleótido y una etapa de detección (b) en la que se detectan uno o más nucleótido(s) modificado(s) incorporado(s) en el polinucleótido al detectar o medir cuantitativamente su fluorescencia.

En una modalidad de la presente descripción, al menos un nucleótido modificado se incorpora en un polinucleótido en una etapa sintética por la acción de una enzima polimerasa.

Sin embargo, pueden usarse otros métodos para unir nucleótidos modificados a polinucleótidos, como por ejemplo síntesis química de oligonucleótidos o ligación de oligonucleótidos marcados a oligonucleótidos no marcados. Para ello, el término "incorporar", cuando se usa en referencia a un nucleótido y polinucleótido, puede abarcar la síntesis de polinucleótidos por métodos químicos, así como también métodos enzimáticos.

En una modalidad específica, se lleva a cabo una etapa sintética y puede comprender opcionalmente incubar una cadena de polinucleótidos molde con una mezcla de reacción que comprende nucleótidos modificados marcados fluorescentemente de la descripción. 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 recocado a la cadena de polinucleótido molde y un grupo fosfato 5' en el nucleótido modificado. Por lo tanto, una etapa sintética puede incluir la formación de una cadena de polinucleótidos según lo dirigido por el emparejamiento de bases complementarias de nucleótidos a una cadena molde.

En todas las modalidades del método, la etapa de detección puede llevarse a cabo mientras que se hibrida la cadena de polinucleótidos en la que los nucleótidos modificados se incorporan a una cadena molde, o después de una etapa de desnaturalización en la que las dos cadenas se separan. Se pueden incluir etapas adicionales, por ejemplo, etapas de reacción química o enzimática o etapas de purificación, entre una etapa sintética y una etapa de detección. En particular, la cadena diana que incorpora el(los) nucleótido(s) modificado(s) puede aislarse o purificarse y luego procesarse adicionalmente o usarse en un análisis subsecuente. A manera de ejemplo, los polinucleótidos diana marcados con nucleótido(s) modificado(s) en una etapa sintética pueden usarse subsecuente como sondas o cebadores

marcados. En otras modalidades, el producto de una etapa sintética expuesto en la presente descripción puede estar sujeto a etapas de reacción adicionales y, si se desea, el producto de estas etapas subsecuentes puede purificarse o aislarse.

5 Las condiciones adecuadas para una etapa sintética serán bien conocidas por aquellos familiarizados con las técnicas estándar de biología molecular. En una modalidad, una etapa sintética puede ser análoga a una reacción de extensión de cebador estándar mediante el uso de precursores de nucleótidos, incluidos los nucleótidos modificados expuestos en la presente descripción, para formar una cadena diana extendida complementaria a la cadena molde en presencia de una enzima polimerasa adecuada. En otras modalidades, una etapa sintética puede formar parte de una reacción de
10 amplificación que produce un producto de amplificación bicatenario marcado comprendido por cadenas complementarias híbridas derivadas de la copia de las cadenas de polinucleótidos diana y molde. Otras etapas sintéticas ilustrativas incluyen traducción de muesca, polimerización por desplazamiento de cadena, marcado de ADN cebado aleatorio, 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 expuestos en la presente descripción. Puede usarse una variedad de
15 polimerasas de origen natural o modificadas. A manera de ejemplo, puede usarse una polimerasa termoestable para una reacción sintética que se lleva a cabo mediante el uso de condiciones de termociclado, mientras que una polimerasa termoestable puede no ser deseable para reacciones de extensión de cebador isotérmicas. Las polimerasas termoestables adecuadas que son capaces de incorporar los nucleótidos modificados de acuerdo con la descripción incluyen aquellas descritas en los documentos WO 2005/024010 o WO06120433. En las reacciones sintéticas que se
20 llevan a cabo a temperaturas más bajas, tales como 37 °C, las enzimas polimerasas no necesitan necesariamente ser polimerasas termoestables, por lo tanto, la elección de la polimerasa dependerá de una serie de factores como la temperatura de reacción, el pH, la actividad de desplazamiento de cadena y similares.

En modalidades específicas no limitantes, la descripción abarca métodos de secuenciación del ácido nucleico, resecuenciación, secuenciación del genoma completo, puntuación única de polimorfismo de nucleótidos, o cualquier otra
25 aplicación que implique la detección del nucleótido o nucleósido modificado marcado con colorantes expuestos en la presente descripción cuando se incorporan en un polinucleótido. Cualquier otra variedad de solicitudes que se benefician del uso de polinucleótidos marcados con los nucleótidos modificados que comprenden colorantes fluorescentes puede usar nucleótidos o nucleósidos modificados marcados con colorantes expuestos en la presente descripción.

30 En una modalidad particular, la descripción proporciona el uso de nucleótidos modificados que comprenden compuestos colorantes de acuerdo con la descripción 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 polinucleotídica en crecimiento en la dirección 5' a 3' mediante el uso de una polimerasa o ligasa con el fin de
35 formar una cadena polinucleotídica extendida complementaria al ácido nucleico molde para ser secuenciado. La identidad de la base presente en uno o más de los nucleótidos añadidos se puede determinar en una etapa de detección o "formación de imágenes". La identidad de la base añadida puede determinarse después de cada etapa de incorporación de nucleótidos. La secuencia del molde se puede entonces inferir mediante el uso de reglas convencionales de emparejamiento de bases de Watson-Crick. El uso de los nucleótidos modificados marcados con colorantes expuestos
40 en la presente descripción para la determinación de la identidad de una sola base puede ser útil, por ejemplo, en la puntuación de polimorfismos de un solo nucleótido, y tales reacciones de extensión de una sola base están dentro del alcance de esta descripción.

45 En una modalidad de la presente descripción, la secuencia de un polinucleótido molde se determina al detectar la incorporación de uno o más nucleótidos en una cadena emergente complementaria al polinucleótido molde que se secuenciará mediante la detección de marcador(es) fluorescente(s) unido(s) al nucleótido(s) incorporado(s). La secuenciación del polinucleótido molde puede cebarse con un cebador adecuado (o prepararse como una construcción de horquilla que contendrá el cebador como parte de la horquilla), y la cadena emergente se extiende paso a paso
50 mediante la adición de nucleótidos al extremo 3' del cebador en una reacción catalizada por la polimerasa.

En modalidades particulares, cada uno de los diferentes nucleótidos trifosfatos (A, T, G y C) puede marcarse con un fluoróforo único y también comprende un grupo de bloqueo en la posición 3' para evitar la polimerización descontrolada. Alternativamente, uno de los cuatro nucleótidos puede estar sin marcar (oscuro). La enzima polimerasa incorpora un nucleótido en la cadena emergente complementaria al polinucleótido molde, y el grupo de bloqueo impide la incorporación
55 adicional de nucleótidos. Cualquier nucleótido no incorporado se puede eliminar por lavado y la señal fluorescente de cada nucleótido incorporado se puede "leer" ópticamente por medios adecuados, como un dispositivo acoplado por carga que usa excitación láser y filtros de emisión adecuados. El grupo de bloqueo-3' y los compuestos colorantes fluorescentes se pueden eliminar (desproteger) (simultáneamente o secuencialmente) para exponer la cadena emergente para la incorporación adicional de nucleótidos. Típicamente, la identidad del nucleótido incorporado se determinará después de
60 cada etapa de incorporación, pero esto no es estrictamente esencial. De manera similar, el documento de la Patente Estadounidense. No. 5,302,509 describe un método para secuenciar polinucleótidos inmovilizados sobre un soporte sólido.

65 El método, como se ejemplificó anteriormente, utiliza la incorporación de nucleótidos A, G, C y T bloqueados en 3' marcados fluorescentemente en una cadena en crecimiento complementaria al polinucleótido inmovilizado, en presencia de ADN polimerasa. La polimerasa incorpora una base complementaria al polinucleótido diana, pero el grupo de bloqueo-

3' impide incorporación adicional. El marcaje del nucleótido incorporado puede entonces ser determinado y el grupo de bloqueo se elimina por escisión química para permitir que se produzca la polimerización adicional. El molde de ácido nucleico que se secuenciará en una reacción de secuenciación por síntesis puede ser cualquier polinucleótido que se desee secuenciar. El molde de ácido nucleico para una reacción de secuenciación comprenderá típicamente una región bicatenaria 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 del molde a secuenciar sobresaldrá de este grupo hidroxilo 3' libre en la cadena complementaria. La región sobresaliente del molde a secuenciar puede ser monocatenaria pero puede ser bicatenaria, dado que "se presenta una muesca" en la cadena complementaria a la cadena molde que se secuenciará para proporcionar un grupo OH 3' libre para el inicio de la reacción de secuenciación. En tales modalidades, la secuenciación puede proceder por desplazamiento de cadena. En ciertas modalidades, se puede agregar un cebador que lleva el grupo hidroxilo 3' libre como un componente separado (por ejemplo, un oligonucleótido corto) que se hibrida con una región monocatenaria del molde a secuenciar. Alternativamente, el cebador y la cadena molde a secuenciar pueden cada una formar parte de una cadena de ácido nucleico parcialmente autocomplementario capaz de formar un dúplex intramolecular, como por ejemplo una estructura de bucle en horquilla. Los polinucleótidos en horquilla y los métodos por los cuales se pueden unir a soportes sólidos se describen en los documentos publicados de solicitud internacional de números WO0157248 y WO2005/047301. Los nucleótidos se pueden agregar sucesivamente a un cebador en crecimiento, lo que resulta en la síntesis de una cadena de polinucleótidos en la dirección 5' a 3'. La naturaleza de la base que se ha agregado puede determinarse, particularmente pero no necesariamente después de cada adición de nucleótidos, lo que proporciona por lo tanto información de secuencia para el molde de ácido nucleico. Por lo tanto, se incorpora un nucleótido en una cadena de ácido nucleico (o polinucleótido) mediante la unión del nucleótido al grupo 3' hidroxilo libre de la cadena de ácido nucleico mediante la formación de un enlace fosfodiéster con el grupo fosfato 5' del nucleótido.

El molde de ácido nucleico a secuenciar puede ser ADN o ARN, o incluso una molécula híbrida compuesta de desoxinucleótidos y ribonucleótidos. El molde de ácido nucleico puede comprender nucleótidos de origen natural y/o no natural y enlaces de cadena principal naturales o no naturales, dado que estos no impiden la copia del molde en la reacción de secuenciación.

En ciertas modalidades, el molde de ácido nucleico que se secuenciará puede unirse a un soporte sólido mediante cualquier método de enlace adecuado conocido en la técnica, por ejemplo, mediante unión covalente. En ciertas modalidades, los polinucleótidos molde pueden unirse directamente a un soporte sólido (por ejemplo, un soporte basado en sílice). Sin embargo, en otras modalidades de la descripción, la superficie del soporte sólido puede modificarse de alguna manera para permitir la unión covalente directa de los polinucleótidos molde, o para inmovilizar los polinucleótidos molde a través de una multicapa de hidrogel o polielectrolito, que puede e sí mismo no estar unido covalentemente al soporte sólido.

Las matrices en las que los polinucleótidos se han unido directamente a soportes a base de sílice son aquellas que se describen, por ejemplo, en el documento WO0006770, en donde los polinucleótidos se inmovilizan en un soporte de vidrio por reacción entre un grupo epóxido colgante sobre el vidrio con un grupo amino interno en el polinucleótido. Además, los polinucleótidos se pueden unir a un soporte sólido por reacción de un nucleófilo a base de azufre con el soporte sólido, por ejemplo, como se describe en el documento WO2005/047301. Un ejemplo todavía adicional de polinucleótidos molde con soporte sólido es cuando los polinucleótidos molde se unen al hidrogel soportado sobre soportes sólidos a base de sílice u otros soportes sólidos, por ejemplo, como se describe en los documentos WO00/31148, WO01/01143, WO02/12566, W003/014392, Patente Estadounidense. No. 6,465,178 y WO00/53812.

Una superficie particular a la cual pueden inmovilizarse los polinucleótidos molde 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 molde de ADN se pueden unir a perlas o micropartículas, por ejemplo como se describe en la Patente Estadounidense. No. 6,172,218. La unión a perlas o micropartículas puede ser útil para las aplicaciones de secuenciación. Se pueden preparar bibliotecas de perlas donde cada perla contiene diferentes secuencias de ADN. Bibliotecas y métodos ilustrativos para su creación se describen en Nature. 437, 376-380 (2005); Science. 309, 5741, 1728-1732 (2005). La secuenciación de matrices de tales perlas mediante el uso de nucleótidos expuestos en la presente descripción está dentro del alcance de la descripción.

Los moldes que se van a secuenciar pueden formar parte de una "matriz" sobre un soporte sólido, en cuyo caso la matriz puede adoptar cualquier forma conveniente. Por lo tanto, el método de la descripción es aplicable a todos los tipos de matrices de alta densidad, que incluye las matrices de una sola molécula, las matrices agrupadas y las matrices de perlas. Los nucleótidos modificados marcados con compuestos colorantes de la presente descripción pueden usarse para secuenciar moldes sobre esencialmente cualquier tipo de matriz, que incluyen, pero no se limitan a las formadas por inmovilización de moléculas de ácido nucleico sobre un soporte sólido.

Sin embargo, los nucleótidos modificados marcados con compuestos colorantes de la descripción son particularmente ventajosos en el contexto de la secuenciación de matrices agrupadas. En matrices agrupadas, distintas regiones en la matriz (a menudo denominadas sitios o características) comprenden múltiples moléculas de molde de polinucleótidos. Generalmente, las moléculas de polinucleótidos múltiples no se pueden resolver individualmente por medios ópticos y, en cambio, se detectan como un conjunto. En dependencia de cómo se forma la matriz, cada sitio en la matriz puede

comprender múltiples copias de una molécula de polinucleótido individual (por ejemplo, el sitio es homogéneo para una especie de ácido nucleico de cadena sencilla o doble particular) o incluso copias múltiples de un pequeño número de diferentes moléculas de polinucleótidos (por ejemplo, copias múltiples de dos especies diferentes de ácido nucleico). Se pueden producir matrices agrupadas de moléculas de ácidos nucleicos mediante el uso de técnicas generalmente conocidas en la técnica. A manera de ejemplo, los documentos WO 98/44151 y WO00/18957, describen métodos de amplificación de ácidos nucleicos en donde tanto el molde como los productos de amplificación permanecen inmovilizados sobre un soporte sólido con el fin de formar matrices compuestas de grupos o "colonias" de moléculas de ácidos nucleicos inmovilizado. Las moléculas de ácidos nucleicos presentes en las matrices agrupadas preparadas de acuerdo con estos métodos son moldes adecuados para la secuenciación mediante el uso de los nucleótidos modificados marcados con los compuestos colorantes de la descripción.

Los nucleótidos modificados marcados con compuestos colorantes de la presente descripción también son útiles en la secuenciación de moldes en matrices de moléculas individuales. El término "matriz de molécula única" o "SMA", como se usa en la presente descripción, se refiere a una población de moléculas de polinucleótidos, distribuidas (o agrupadas) sobre un soporte sólido, en donde la separación de cualquier polinucleótido individual de todas las demás de la población es tal que es posible resolver individualmente las moléculas de polinucleótidos individuales. Las moléculas de ácidos nucleicos diana inmovilizadas sobre la superficie del soporte sólido pueden por lo tanto poder resolverse por medios ópticos en algunas modalidades. Esto significa que una o más señales distintas, cada una en representación de un polinucleótido, ocurrirá dentro del área resoluble del dispositivo de imagen particular usado.

Se puede lograr la detección de una sola molécula en donde el espacio entre las moléculas de polinucleótidos adyacentes en una matriz es de al menos 100 nm, más particularmente al menos 250 nm, aún más particularmente al menos 300 nm, incluso más particularmente al menos 350 nm. Por lo tanto, cada molécula es individualmente resoluble y detectable como un punto fluorescente de molécula única, y la fluorescencia de dicho punto fluorescente de molécula única también exhibe fotoblanqueamiento en una sola etapa.

Los términos "resuelto individualmente" y "resolución individual" se usan en la presente descripción para especificar que, cuando se visualiza, es posible distinguir una molécula en la matriz de sus moléculas vecinas. La separación entre las moléculas individuales sobre la matriz se determinará, en parte, por la técnica particular usada para resolver las moléculas individuales. Las características generales de las matrices de moléculas individuales se entenderán por referencia a las solicitudes publicadas WO00/06770 y WO 01/57248. Aunque un uso de los nucleótidos modificados de la descripción es en reacciones de secuenciación por síntesis, la utilidad de los nucleótidos modificados no se limita a tales métodos. De hecho, los nucleótidos pueden usarse 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 colorantes de la descripción pueden usarse en protocolos de secuenciación fluorescente automatizada, particularmente secuenciación fluorescente del ciclo colorante-terminador basado en el método de secuenciación de terminación de cadena de Sanger y colaboradores. Tales métodos generalmente usan enzimas y secuenciación de ciclo para incorporar didesoxinucleótidos marcados con fluorescencia en una reacción de secuenciación de extensión del cebador. Los llamados métodos de secuenciación de Sanger y los protocolos relacionados (tipo Sanger) usan la terminación de cadena aleatoria con didesoxinucleótidos marcados.

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

Los nucleótidos modificados marcados con compuestos colorantes de la presente descripción que incorporan grupos de bloqueo 3', serán reconocidos, también pueden ser útiles en los métodos de Sanger y protocolos relacionados, ya que el mismo efecto logrado mediante el uso de los nucleótidos didesoxi modificados puede lograrse mediante el uso de los nucleótidos modificados que tienen grupos de bloqueo 3'-OH: ambos evitan la incorporación de nucleótidos posteriores. Cuando se usan los nucleótidos de acuerdo con la presente descripción, y que tienen un grupo de grupo de bloqueo 3', se usarán en los métodos de secuenciación tipo Sanger, se apreciará que los compuestos colorantes o marcadores detectables unidos a los nucleótidos no necesitan estar conectados a través de enlazadores escindibles, ya que en cada instancia en la que se incorpora un nucleótido marcado de la descripción, no existen nucleótidos que necesitan ser incorporados subsecuetemente y por lo tanto, el marcador no necesita ser retirado del nucleótido.

La presente descripción también proporciona kits que incluyen nucleósidos y/o nucleótidos modificados marcados con colorantes. Tales kits generalmente incluirán al menos un nucleótido o nucleósido modificado marcado con un colorante expuesto en la presente descripción junto con al menos un componente adicional. Los componentes adicionales pueden ser uno o más de los componentes identificados en un método expuesto anteriormente o en la sección de Ejemplos a continuación. Algunos ejemplos no limitantes de componentes que pueden combinarse en un kit de la presente descripción se exponen a continuación.

En una modalidad particular, un kit puede incluir al menos un nucleótido o nucleósido modificado marcado con un colorante expuesto en la presente descripción junto con nucleótidos o nucleósidos modificados o no modificados. Por ejemplo, los nucleótidos modificados marcados con colorantes de acuerdo con la descripción pueden suministrarse en combinación

con nucleótidos nativos o sin marcar, y/o con nucleótidos marcados con fluorescencia o cualquier combinación de los mismos. En consecuencia, los kits pueden comprender nucleótidos modificados marcados con colorantes de acuerdo con la descripción y nucleótidos modificados marcados con otros, por ejemplo, compuestos colorantes de la técnica anterior. Las combinaciones de nucleótidos pueden proporcionarse 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 colorante, los diferentes nucleótidos pueden estar marcados con diferentes compuestos colorantes, o uno puede ser oscuro, sin compuestos colorantes. Cuando los diferentes nucleótidos están marcados con diferentes compuestos colorantes, una característica de los kits es que dichos compuestos colorantes son colorantes fluorescentes espectralmente distinguible. Como se usa en la presente descripción, el término "colorantes fluorescentes espectralmente distinguible" se refiere a los colorantes fluorescentes que emiten energía fluorescente a longitudes de onda que pueden distinguirse mediante un equipo de detección fluorescente (por ejemplo, una plataforma de secuenciación de ADN con base capilar comercial) cuando están presentes dos o más de estos colorantes en una muestra. Cuando dos nucleótidos modificados marcados con compuestos colorantes fluorescentes se suministran en forma de kit, una característica de algunas modalidades es que los colorantes fluorescentes espectralmente distinguible pueden excitarse a la misma longitud de onda, tal como, por ejemplo, por el mismo láser. Cuando cuatro nucleótidos modificados marcados con compuestos colorantes fluorescentes se suministran en forma de kit, una característica de algunas modalidades es que dos de los colorantes fluorescentes espectralmente distinguible pueden excitarse en una longitud de onda y los otros dos colorantes espectralmente distinguibles pueden excitarse 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 modalidad, un kit incluye un nucleótido modificado marcado con un compuesto de la presente descripción y un segundo nucleótido modificado marcado con un segundo colorante en donde los colorantes tienen una diferencia en la absorbancia máxima de al menos 10 nm, particularmente de 20 nm a 50 nm. Más particularmente, los dos compuestos colorantes tienen desplazamientos de Stokes de entre 15-40 nm, donde el "desplazamiento de Stokes" es la distancia entre las longitudes de onda del pico absorción y del pico de emisión.

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

En una modalidad alternativa, los kits de la descripción pueden contener nucleótidos donde la misma base se marca con dos compuestos diferentes. Un primer nucleótido puede marcarse con un compuesto de la descripción. Un segundo nucleótido puede marcarse con un compuesto espectralmente distinto, por ejemplo, un colorante 'verde' que absorbe a menos de 600 nm. Un tercer nucleótido puede marcarse como una mezcla del compuesto de la descripción 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 pueden marcarse como 'verde', 'rojo', 'rojo/verde' y oscuro. Para simplificar la instrumentación adicional, se pueden marcar cuatro nucleótidos con dos colorantes excitados con un solo láser y, por lo tanto, el marcado de los nucleótidos 1-4 puede ser 'rojo 1', 'rojo 2' 'rojo 1/rojo 2', y oscuro o 'verde 1', 'verde 2' verde 1/verde 2 'y oscuro.

Los nucleótidos pueden contener dos colorantes de la presente descripción. Los colorantes donde Ra_1 o Ra_2 es un anillo aromático adicional fusionado a carbonos adyacentes del anillo de indol que absorben a una longitud de onda más larga que donde los colorantes no tienen la conjugación aromática adicional. Un kit puede contener dos o más nucleótidos marcados con colorantes de la descripción. Un kit puede contener un nucleótido marcado con un compuesto de la descripción donde 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 descripción donde uno o ambos Ra_1 y Ra_2 es un anillo adicional fusionado a un átomo de carbono adyacente. Los kits pueden contener un nucleótido adicional donde el nucleótido se marca con un colorante que absorbe en la región de 520 nm a 560 nm. Los kits pueden contener además un nucleótido no marcado.

Los nucleótidos pueden contener dos colorantes de la presente descripción. Los colorantes donde p es uno absorben a una longitud de onda más corta que donde p es dos. Un kit puede contener dos o más nucleótidos marcados con colorantes de la descripción. Un kit puede contener un nucleótido marcado con un compuesto de la descripción donde p es uno, y un nucleótido marcado con un compuesto de la descripción donde p es dos. Los kits pueden contener un nucleótido adicional donde el nucleótido se marca con un tercer marcador. Los kits pueden contener además un nucleótido no marcado o un nucleótido marcado con un cuarto marcador.

Aunque los kits se ejemplifican anteriormente con respecto a las configuraciones que tienen diferentes nucleótidos que están marcados con diferentes compuestos colorantes, se entenderá que los kits pueden incluir 2, 3, 4 o más nucleótidos diferentes que tienen el mismo compuesto colorante.

5 En modalidades 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 tales kits pueden incluir tampones y similares. Los nucleótidos modificados marcados con colorantes de acuerdo con la descripción, y otros componentes de nucleótidos que incluyen mezclas de diferentes nucleótidos, pueden proporcionarse en el kit en una forma concentrada para diluirse antes de su uso. En tales modalidades, también se puede incluir un tampón de dilución adecuado.

10 Nuevamente, uno o más de los componentes identificados en un método expuesto en la presente descripción pueden incluirse en un kit de la presente descripción.

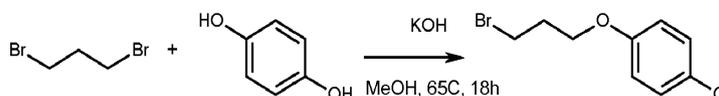
Se nota que, como se usa en esta descripción y en las reivindicaciones adjuntas, las formas singulares "un", "una", y "el/la" incluyen los referentes plurales a menos que expresa e inequívocamente se limite a una referente.

15

Detalles experimentales

Preparación de 4-(3-bromopropoxi)fenol

20



25

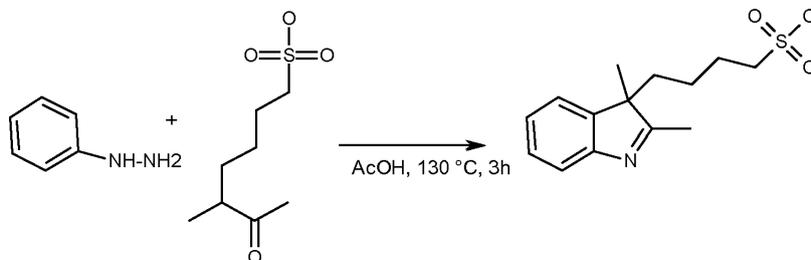
30

A una solución de hidroquinona (3 g; 27,2 mmol) en metanol (30 mL) se le añadió bajo N₂ hidróxido de potasio (2 g; 35,4 mmol) y 1,3-dibromopropano (8,3 mL; 81,7 mmol). La reacción se agitó toda la noche a 65 °C. Luego, los volátiles se eliminaron al vacío y se añadió éter dietílico (30 mL). La capa orgánica se lavó con una solución acuosa saturada de cloruro de amonio y agua (30 mL cada una). La capa orgánica se secó con sulfato de magnesio, se filtró y los volátiles se eliminaron al vacío. La suspensión beige resultante se purificó por cromatografía instantánea sobre gel de sílice (eluyente: éter de petróleo/acetato de etilo 70/30) al proporcionar 2,6 g de una suspensión blanca. Rendimiento: 41 %. RMN de protón (CDCl₃) : 6,87-6,73 (4H, m, H-Ar); 5,36 (1H, s, OH); 4,04 (2H, t, J = 6,2 Hz); 3,60 (2H, J = 6,2 Hz); 2,28 (2H, q, J = 6,2 Hz, H-2). ¹³C (CDCl₃) : 152,8; 149,6; 116,1; 116,0; 66,7; 34,4; 30,2.

35

Preparación de ácido 4-(2,3-dimetil-3H-indol-3-il)butano-1-sulfónico

40



45

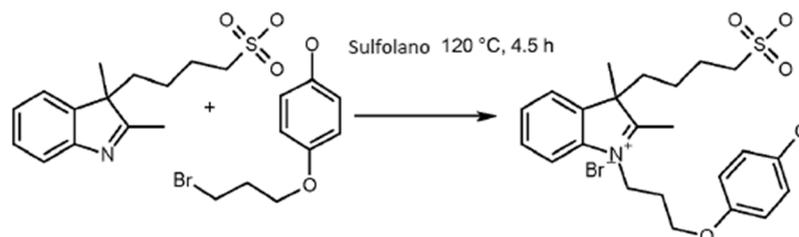
50

La fenilhidrazina (8,6 mmol; 850 μL) y el ácido 5-metil-6-oxoheptanosulfónico (7,2 mmol; 1,5 g) en ácido acético glacial (19 mL) se agitaron a 130 °C durante 3 horas. La mezcla de reacción se enfrió a temperatura ambiente y el disolvente se eliminó al vacío. El residuo se purificó mediante cromatografía instantánea sobre gel de sílice (eluyente: diclorometano/metanol 80/20) que proporciona 1,23 g de espuma de color amarillo pálido. Rendimiento: 61 %. EM (DUIS): E⁻: 280 (M-1); RMN de protón (D₂O): 7,42 (1H, d, J = 7,8Hz, H-4); 7,27 (1H, t, J = 7,8Hz, H-5); 7,24 (1H, d, J = 7,8Hz, H-7); 7,17 (1H, t, J = 7,8Hz, H-6); 2,59 (2H, dd, J = 7,6; 8,9Hz, H-11); 2,21-2,13 (1H, m); 1,88-1,73 (2H, m, H-8); 1,52-1,43 (2H, m, H-10); 1,13 (3H, s, CH₃ on C-3); 0,70-0,63 (1H, m, H-9a); 0,51-0,44 (1H, m, H-9b). ¹³C (D₂O): 190,8; 152,3; 143,7; 127,8; 125,7; 122,2; 118,6; 57,9; 50,5; 48,9; 35,6; 24,0; 22,5; 21,8; 14,8.

55

Preparación de 1-(3-(4-hidroxifenoxi)propil)-2,3-dimetil-3-(4-sulfobutil)-3H-indol-1-ilo

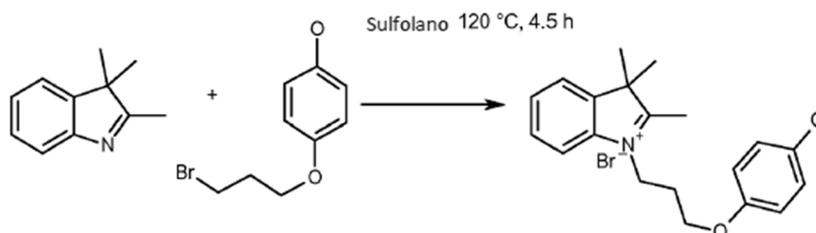
60



65

4-(3-bromopropoxi)fenol (1,90 mmol, 438 mg) y de ácido 4-(2,3-dimetil-3H-indol-3-il)butano-1-sulfónico (1,58 mmol, 504 mg) en sulfolano (2 mL) se agitaron a 120 °C durante 4,5 h. La mezcla de reacción se enfrió a temperatura ambiente y se purificó directamente por cromatografía instantánea sobre gel de sílice (eluyente: diclorometano/metanol 65/35) que proporciona 329 mg de un aceite gris pálido. Rendimiento: 40 %. EM (DUIS): E⁻: 430 (M-1), E⁺: 432 (M+1); RMN de protón (TFA): 7,75-7,65 (4H, m, H-4/5/6/7); 6,98 (2H, d, J = 8,9 Hz, Har on phenyl); 6,85 (2H, d, J = 8,9 Hz, Har on phenyl); 4,82 (2H, bt, J = 6,6 Hz, H-12); 4,24-4,13 (2H, m, H-14); 3,69 (1H, s); 3,20-3,06 (2H, m, H-11); 2,91 (3H, s, CH₃ on C-2); 2,63-2,52 (2H, m, H-13); 2,41-2,75 (2H, m, H-8); 1,90-1,83 (2H, m, H-10); 1,66 (3H, s, CH₃ on C-3); 1,13-0,93 (2H, m, H-9).

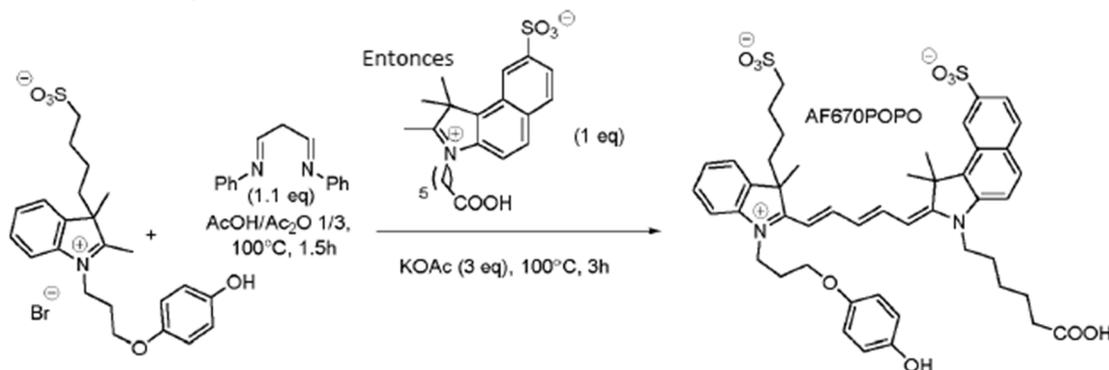
Preparación de 1-(3-(4-hidroxifenoxi) propil)-2,3,3-trimetil-3H-indol-1-io



Se agitaron el 4-(3-bromopropoxi)fenol (1,40 mmol, 324 mg) y trimetilindolenina (0,80 mmol, 129 µL) en sulfolano (1 mL) a 120 °C durante 3 horas. La mezcla de reacción se enfrió a temperatura ambiente y se purificó directamente por cromatografía instantánea sobre gel de sílice (eluyente: diclorometano/metanol 90/10) que proporciona 260 mg de un aceite gris pálido. Rendimiento: 84 %. RMN de protón (MeOD): 7,06 (1H, d, J = 7,2 Hz, H-4); 7,00 (1H, bt, J = 7,8 Hz, H-5); 6,77-6,64 (5 H, m, H-6 y H-on fenil orto y meta); 6,58 (1H, d, J = 7,8 Hz, H-7); 3,90 (2H, t, J = 5,8 Hz, H-8); 3,71 (2H, J = 6,5 Hz, H-10); 2,04 (2H, bq, J = 6,1 Hz, H-9); 1,30 (6H, s, 2xCH₃ en C-3). ¹³C (MeOD): 162,7; 153,6; 152,3; 147,2; 138,6; 128,6; 122,7; 119,5; 116,8; 116,6; 116,5; 103,3; 66,8; 45,1; 39,7; 30,6; 27,5.

Preparación de AF670POPO

Esquema de síntesis del producto:



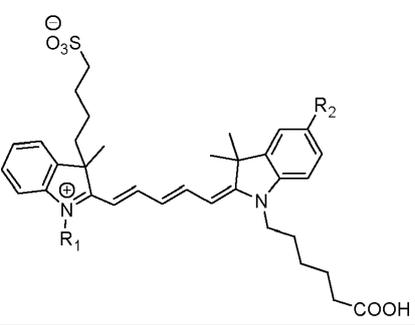
Procedimiento de síntesis del colorante:

El ácido 4-(2,3-dimetil-3H-indol-3-il) butano-1-sulfónico (204 mg, 400 µmol) y el monoclorhidrato de malonaldehído Bis (fenilimina) (114 mg, 440 µmol) se agitaron en una mezcla de ácido acético anhídrido/ácido acético (2/0,6 mL) a 100 °C durante 1,5 h. Se añadieron acetato de potasio (118 mg; 1,20 mmol) y la sal de benzindolio sulfonada X (193 mg; 400 µmol). La mezcla de reacción se agitó a 100 °C por 3 h. Luego se enfrió a temperatura ambiente y se añadió éter dietílico (40 mL). Un sólido azul oscuro se estrelló. Se filtró, se lavó con éter dietílico extra (40 mL) y se disolvió en agua (25 mL) que contiene ~30 % de acetonitrilo. La solución azul oscuro resultante se filtró y se purificó por HPLC preparativa. Las fracciones principales coloreadas de azul (absorción máxima ~ 665 nm) se recogieron y los disolventes se eliminaron al vacío. EM (DUIS): E⁻: 869 (M-1); 434,1 (M-2/2).

Los colorantes DX-DX se prepararon de manera similar mediante el uso de materiales de partida apropiados

En la tabla 1, las propiedades espectrales de algunos colorantes preparados de esta manera (soluciones en agua, DO ~ 0,2) en comparación con parámetros similares del colorante análogo estructural conocido Std (Patente: P6240WO)

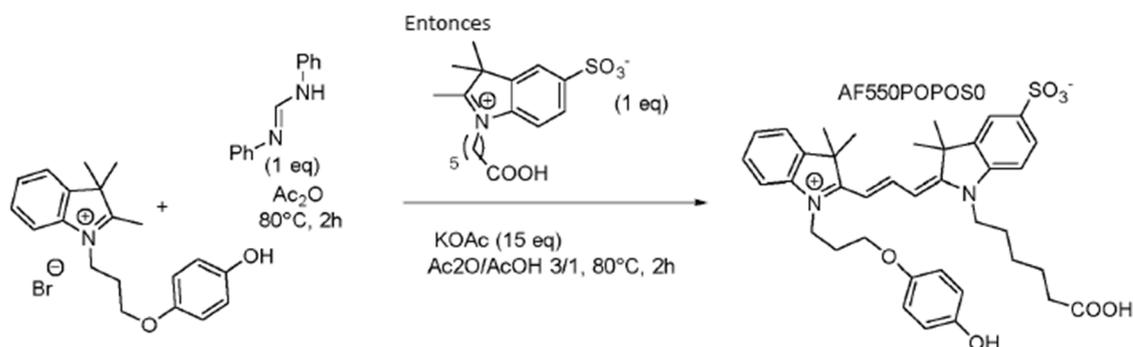
Tabla 1



Colorante	R ₁	R ₂	Máxima Absorbancia (nm)	Máxima Fluorescencia (nm)
Std (NR650C5)	Ph	SO ₂ NH ₂	650	673
AF670POPO	-(CH ₂) ₃ -O-C ₆ H ₄ -pOH	-Benz-SO ₃ ⁻	666	684
AF670POP	-(CH ₂) ₃ -O-C ₆ H ₅	-Benz-SO ₃	665	684
AF670PPO	-(CH ₂) ₃ -C ₆ H ₄ -pOH	-Benz-SO ₃	665	683

Preparación de AF550POPOS0

Esquema de síntesis del producto:



Procedimiento de síntesis del colorante:

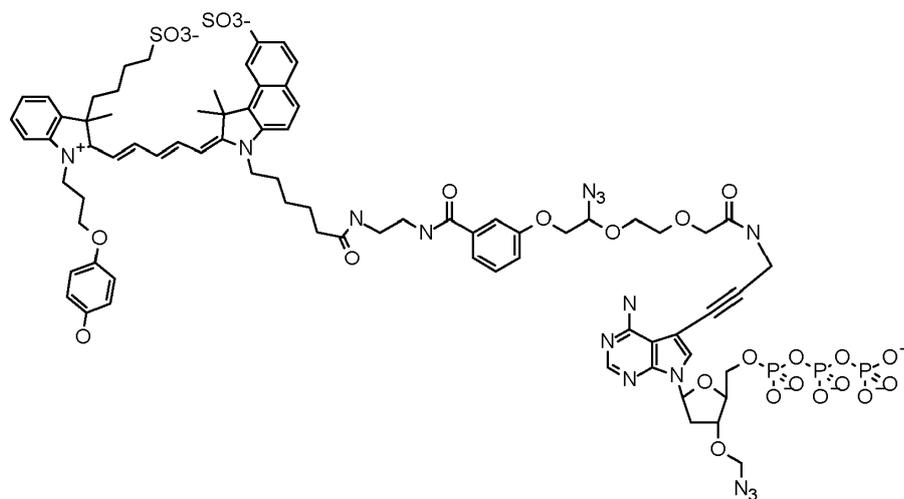
1-(3-(4-hidroxifenoxi) propil)-2,3,3-trimetil-3H-indol-1-ilo (150 mg, 385 μ mol) y *N,N'*-difenilformamida (74 mg, 385 μ mol) se agitó en ácido acético anhídrido (2 mL) a 80 °C durante 2 h. Se añadieron ácido acético (0,5 mL), acetato de potasio (567 mg, 5,78 mmol) y la sal de indolio sulfonada X (166 mg; 385 μ mol). La mezcla de reacción se agitó a 80 °C por 2,5 h. Luego los volátiles se eliminaron bajo presión reducida y se añadió éter dietílico (20 mL). Un sólido rojo oscuro se estrelló. Se filtró, se lavó con éter dietílico extra (20 mL) y se disolvió en agua (15 mL) que contiene ~5 % de acetonitrilo. La solución roja oscura resultante se filtró y se purificó por HPLC preparativa. Las fracciones principales coloreadas de rojo (absorción máxima ~ 550 nm) se recogieron y los disolventes se eliminaron al vacío. EM (DUIS): No acetilado E⁻: 671,0 (M-1); E⁺: 673,3 (M+1); 774,4 (Et₃N⁺ sal).

En la tabla 2, las propiedades espectrales de algunos colorantes preparados de esta manera (soluciones en agua, DO ~ 0,2) en comparación con parámetros similares de los colorantes análogos estructurales conocidos (Patente: P6240WO)

Tabla 2

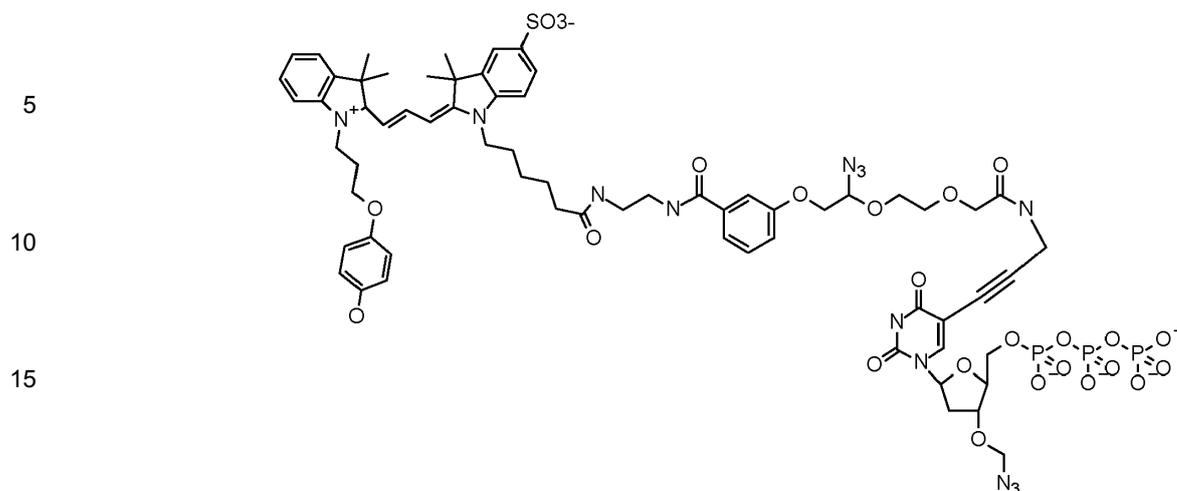
Colorante	R ₁	Máxima Absorbancia (nm)	Máxima Fluorescencia (nm)
Std 1 (NR550S0)	Ph	550	576
Std 2 (NR545S0)	Me	545	559
AF550POPOS0	-(CH ₂) ₃ -O-C ₆ H ₄ -pOH	548	563
AF550POPS0	-(CH ₂) ₃ -O-C ₆ H ₅	547	566
AF550PPS0	-(CH ₂) ₃ -C ₆ H ₄ -pOH	548	567
AF550PPS0	-(CH ₂) ₃ -C ₆ H ₅	548	567

Conjugado de nucleótidos con colorante AF670POPO (FFA-AF670POPO)



Se añadieron DMA anhidro (3 mL), TSTU (18 mg; 58,8 μmol) y Base de Hunig (0,05 mL) a la muestra seca del colorante AF670POPO (60 mg; 53,5 μmol). Se desarrolló el éster activado de color azul. La mezcla de reacción se agitó a temperatura ambiente durante 30 minutos. De acuerdo con la TLC (20 % de H₂O en CH₃CN) la activación se completó. Una vez completada la activación, la solución de 5 pppA-LN3 como una sal de trietilamonio (53 mg; 56,1 μmol) en DMA/agua (0,5/0,3 mL) se añadió a la reacción. La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno por 18 h. El progreso del acoplamiento se verificó por TLC (20 % de H₂O en acetonitrilo). La mezcla de reacción se enfrió hasta ~4 °C con un baño de hielo, luego se añadió una solución 10 de TEAB 0,1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 minutos. La mezcla de reacción se aplicó a la columna con ~ 25 g de suspensión de resina DEAE Sephadex en solución de TEAB 0,05 M en agua y lavada con TEAB (gradiente de concentración de 0,1 M a 1 M). Las fracciones coloreadas se recogieron y se evaporaron, luego se coevaporaron de nuevo con agua para eliminar más TEAB y aspirar a sequedad. El residuo se redisolvió luego en TEAB 0,1 M. Esta solución se filtró a través de un filtro de jeringa de 0,2 μm de tamaño de poro en un frasco Corning 20. El producto se purificó por HPLC mediante el uso de una columna de fase reversa C18 con acetonitrilo-TEAB 0,1 M. Rendimiento del 74 % (basado en la densidad óptica de la solución mediante el uso del coeficiente de extinción estimado).

Conjugado de nucleótidos con colorante AF550POPOS0 (FFT-AF550POPOS0)



25

30

Se añadieron DMA anhidro (3 mL), TSTU (19 mg; 63,5 μ mol) y Base de Hunig (0,05 mL) a la muestra seca del colorante AF550POPOS0 (39 mg; 57,7 μ mol). Se desarrolló el éster activado de color rojo. La mezcla de reacción se agitó a temperatura ambiente durante 30 minutos. La solución de pppT-LN3 como una sal de trietilamonio (59 mg; 63,5 μ mol) en DMA/agua (0,5/0,3 mL) se añadió a la reacción. La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno por 18 h. El progreso del acoplamiento se verificó por TLC (20 % de H₂O en acetonitrilo). La mezcla de reacción se enfrió hasta \sim 4 °C con un baño de hielo, luego se añadió una solución 10 de TEAB 0,1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 minutos. La mezcla de reacción se aplicó a la columna con \sim 25 g de suspensión de resina DEAE Sephadex en solución de TEAB 0,05 M en agua y lavada con TEAB (gradiente de concentración de 0,1 M a 1 M). Las fracciones coloreadas se recogieron y se evaporaron, luego se coevaporaron de nuevo con agua para eliminar más TEAB y aspirar a sequedad. El residuo se redisolvió luego en TEAB 0,1 M. Esta solución se filtró a través de un filtro de jeringa de 0,2 μ m de tamaño de poro en un frasco Corning 20. El producto se purificó por HPLC mediante el uso de una columna de fase reversa C18 con acetonitrilo-TEAB 0,1 M. Rendimiento del 48 % (basado en la densidad óptica de la solución mediante el uso del coeficiente de extinción estimado).

35 Descripción de las figuras

40 **Figura 1** Los FFAs marcados con AF670POPO y sus análogos junto con ffa-NR650C5 se probaron en sistemas de secuenciación Illumina HiSeqX modificados, en los que aumentaba la energía de excitación de ambos láseres (rojo y verde) en aproximadamente el 33 % se desplegó en el ciclo 101 y 251 para la lectura 1 y 2 respectivamente (véase las señales en la figura 1).

45 **Figura 2** Inesperadamente, en el caso de ffa-NR650C5 (figura 2), se observó que las señales rojas no respondían proporcionalmente al cambio de energía del láser rojo. Consecuentemente, la forma de los diagramas de dispersión de separación de base se alteró en el ciclo 101. El círculo superior derecho está más cerca del círculo superior izquierdo, lo que significa que es más difícil distinguir entre los marcadores respectivos. Los fenómenos que ocurrieron indicaron que NR650C5 probablemente estaba fotosaturado o casi saturado a la densidad de la energía máxima usada en HiSeqX. Este cambio de diagrama de dispersión afecta negativamente la calidad de la secuencia. Curiosamente, con AF670POPO y sus análogos, se observó una mejora notable. Su señal se incrementó proporcionalmente con el cambio de densidad de la energía del láser. El diagrama de dispersión mantuvo sus separaciones y la forma del ciclo 100 al ciclo 101 cuando se desplegó la energía máxima (figura 2). El círculo superior derecho en la figura 2 está más alejado del círculo superior izquierdo que el círculo superior derecho en la figura 1. Por lo tanto, los marcadores están mejor separados en la figura 2.

55 **Figura 3** En los análogos de 550 nm, fue interesante observar que la presencia de hidroxilo en el anillo de fenilo tiene un impacto perjudicial en la intensidad de fluorescencia (figura 3). La fluorescencia de AF550POPOS0 sin el OH fenólico fue 3 veces más fuerte que el AF550POPOS0 que tiene el OH.

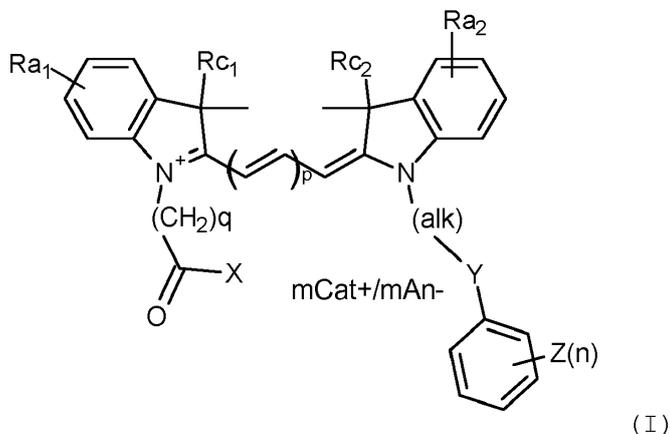
60 **Figura 4** Estos nuevos análogos ofrecieron oportunidades para modular y optimizar la distribución del diagrama de dispersión para las aplicaciones de secuenciación. Por ejemplo, cuando T se marcó con AF550POPOS0 y A con una mezcla de AF670POPO y NR550S0, se logró un buen diagrama de dispersión cuadrado (figura 4a). Adicionalmente, el reemplazo de NR550S0 con AF550POPOS0 para la A verde mostró un muy buen diagrama de dispersión cuadrado (figura 4b).

65 **Figura 5** En base a estas nuevas propiedades, ffa-LN3-AF550POPOS0 y ffa-LN3-AF670POPO se emplearon para secuenciar un molde humana en un HiSeqX modificado. Se logró una secuenciación de 2x150 pb de alta calidad tasa de error <1 % para la lectura 1 y tasa de error <1,5 % para la lectura 2 (las principales métricas de secuencia se

muestran en la figura 5). Los gráficos de dispersión obtenidos mediante el uso de los marcadores de la técnica anterior significaron que no se pudo completar una lectura de 2x150 pb.

REIVINDICACIONES

1. Un compuesto de la fórmula (I) o formas mesoméricas del mismo:



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

m es un entero de 0-3;

p es un entero de 1-2;

q es un entero de 1-5;

alk es una cadena de 1-5 átomos de carbono que contiene opcionalmente uno o más dobles o triples enlaces;

Y es O;

Z es OH;

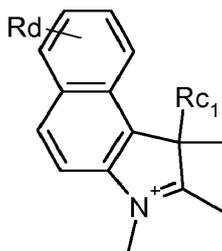
n es 1;

X es OH u O⁻ o un conjugado de éster o amida del mismo;

cada uno de Ra₁ y Ra₂ es independientemente H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; y

cada uno de Rc₁ y Rc₂ es independientemente alquilo o alquilo sustituido, en donde al menos uno de Ra₁ o Ra₂ es SO₃⁻ o Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente, el anillo adicional tiene un SO₃⁻, o Rc₁ o Rc₂ es un grupo de ácido alquilsulfónico.

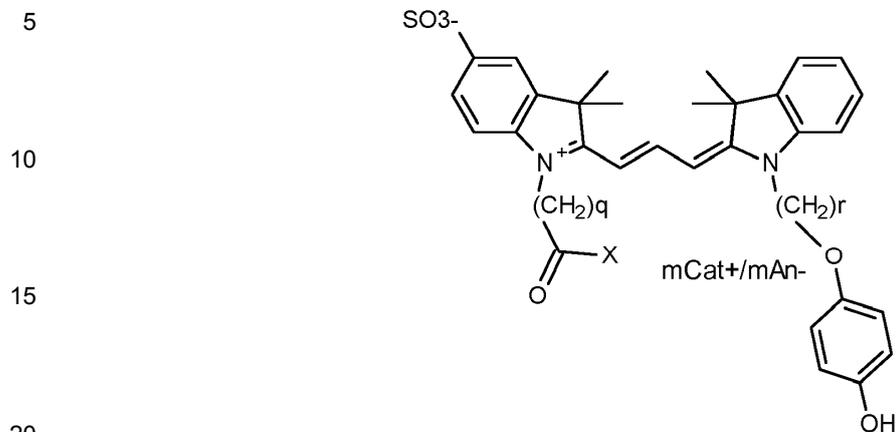
2. Un compuesto de acuerdo con la reivindicación 1 en donde (alk) es (CH₂)₃.
3. Un compuesto de acuerdo con la reivindicación 1 o la reivindicación 2 en donde Ra₁ es H o SO₃⁻.
4. Un compuesto de acuerdo con la reivindicación 1 o la reivindicación 2 en donde Ra₂ es H o SO₃⁻.
5. Un compuesto de acuerdo con una cualquiera de las reivindicaciones 1 a 4 en donde Ra₁ o Ra₂ es un anillo adicional fusionado a un átomo de carbono adyacente.
6. Un compuesto de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde Rc₁ o Rc₂ es metilo, etilo, propilo o - (CH₂)_tSO₃⁻ donde t es 1-6.
7. Un compuesto de acuerdo con cualquiera de las reivindicaciones anteriores en donde el grupo OH de Z está en la posición 4.
8. Un compuesto de acuerdo con cualquier reivindicación precedente en donde cualquier indol es parte de una estructura:



donde Rd puede ser H, alquilo, alquilo sustituido, arilo, arilo sustituido, halógeno, carboxi, sulfonamida o ácido sulfónico.

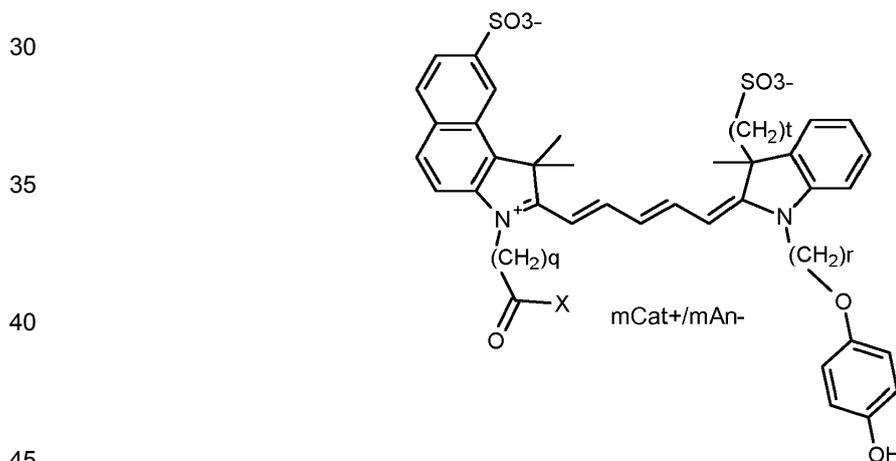
en donde Rc₁ es alquilo o alquilo sustituido.

9. Un compuesto de acuerdo con la reivindicación 1 la que se representa por la fórmula (VIIIa) o fórmula (VIId):



(VIIIa)

en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico o inorgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 r es un entero de 1-5; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo;

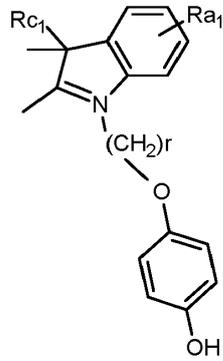


(VIId)

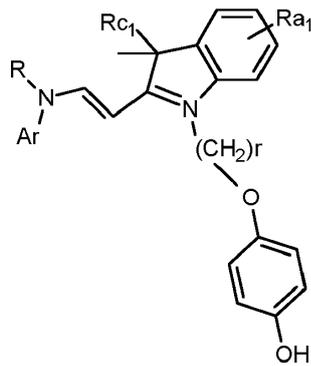
en donde mCat+ o mAn- es un contraión cargado positivamente/negativamente orgánico y
 m es un entero de 0-3;
 q es un entero de 1-5;
 r es un entero de 1-5;
 t es un entero de 1-6; y
 X es OH u O⁻ o un conjugado de éster o amida del mismo.

- 55 10. Un nucleótido u oligonucleótido marcado con un compuesto de acuerdo con las reivindicaciones 1-9.
11. Un nucleótido u oligonucleótido marcado de acuerdo con la reivindicación 10 en donde el marcador está unido a través de un enlace amida formado a partir la fracción C(=O)-X.
- 60 12. Un nucleótido u oligonucleótido marcado de acuerdo con las reivindicaciones 10 u 11, en donde el marcador está unido a la posición C5 de una base pirimidina o la posición C7 de una base de 7-desaza purina a través de una fracción enlazadora.
13. Un kit que comprende dos o más nucleótidos en donde al menos un nucleótido es un nucleótido marcado de acuerdo con las reivindicaciones 10-12.

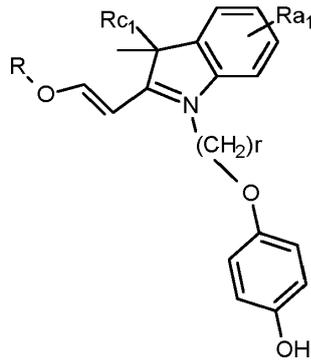
14. Uso de un nucleótido de acuerdo con cualquiera de las reivindicaciones 10-12, un oligonucleótido de acuerdo con las reivindicaciones 10-12 o un kit de acuerdo con la reivindicación 13 en la secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN o ensayos de unión a proteínas.
15. Un método para sintetizar un compuesto de acuerdo con las reivindicaciones 1-9 utilizando uno de los siguientes materiales de partida:



(X)

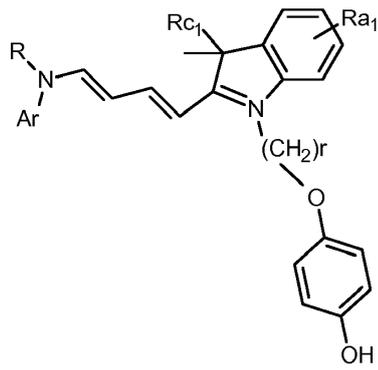


(X1)



(X2)

5

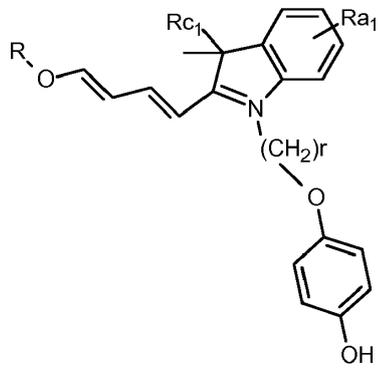


10

15

(X3)

20



25

30

35

(X4)

40

o una sal de los mismos en donde Ra₁ es H, SO₃⁻, sulfonamida, halógeno o un anillo adicional fusionado a un átomo de carbono adyacente; R_{c1} es alquilo o alquilo sustituido; Ar es un grupo aromático y R es un grupo alquilo.

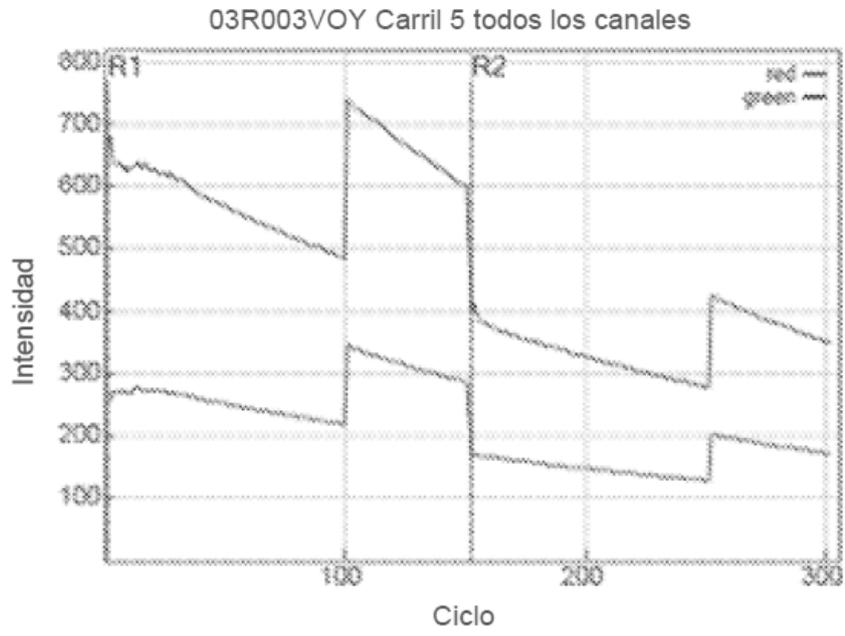


Figura 1

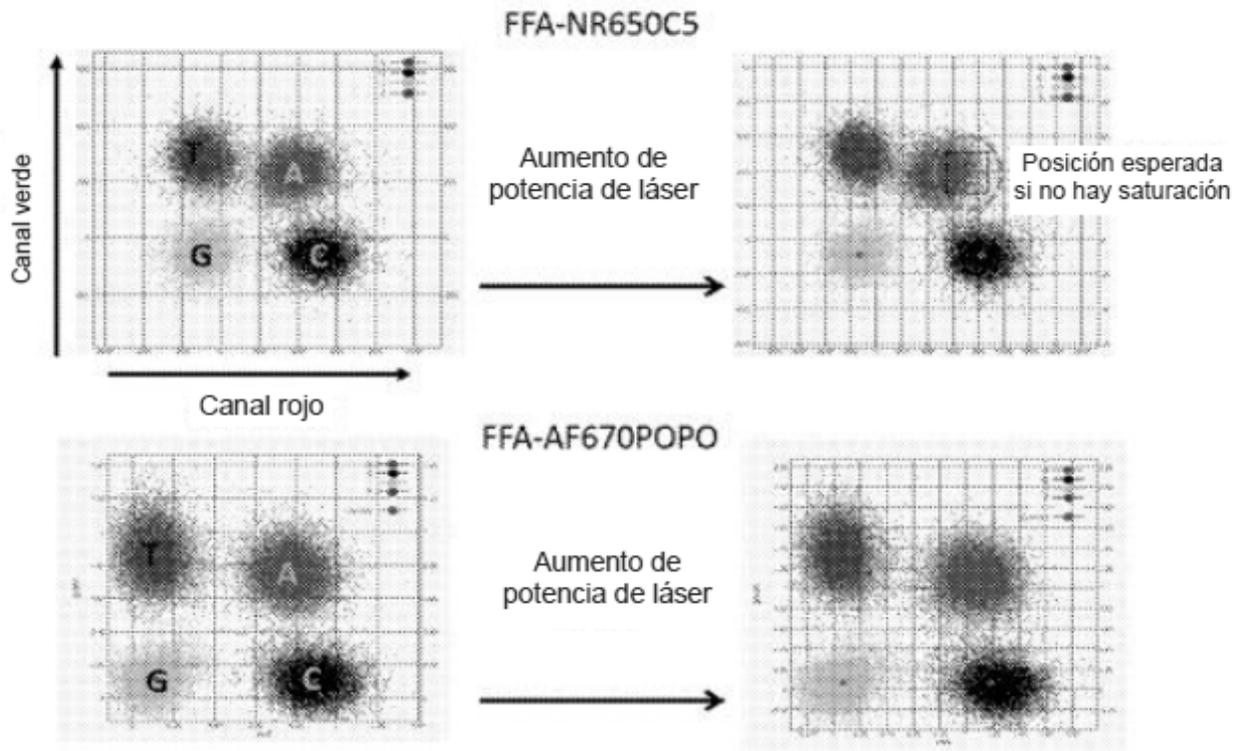


Figura 2

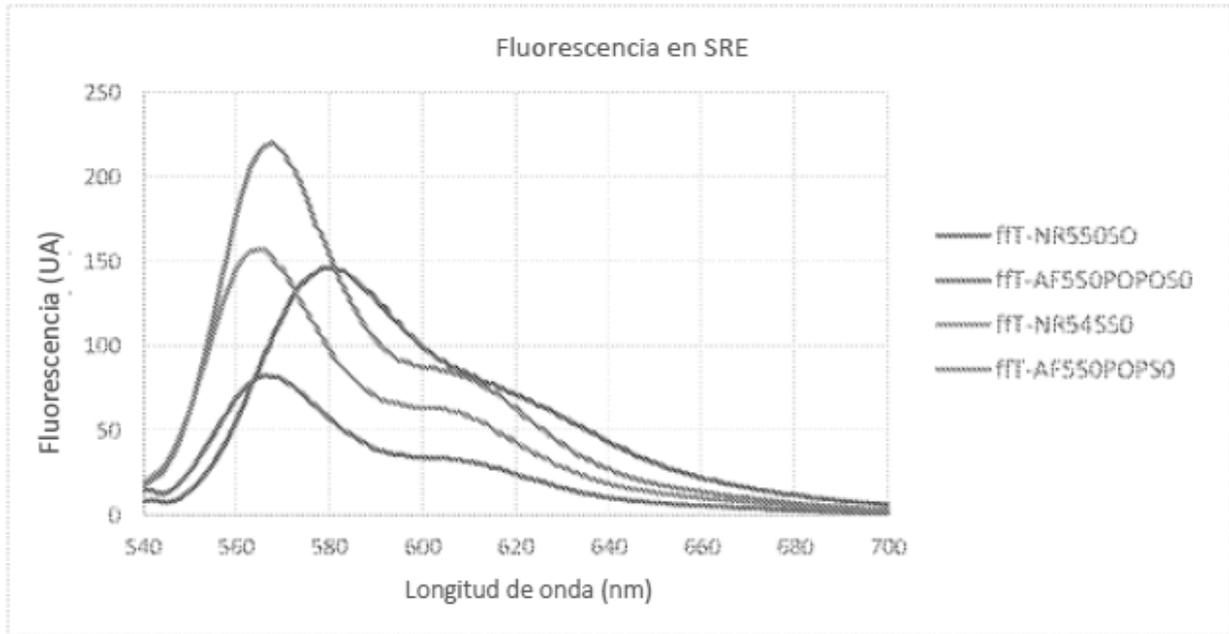


Figura 3

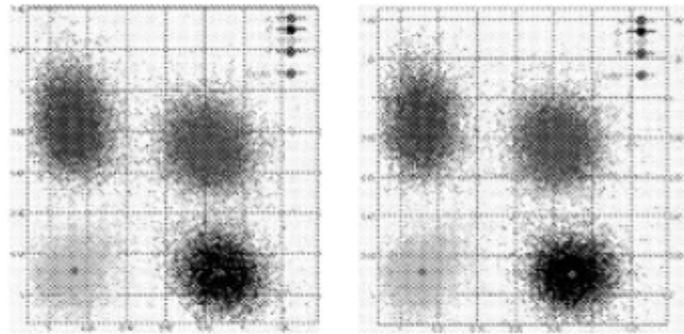


Figura 4

Carril	Grupos PF(%)	Fase/Prefase R1	% \geq Q30 R1	Alineado R1	Tasa de error (%) R1	Fase/Prefase R2	% \geq Q30 R2	Alineado R2	Tasa de error (%) R2
350 pb-Humana	72,32 \pm 3,79	0,151 / 0,132	88,7	0,92 \pm 1,46	0,92 \pm 0,16	0,141 / 0,140	81,5	93,09 \pm 1,04	1,46 \pm 0,19

Figura 5