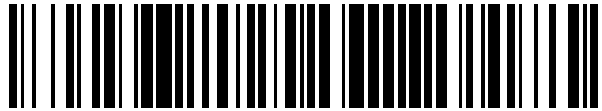


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 620 427**

51 Int. Cl.:

**C07D 311/82** (2006.01)

**C07H 21/00** (2006.01)

**C09B 11/24** (2006.01)

**G01N 33/58** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.03.2013 PCT/EP2013/054793**

87 Fecha y número de publicación internacional: **12.09.2014 WO2014135223**

96 Fecha de presentación y número de la solicitud europea: **08.03.2013 E 13709182 (3)**

97 Fecha y número de publicación de la concesión europea: **04.01.2017 EP 2964624**

54 Título: **Compuestos rodamina y su uso como marcadores fluorescentes**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**28.06.2017**

73 Titular/es:

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

72 Inventor/es:

**ROMANOV, NIKOLAI, NIKOLAEVICH y  
LIU, XIAOHAI**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

ES 2 620 427 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 rodamina y su uso como marcadores fluorescentes

5 La presente invención se refiere a compuestos rodamina novedosos y su uso como marcadores fluorescentes. Los compuestos se pueden utilizar como marcadores fluorescentes, particularmente para marcar nucleótidos en aplicaciones de secuencias de ácidos nucleicos.

Antecedentes

10

En esta solicitud hace referencia a diversas publicaciones y documentos de patentes con el fin de describir más completamente el estado de la técnica a la que pertenece esta invención. La divulgación de cada una de estas publicaciones y documentos se incorpora aquí por referencia.

15

La detección de ácidos nucleicos no radiactivos que utilizan marcadores fluorescentes es una tecnología importante en la biología molecular. Muchos procedimientos empleados en la tecnología de ADN recombinante se basan en gran medida previamente en el uso de nucleótidos o polinucleótidos marcados radioactivamente con, por ejemplo <sup>32</sup>P. 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 como su costo, vida útil limitada y más importante, consideraciones de seguridad. Eliminar la necesidad de marcadores radiactivos mejora la seguridad mientras reduce el impacto ambiental y los costos asociados con, por ejemplo, desechos de reactivos. Métodos aceptables para la detección fluorescente no radiactiva incluyen, por vía de ejemplos no limitantes, secuenciamiento de ADN automatizado, métodos de hibridación, detección en tiempo real de productos de reacción de cadena de polimerasa e inmunoensayos.

20

25

Para muchas aplicaciones es deseable emplear múltiples marcadores fluorescentes distinguibles espectralmente con el fin de alcanzar detección independiente de una pluralidad de analitos que se superponen espacialmente. En dichos métodos múltiples se puede reducir el número de recipientes de reacción simplificando protocolos experimentales y facilitando la producción de kits de reactivos específicos de aplicaciones. En el secuenciamiento de ADN automatizado multicolor, por ejemplo, la detección fluorescente múltiple permite el análisis de múltiples bases de nucleótidos en una única línea de electroforesis, aumentando por lo tanto el rendimiento sobre un color y reduciendo la incertidumbre asociada con las variaciones de movilidad electroforética interlínea.

30

35

Sin embargo, la detección fluorescente múltiple puede ser problemática y existe una serie de factores importantes que restringen la selección de marcadores fluorescentes. En primer lugar, es difícil encontrar compuestos de tinte cuyo espectro de emisión se resuelve espectralmente de forma adecuada. Adicionalmente, cuando se utilizan diversos tintes fluorescentes juntos, puede ser difícil la excitación simultánea debido a que las bandas de absorción de los tintes frecuentemente se separan ampliamente. Muchos métodos de excitación utilizan láseres de alta energía y por lo tanto el tinte puede tener suficiente fotoestabilidad para soportar dicha excitación láser. Una consideración final de importancia particular en los métodos de biología molecular es que los tintes fluorescentes deben ser compatibles con la química de reactivos utilizados tal como por ejemplo reactivos y disolventes sintética de ADN, reguladores, enzimas de polimerasa y enzimas ligasa.

40

45

A medida que avanza la tecnología de secuenciamiento se ha desarrollado la necesidad de compuestos de tintes fluorescentes, sus conjugados de ácidos nucleicos y grupos de tintes que satisfagan todas las restricciones anteriores y que sean particularmente compatibles con los métodos moleculares de alto rendimiento tal como el secuenciamiento de fase sólida y similares.

50

La solicitud WO2007135368 describe una clase de compuestos rodamina adecuados para uso como marcadores fluorescentes. Los compuestos descritos allí son adecuados para su uso en protocolos de secuenciamiento de ácidos nucleicos de fase sólida. Los avances en la tecnología y rendimiento de secuenciamiento de ácidos nucleicos de fase sólida han conducido a desarrollos y mejoras adicionales al diseño molecular de marcadores fluorescentes, particularmente en el contexto de la interacción entre reactivos fluorescentes y secuencias de ácidos nucleicos particulares.

55

La publicación European Journal of Organic Chemistry 2002, (7), 1149-1162 describe compuestos rodaminas similares con un ácido carboxílico y un amino sustituido como sustituyentes del anillo fenilo, que se utilizan como marcadores fluorescentes.

60

Las moléculas de tintes fluorescentes con propiedades fluorescentes mejoradas (tal como intensidad de fluorescencia, posición de máxima de fluorescencia y forma de la banda de fluorescencia) pueden mejorar la velocidad y precisión del secuenciamiento de ácido nucleico. La intensidad de señal de fluorescencia es especialmente importante cuando se hacen mediciones en agua basadas en reguladores biológicos y/o a temperaturas de fluorescencia mayores de la mayoría de tintes lo que es significativamente menor en dichas condiciones. Más aún, la naturaleza de la base en la que se une un tinte también afecta la fluorescencia máxima, la intensidad de fluorescencia y otras propiedades espectrales del tinte. Las interacciones específicas de secuencia entre el tinte fluorescente y la nucleobase se pueden adaptar mediante diseño específico de tintes fluorescentes. La optimización de la estructura de los tintes fluorescentes puede

65

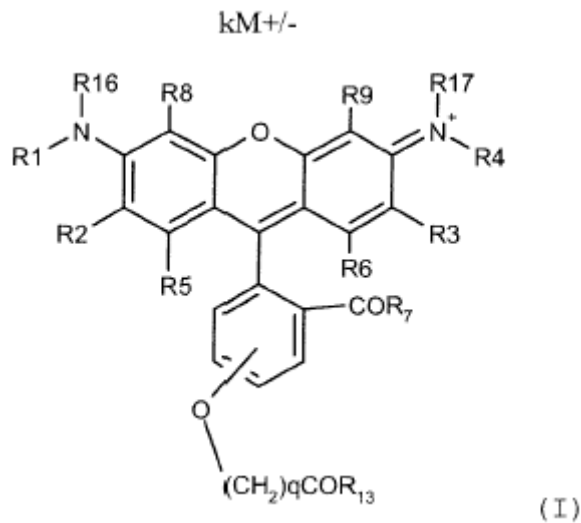
mejorar sus propiedades fluorescentes y también mejorar la eficiencia de la incorporación de nucleótidos, reducir el nivel de errores de secuenciación y reducir el uso de reactivos en, y por lo tanto los costos del secuenciación de ácidos nucleicos.

5 Se describen aquí construcciones de rodamina mejoradas y su uso como marcadores de biomoléculas, particularmente como marcadores para nucleótidos utilizados en secuenciación de ácidos nucleicos. Se pueden observar mejoras en mayores intensidades de fluorescencia de dichos tintes cuando se preparan como conjugados de biomoléculas y en la longitud y calidad de la lectura de secuenciación que se puede obtener utilizando las nuevas construcciones fluorescentes.

10

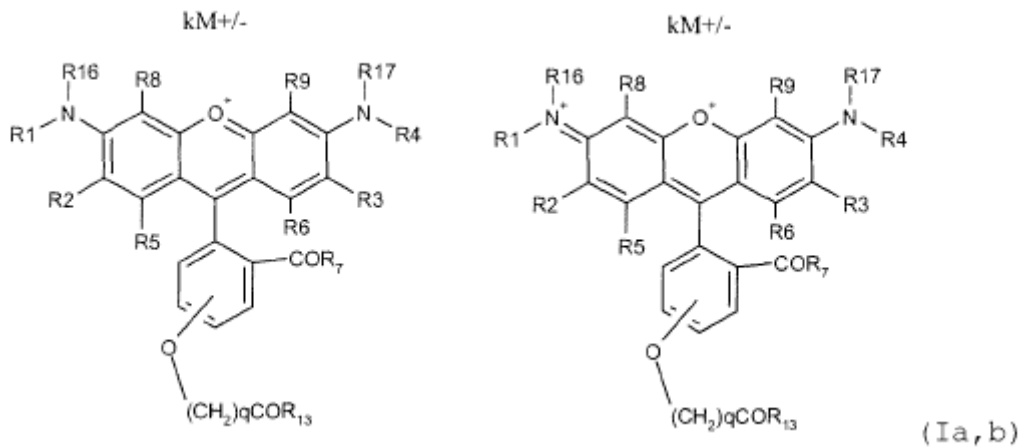
Resumen

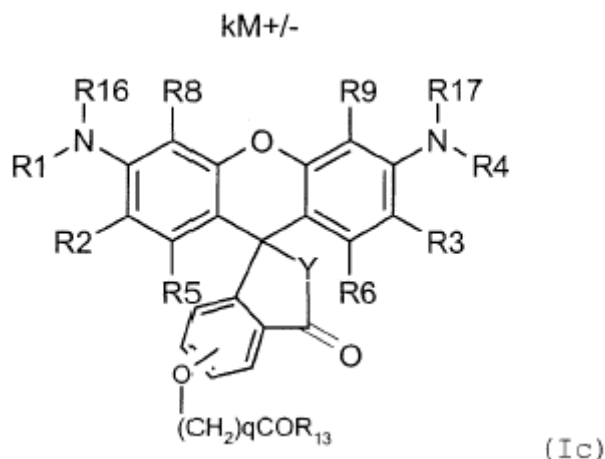
De acuerdo con un primer aspecto la invención proporciona compuestos tintes de rodamina de la fórmula (I) y mesómeros del mismo:



15

Los mesómeros de la invención pueden incluir las fórmulas representadas por 1a, 1b o forma cíclica 1c:





En las fórmulas I, Ia, Ib o Ic,

- 5 M<sup>+/-</sup> es un contraión común,  
 k es un entero desde 0 hasta 6,  
 q es un entero desde 1 hasta 6,  
 10 Y = O, NR<sub>11</sub>,  
 R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,  
 15 R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es una cadena de carbonos o heterosustituida que forma un anillo,  
 R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es una cadena de carbonos o heterosustituida que forma un anillo,  
 20 R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,  
 R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxilo- o alcoxi,  
 25 R<sub>8</sub> es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,  
 R<sub>9</sub> es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,  
 30 R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido,  
 R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido, y  
 35 R<sub>16</sub> y R<sub>17</sub> son independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido.

En otra realización los compuestos de la presente invención se pueden conjugar con una variedad de grupos funcionales de sustrato tal como, por ejemplo, nucleósidos, nucleótidos, polinucleótidos, polipéptidos, carbohidratos, ligandos, partículas y superficies.

De acuerdo con un aspecto adicional de la invención por lo tanto, se proporcionan compuestos tintes que comprenden grupos enlazadores para permitir, por ejemplo, unión covalente a dichos grupos funcionales de sustrato.

45 De acuerdo con un aspecto adicional la invención proporciona un compuesto de nucleósido o nucleótido definido por la fórmula: N-L-Dye, en la que N es un nucleótido, L es un grupo funcional enlazador opcional y Dye es un compuesto fluorescente de acuerdo con la presente invención.

50 En un aspecto adicional la invención incluye métodos de secuenciación utilizando los compuestos tintes de la presente invención.

De acuerdo con un aspecto adicional la invención también proporciona kits que comprenden compuestos tintes de la invención (libres o en forma conjugada) que se pueden utilizar en diversos ensayos inmunológicos, marcado de oligonucleótidos y ácidos nucleicos y para secuenciación de ADN mediante síntesis. En aún otro aspecto la invención proporciona kits que comprenden 'grupos' de tintes particularmente adecuados a ciclos de secuenciación mediante síntesis sobre una plataforma de instrumentos automatizada.

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

10 Descripción de Figuras

La figura 1 muestra la menor reducción de temperatura de los tintes como describe aquí en comparación con los tintes de la técnica anterior. Las intensidades de fluorescencia normalizadas de soluciones  $1.10^{-6}M$  de tintes (I-1) y (I-3) se compararon con tintes comercialmente disponibles Atto532 para la misma región espectral a diferente temperatura. Se mide la intensidad de los tintes en 20, 40 y 60°C. La figura 1 muestra la intensidad relativa de los tintes en cada temperatura. Los tintes comerciales Atto532 muestran una mayor pérdida de intensidad de fluorescencia a mayores temperaturas con relación a los tintes I-1 e I-3. La figura 1 demuestra que la fluorescencia de los nuevos tintes en soluciones a base de agua es menos variable con la temperatura.

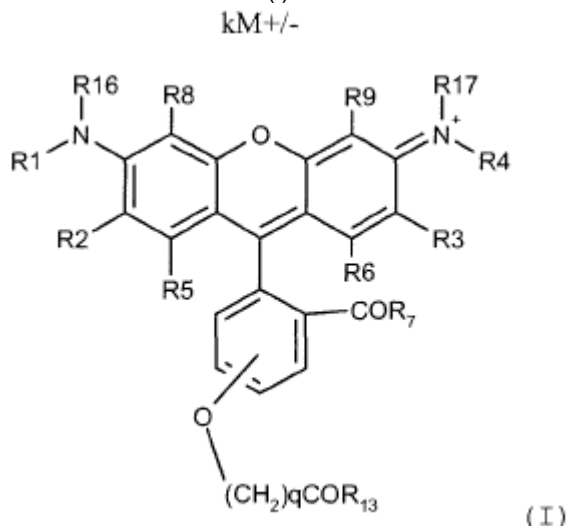
La figura 2 demuestra que la fluorescencia de conjugados de nucleobase basados en estos nuevos tintes en soluciones a base de agua es mayor que los tintes comercialmente disponibles cuando se excitan mediante luz a 532 nm. Se comparó el espectro de Fluorescencia Normalizado de  $1.10^{-6}M$  de conjugados de tinte-nucleobases (I-13) -T y (I-11) -T con análogos estructurales cuando se conjuga pppT con el tinte Atto532 disponible comercialmente. El tinte I-3 es más brillante que el Atto 532 en la misma concentración de nucleótidos. El tinte I-1 se desplaza en rojo en comparación con el tinte Atto532.

La figura 3 demuestra que la fluorescencia de nuevos tintes en soluciones a base de agua depende menos de la temperatura. Soluciones  $1.10^{-6}M$  de Intensidades de Fluorescencia Normalizada de tintes cuando se conjuga con nucleobase, T-(I-11) y T-(I-13), cuando se comparan con tinte Atto532 disponible comercialmente conjugado con la mismo T-nucleobase. Tanto los tintes I-1 como I-3 muestran mayor intensidad de fluorescencia a temperaturas elevadas en comparación con Atto-532.

La figura 4 demuestra mejor distinción señales de fluorescencia cuando se marca una nucleobase con un nuevo tinte de acuerdo con la invención (I-3) (fila 6) en comparación con fluoróforo estándar fijados cuando la misma nucleobase ha conjugado con el tinte Atto532 disponible comercialmente (control 1). La figura 4 muestra una gráfica de intensidad rojo versus intensidad verde en una serie de secuenciamiento de 4 colores Illumina. La mayor distancia entre los grupos de tintes señala menores oportunidades de llamadas fallidas, y por lo tanto aumenta la precisión del secuenciamiento. El aumento del brillo del tinte I-3 comparado con los tintes comerciales significa que se mejoran los datos de secuenciamiento.

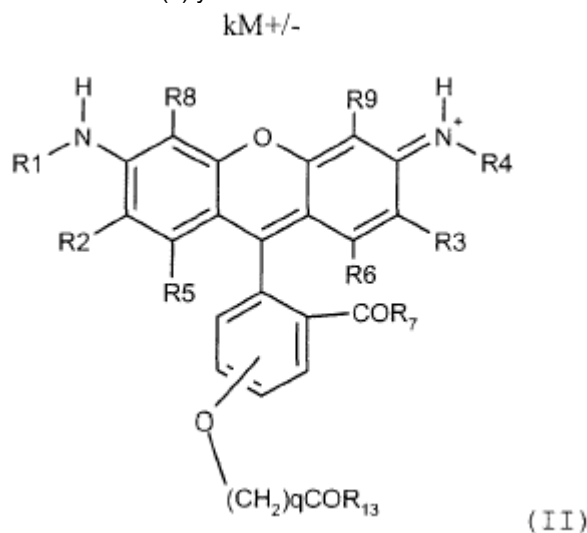
40 Descripción detallada

La invención se relaciona con compuestos tintes de rodamina novedosos particularmente adecuados para métodos de detección de fluorescencia y secuenciación mediante síntesis. De acuerdo con un primer aspecto la invención proporciona compuestos tintes de rodamina de la fórmula (I):



en la que

- M<sup>+/-</sup> es un contraión común,
- k es un entero desde 0 hasta 6,
- 5 q es un entero desde 1 hasta 6,
- R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,
- R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es un carbono o cadena heterosustituida que forma un anillo,
- 10 R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es un carbono o cadena heterosustituida que forma un anillo,
- 15 R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,
- R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxilo- o alcoxi,
- R<sub>8</sub> es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,
- 20 R<sub>9</sub> es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,
- 25 R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido,
- R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido, y
- 30 R<sub>16</sub> y R<sub>17</sub> son independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido.
- R<sub>16</sub> y R<sub>17</sub> puede ser independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido. El grupo alquilo se puede sustituir con SO<sub>3</sub><sup>-</sup>. En el que R<sub>16</sub> o R<sub>17</sub> es un alquilo sustituido con SO<sub>3</sub><sup>-</sup>, el grupo SO<sub>3</sub><sup>-</sup> se puede coordinar con un contraión, por ejemplo iones de metal o iones de amonio. R<sub>16</sub> y R<sub>17</sub> pueden ser H. Los compuestos de la invención por lo tanto incluyen compuestos de la fórmula (II) y mesómeros del mismo:
- 35



- en el que
- 40 M<sup>+/-</sup> es un contraión común,
- k es un entero desde 0 hasta 6,
- q es un entero desde 1 hasta 6,
- 45 R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,
- R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o

R<sub>5</sub> es un carbono o cadena heterosustituida que forma un anillo,

R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxil- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es un carbono o cadena heterosustituida que forma un anillo,

5

R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxil- o alcoxi,

10

R<sub>8</sub> es H, halógeno, grupo hidroxil- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,

R<sub>9</sub> es H, halógeno, grupo hidroxil- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,

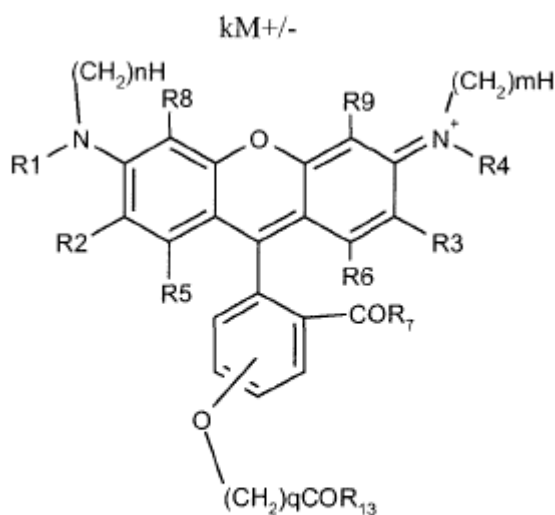
15

R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y

R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido.

20

Cuando R<sub>16</sub> y/o R<sub>17</sub> es alquilo no sustituido -(CH<sub>2</sub>)<sub>n</sub>H o -(CH<sub>2</sub>)<sub>m</sub>H, n y m pueden ser 1 a 6. Los compuestos de la invención por lo tanto incluyen compuestos de la fórmula (IIa):



25

en la que M<sup>-</sup> es un contraión común,

k, n, y m son independientemente enteros desde 0 hasta 6,

q es un entero desde 1 hasta 6,

30

R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxil- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es una cadena de carbonos o heterosustituida que forma un anillo,

35

R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxil- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es una cadena de carbonos o heterosustituida que forma un anillo,

R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

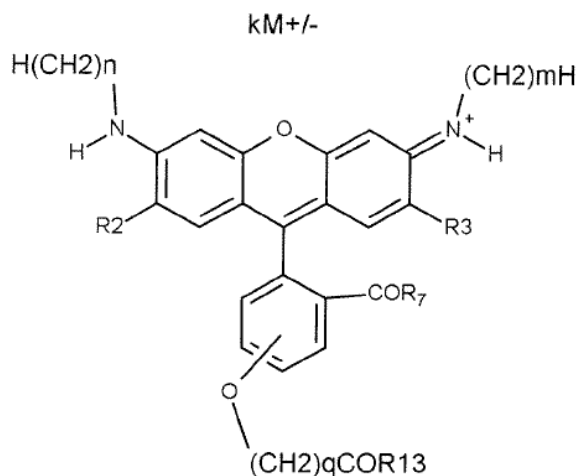
40

R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxil- o alcoxi,

R<sub>8</sub> es H, halógeno, grupo hidroxil- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,

- R<sub>9</sub> es H, halógeno, grupo hidroxilo o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,
- 5 R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y
- R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido.
- 10 n y m pueden ser los mismos o diferentes. n puede ser 1, 2, 3, 4, 5 o 6. m puede ser 1, 2, 3, 4, 5 o 6. El -(CH<sub>2</sub>)<sub>n</sub>-H o -(CH<sub>2</sub>)<sub>m</sub>-H puede ser grupos alquilo C<sub>1-6</sub>, por ejemplo metilo, etilo o propilo, y se puede sustituir opcionalmente.
- 15 q puede ser 1, 2, 3, 4, 5 o 6. Adicionalmente al enlazador (CH<sub>2</sub>)<sub>q</sub>, el enlazador puede contener sustituyentes en cualesquier átomos de carbono o heteroátomos adicionales. Por ejemplo el enlazador puede contener átomos de oxígeno adicionales en la forma de separadores del tipo etilenglicol -(CH<sub>2</sub>CH<sub>2</sub>O)<sub>n</sub>-. El enlazador está presente para unir la biomolécula a través de residuo de COR<sub>13</sub> en la forma de un grupo ácido, éster o amida al resto de la construcción responsable para fluorescencia y para separar la biomolécula de la molécula de tinte.
- 20 R<sub>1</sub> puede ser H o un grupo alquilo o alquilo sustituido. R<sub>1</sub> se puede escoger de tal manera que R<sub>1</sub> puede no ser H cuando n es cero. R<sub>1</sub> puede ser metilo o etilo. R<sub>1</sub>, se puede unir a R<sub>2</sub> y/o R<sub>8</sub> para formar una estructura de anillo. El anillo puede tener un anillo de 5 o 6 miembros. El anillo puede tener un anillo todo de carbono, o puede contener heteroátomos adicionales.
- 25 R<sub>2</sub> puede ser H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo o alcoxi. Opcionalmente, R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es una cadena de carbonos o heterosustituida que forma un anillo.
- 30 R<sub>3</sub> puede ser H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo o alcoxi. Opcionalmente R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es un carbono o cadena heterosustituida que forma un anillo.
- R<sub>4</sub> puede ser H o un grupo alquilo o alquilo sustituido. R<sub>4</sub> no es H cuando m es cero. R<sub>4</sub> puede ser metilo o etilo. R<sub>4</sub> se puede unir a R<sub>3</sub> y/o R<sub>9</sub> para formar una estructura de anillo. El anillo puede ser un anillo de 5 o 6 miembros. El anillo puede ser un anillo todo de carbono en, o puede contener heteroátomos adicionales.
- 35 R<sub>5</sub> y R<sub>6</sub> puede ser H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxilo o alcoxi. Opcionalmente R<sub>5</sub> se puede unir a R<sub>2</sub>, y R<sub>6</sub> se puede unir a R<sub>3</sub>.
- 40 R<sub>8</sub> puede ser H, halógeno, grupo hidroxilo o alcoxi, grupo alquilo o alquilo sustituido. Opcionalmente R<sub>8</sub> se puede unir junto con R<sub>1</sub> para formar un carbono o cadena de carbono heterosustituidos que forman un anillo.
- 45 R<sub>9</sub> puede ser H, halógeno, grupo hidroxilo o alcoxi, grupo alquilo o alquilo sustituido. Opcionalmente R<sub>4</sub> se puede unir junto con R<sub>4</sub> para formar un carbono o cadena de carbono heterosustituidos que forman un anillo.
- R<sub>7</sub> junto con el C=O forma un grupo ácido, éster o amida. En particular R<sub>7</sub> puede ser OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, arilo o un arilo sustituido. Cuando R<sub>11</sub> y R<sub>12</sub> son H R<sub>7</sub> puede ser OH o NH<sub>2</sub>. R<sub>7</sub> puede ser un grupo alcoxi o un grupo amina primario o secundario con uno o dos grupos alquilo y/o arilo. El éster o amida COR<sub>7</sub> adicionalmente se puede sustituir.
- 50 R<sub>13</sub> junto con el C=O forma un grupo ácido, éster o amida. En particular R<sub>13</sub> puede ser OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido, arilo o arilo sustituido. R<sub>13</sub> puede ser OH o NH<sub>2</sub>. R<sub>13</sub> puede ser un alcoxi o una amina primaria o secundaria con uno o dos grupos alquilo/arilo. El éster o amida adicionalmente se puede sustituir. R<sub>13</sub> puede ser NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> es H o alquilo y R<sub>15</sub> es alquilo o un alquilo sustituido. La sustitución puede permitir la conjugación a biomoléculas. Las moléculas se pueden unir a fragmento estructural de núcleo de tinte de rodamina a través de R<sub>13</sub> para uso adicional.
- 55 Los compuestos de la invención pueden incluir el compuesto en el que R<sub>16</sub> y R<sub>17</sub> son alquilo no sustituido -(CH<sub>2</sub>)<sub>n</sub>H y -(CH<sub>2</sub>)<sub>m</sub>H en el que n es 1 a 3, R<sub>1</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, R<sub>8</sub> y R<sub>9</sub> todos son H, R<sub>2</sub> y R<sub>3</sub> son H o CH<sub>3</sub>. Dichos compuestos se muestran en la fórmula (III) adelante:





en la que k, n, m son independientemente enteros desde 1 hasta 3,

5 q es un entero desde 1 hasta 6,

R<sub>2</sub> y R<sub>3</sub> son independientemente H o CH<sub>3</sub>,

R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y

10

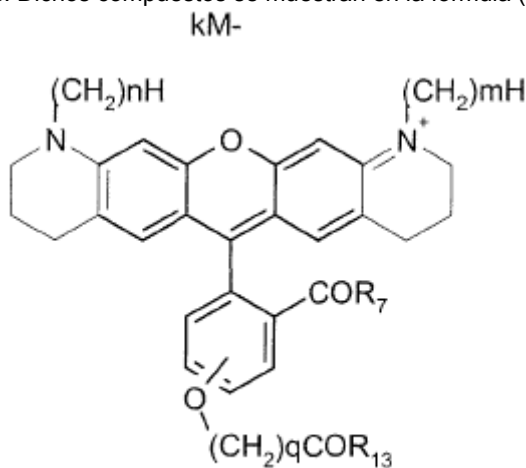
R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido.

15

Los compuestos de la invención pueden incluir el compuesto en el que R<sub>16</sub> y R<sub>17</sub> son H, R<sub>1</sub> se une a R<sub>2</sub> o R<sub>3</sub> a través de una cadena de grupos CH<sub>2</sub> para formar un anillo, y R<sub>4</sub> se une a R<sub>3</sub> o R<sub>6</sub> a través de una cadena de grupos CH<sub>2</sub> para formar un anillo.

20

Los compuestos de la invención pueden incluir el compuesto en el que R<sub>16</sub> y R<sub>17</sub> son H, R<sub>1</sub> se une a R<sub>2</sub> a través de una cadena de grupos CH<sub>2</sub> para formar un anillo de 6 miembros, y R<sub>4</sub> se une a R<sub>3</sub> a través de una cadena de grupos CH<sub>2</sub> para formar un anillo de 6 miembros. Dichos compuestos se muestran en la fórmula (IV) adelante:



(IV)

en la que k, n, m son independientemente enteros desde 1 hasta 3, en la que q es un entero desde 1 hasta 6,

25

R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y

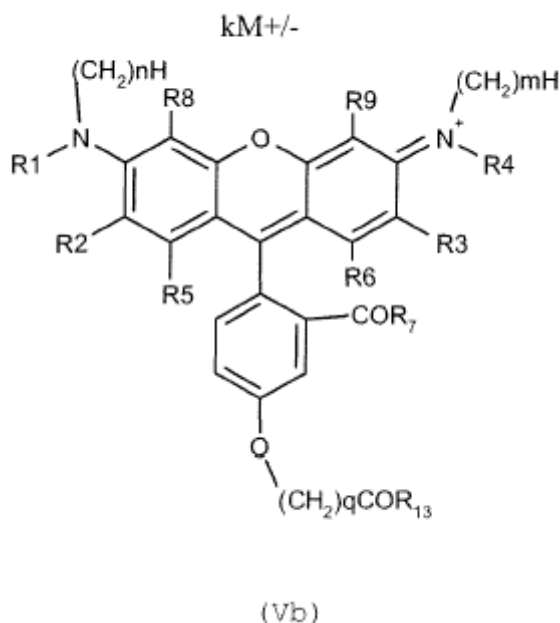
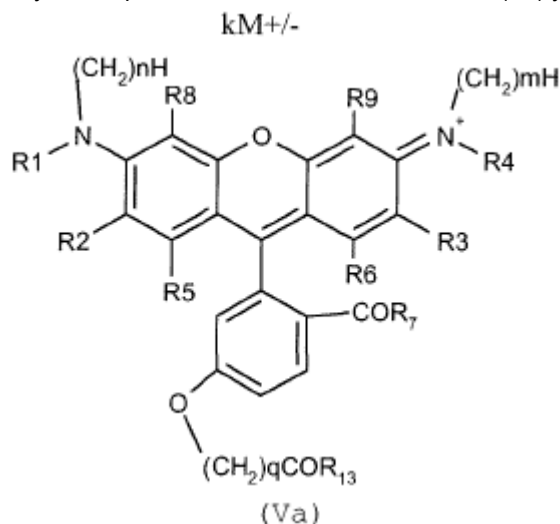
R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido.

Los tintes de la invención pueden incluir los compuestos en los que cualquiera de R<sub>1</sub>, R<sub>4</sub>, R<sub>16</sub> o R<sub>17</sub> son H o alquilo. El alquilo se puede sustituir con SO<sub>3</sub><sup>-</sup> compuestos de la invención pueden incluir el compuesto en el que R<sub>1</sub> y R<sub>4</sub> son H y

R<sub>16</sub> y/o R<sub>17</sub> es SO<sub>3</sub><sup>-</sup>.

El grupo R<sub>13</sub>CO-(CH<sub>2</sub>)<sub>q</sub>- se une al fragmento estructural de núcleo de tinte de rodamina a través de un heteroátomo, por ejemplo un átomo de oxígeno. Este heteroátomo, se puede unir a cualquiera de los átomos de carbono del anillo fenilo del fragmento estructural de núcleo de tinte de rodamina. Los compuestos se pueden preparar y utilizar como mezclas de isómeros posicionales en las que el heteroátomo o átomo de oxígeno está presente en diferentes posiciones del anillo benceno, o los compuestos pueden prepararse y usarse como isómeros individuales: Opcionalmente el anillo benceno de tales nuevos colorantes de rodamina pueden contener sustituyentes adicionales para una sintonización adicional fina de sus parámetros espectrales.

Los compuestos de la invención incluyen compuestos de acuerdo con la fórmula (Va) y (Vb), y mezclas de los mismos:



en la que M<sup>+/-</sup> es un contraión común,

k es un entero desde 0 hasta 6,

q es un entero desde 1 hasta 6,

R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxi- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es una cadena de carbonos o heterosustituida que forma un anillo,

R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxi- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub>

es una cadena de carbonos o heterosustituida que forma un anillo,

R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

5 R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxil- o alcoxi,

R<sub>8</sub> es H, halógeno, grupo hidroxil- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,

10 R<sub>9</sub> es H, halógeno, grupo hidroxil- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,

R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido,

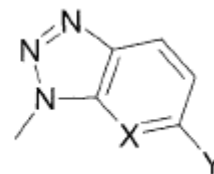
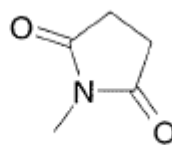
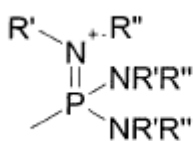
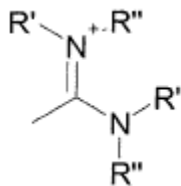
15 R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido, y

R<sub>16</sub> y R<sub>17</sub> son independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido.

20 Los compuestos de la invención se pueden unir a biomoléculas. Los compuestos de la invención se pueden unir a oligonucleótidos. Los compuestos de la invención se pueden unir a nucleótidos. Los compuestos se pueden unir a oligonucleótidos o nucleótidos a través del nucleótido base.

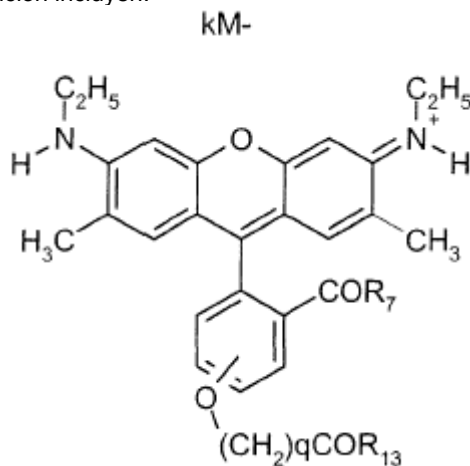
25 La unión a las biomoléculas puede ser a través de un grupo de COR<sub>7</sub> y/o R<sub>13</sub>CO(CH<sub>2</sub>)<sub>q</sub>-O- o a través de ambos de estos. El grupo R<sub>7</sub> y/o R<sub>13</sub> ser un grupo alquil-, aril-, o heteril-oxi, amina primaria o secundaria con un grupo R<sub>15</sub> sustituido, que se puede utilizar para unión. En particular un grupo de COR<sub>7</sub> y/o R<sub>13</sub>CO(CH<sub>2</sub>)<sub>q</sub>-O- puede ser el residuo de éster activado más adecuado para formación de enlace de amida/péptido adicional.

Por ejemplo, R<sub>7</sub> y/o R<sub>13</sub> pueden ser:



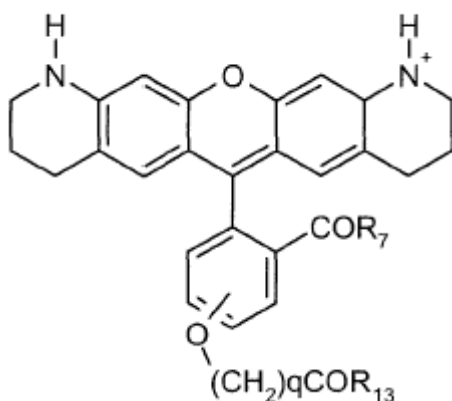
30

Ejemplos de compuestos de la invención incluyen:



y

kM-



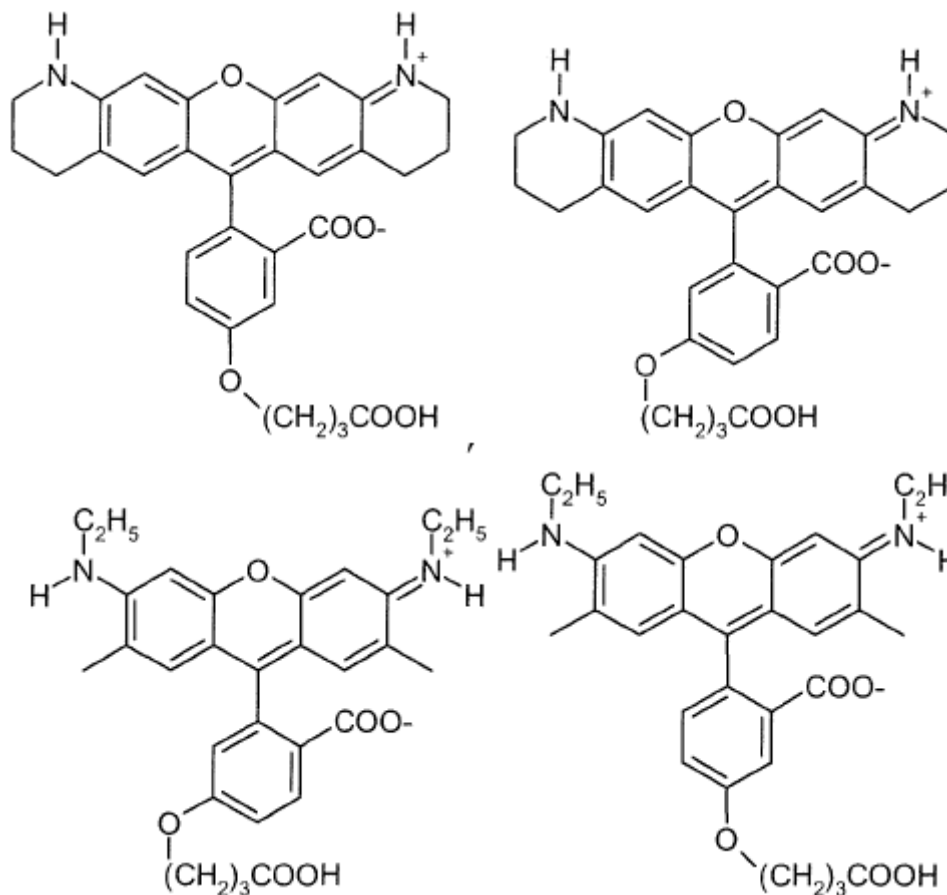
en el que q es un entero desde 1 hasta 6,

5

R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y

R<sub>13</sub> es OR<sub>14</sub> o NR<sub>14</sub>R<sub>15</sub> en el que R<sub>14</sub> y R<sub>15</sub> son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido.

Ejemplos adicionales incluyen



10

15

Un aspecto de la invención es un nucleótido u oligonucleótido marcado con un compuesto fluorescente como se describe aquí. El nucleótido u oligonucleótido marcado puede tener el marcador unido a través de un grupo alquilo sustituido R<sub>15</sub>. El nucleótido u oligonucleótido marcado puede tener el marcador unido a la posición C5 de una base de pirimidina o la posición C7 de una base de 7-deaza purina a través de un grupo funcional enlazador.

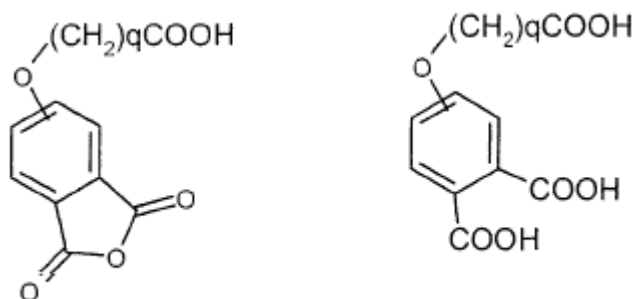
El nucleótido u oligonucleótido marcado también puede tener un grupo de bloqueo 3'OH unido covalentemente al azúcar de ribosa o desoxirribosa del nucleótido.

Se describen aquí kits que comprenden dos o más nucleótidos en los que por lo menos un nucleótido es un nucleótido marcado con un compuesto de la presente invención. El kit puede comprender dos o más nucleótidos marcados. Los nucleótidos se pueden marcar con dos o más marcadores fluorescentes. Dos o más de los marcadores se pueden excitar utilizando una fuente de excitación única, que puede ser un láser.

El kit puede contener cuatro nucleótidos marcados, en los que el primero de cuatro nucleótidos se marca con un compuesto como se divulga aquí, y el segundo, tercero, y cuarto nucleótidos cada uno se marca con un compuesto diferente, en el que cada compuesto tiene un máximo de absorbancia distinto y cada uno de los compuestos se puede distinguir de los otros tres compuestos. El kit puede ser de tal manera que dos o más de los compuestos tienen un máximo de absorbancia distinto por encima de 600 nm.

Los compuestos de la invención, nucleótidos o kits se pueden utilizar en secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN o ensayos de unión de proteína. El uso puede ser sobre un instrumento de secuenciación automatizado. El instrumento de secuenciación puede contener dos láseres que operan a diferentes longitudes de onda.

Se describe aquí nuevos compuestos de la fórmula (Xa, b) como materiales de partida para síntesis de tintes de la fórmula (I) de acuerdo con la invención.



(Xa, b)

en la que q es 1 a 6.

Se describe aquí un método para sintetizar compuestos de la invención. Un compuesto de la fórmula (Xa, b) se puede utilizar en una reacción de condensación con un derivado de 3-amino fenol sustituido o no sustituido o derivado heterocíclico que contiene este fragmento estructural.

Se han sintetizado los tintes de la fórmula (I) mediante una condensación de derivados de ácido ftálico (Xa, b) con derivados de 3-amino-fenol, preferiblemente a alta temperatura, con o sin catalizadores de ácido de Lewis. La reacción también se puede cumplir al utilizar ácido fosfórico o ácidos poli-fosfóricos como solvente y/o un catalizador. Se puede lograr la reacción de condensación con o sin catalizadores en líquidos iónicos. La reacción de condensación se puede lograr con o sin catalizadores en líquidos iónicos o solventes orgánicos polares de alto punto de ebullición como DMF, DMA, sulfolano o 1,2-diclorobenceno.

Como se utiliza aquí, el término "alquilo" se refiere a hidrocarburo C1-C20 y puede incluir anillos carbocíclicos no aromáticos C3-C10. En realizaciones particulares los grupos alquilo son alquilo C1-C6 que se refiere a radicales hidrocarburo de cadena lineal o ramificada, saturado que contiene entre uno y seis átomos de carbono, respectivamente.

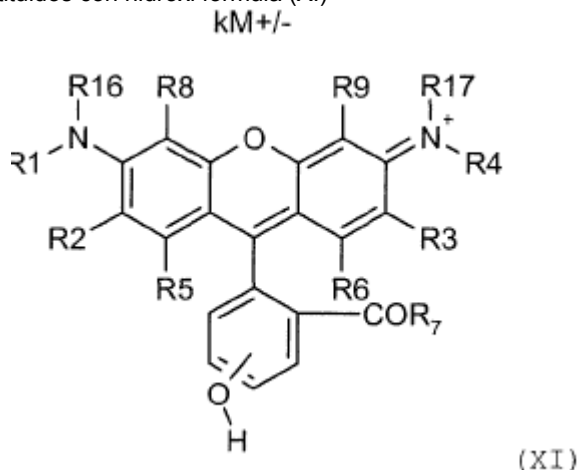
El término "halógeno" como se utiliza aquí se refiere a flúor (en lo sucesivo designado como F), cloro (en lo sucesivo designado como Cl), bromo (en lo sucesivo designado como Br) o yodo (en lo sucesivo designado como I), y usualmente se relaciona con la sustitución de un átomo de hidrógeno en un compuesto orgánico, esta sustitución es opcionalmente una sustitución completa para el hidrógeno.

El término "alquilo sustituido", se refiere a grupos alquilo, alquenoilo o alquiniilo como se definió anteriormente en el que pueden estar opcionalmente sustituidos adicionalmente con, pero no limitado a, halo, ciano,  $\text{SO}_3^-$ , SRa, ORa, NRbRc, oxo, CONRbRc, COOH y COORb. Ra, Rb y Rc se pueden seleccionar cada uno independientemente de H, alquilo, alquilo sustituido, alquenoilo, alquenoilo sustituido, alquiniilo, alquiniilo sustituido, arilo y arilo sustituido. Adicionalmente, dicho alquilo sustituido, alquenoilo sustituido y alquiniilo sustituido se puede interrumpir opcionalmente por lo menos mediante un heteroátomo o grupo seleccionado entre O, NRb, S(O)<sub>t</sub>, en el que t es 0 a 2, y similares. El alquilo sustituido también cubre grupos tal como bencilo donde los grupos alquilo comprenden un arilo adicional o un grupo funcional arilo sustituido.

Los tintes de acuerdo con la presente invención se pueden sintetizar a partir de una variedad de diferentes materiales de partida, que derivados sustituidos de N y/o C de 3-aminofenol, 5-hidroxil- y/o 7-hidroxil-1,2,3,4-tetrahidroquinolina y

5 otros materiales de partida aromáticos o heterocíclicos similares que contienen fragmento estructural de m-aminofenol. La condensación de estos compuestos con derivados de ácido ftálico adicionalmente sustituidos o no sustituidos, fórmula Xa, b da los tintes como se describe. La reacción de condensación normalmente se lleva a cabo a alta temperatura con o sin solvente adecuado, y es asistida por el uso de irradiación de microondas. El uso de líquidos iónicos como solvente en dichas reacciones de condensación es especialmente ventajoso. La reacción se puede catalizar por ácidos de Lewis.

Se pueden sintetizar los tintes de acuerdo con la invención también mediante métodos de alquilación convencional a partir de tintes de rodamina sustituidos con hidroxí fórmula (XI)



- 10 M<sup>+/-</sup> es un contraión común,  
 k es un entero desde 0 hasta 6,
- 15 R<sub>1</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,  
 R<sub>2</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxí- o alcoxi, o R<sub>2</sub> junto con R<sub>1</sub> o R<sub>5</sub> es una cadena de carbonos o heterosustituida que forma un anillo,
- 20 R<sub>3</sub> es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxí- o alcoxi o R<sub>3</sub> junto con R<sub>4</sub> o R<sub>6</sub> es una cadena de carbonos o heterosustituida que forma un anillo,  
 R<sub>4</sub> es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,
- 25 R<sub>5</sub> y R<sub>6</sub> son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxí- o alcoxi,  
 R<sub>8</sub> es H, halógeno, grupo hidroxí- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>1</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,
- 30 R<sub>9</sub> es H, halógeno, grupo hidroxí- o alcoxi, grupo alquilo o alquilo sustituido o junto con R<sub>4</sub> es un carbono o cadena de carbono heterosustituidos que forman un anillo,  
 R<sub>7</sub> es OR<sub>11</sub> o NR<sub>11</sub>R<sub>12</sub> en el que R<sub>11</sub> y R<sub>12</sub> son independientemente H, alquilo o un alquilo sustituido, y
- 35 R<sub>16</sub> y R<sub>17</sub> son independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido.

Se pueden preparar tintes de fórmula (XI) mediante condensación de derivados 3-aminofenol con derivados de ácido hidroxilo-ftálico como se describe adelante.

40 Se puede llevar a cabo preparación de N-alquil-5-hidroxi-1,2,3,4-tetrahidro-quinolina o N-sulfonatoalquil-7-hidroxi-1,2,3,4-tetrahidro-quinolina utilizando hidrogenación química o catalítica, por ejemplo utilizando Níquel Raney como catalizador. La adición de una base orgánica o inorgánica, por ejemplo trietilamina, mejora en gran medida el índice de reacción.

45 Se puede llevar a cabo mono N-alquilación de 3-aminofenoles con alquilsulfatos utilizando un equivalente o más del aminofenol. La di-alquilación del 3-amino fenol se puede lograr utilizando más equivalentes de alquilsufona. Ambos derivados fenólicos resultantes se pueden condensar con anhídrido ftálico para hacer fluoróforos como se describe.

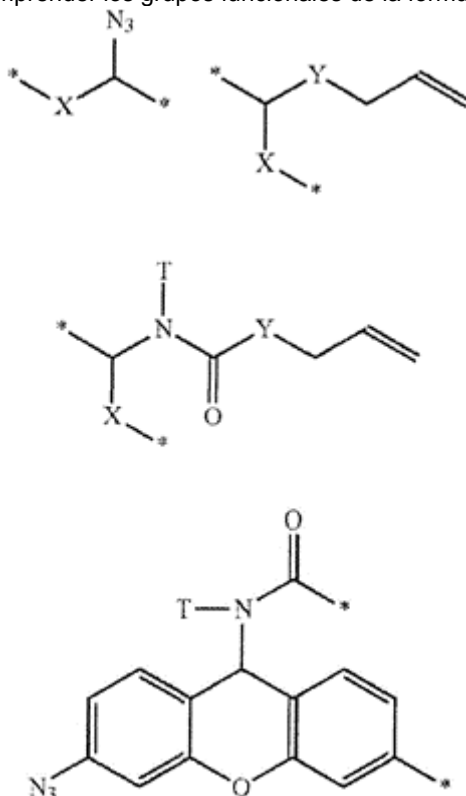
50 De acuerdo con un aspecto de la invención, se proporcionan compuestos de tinte adecuados para unión a grupos funcionales de sustratos, que comprenden particularmente grupos ligadores que permiten la adhesión a los grupos

funcionales de sustratos. Los grupos funcionales de sustratos pueden ser virtualmente cualquier molécula o sustancia a la que se puede conjugar los tintes de la invención y, por vía de ejemplo no limitante, pueden incluir nucleósidos, nucleótidos, polinucleótidos, carbohidratos, ligandos, partículas, superficies sólidas, polímeros orgánicos e inorgánicos y combinaciones o ensambles de los mismos, tal como cromosomas, núcleos, células vivas y similares. Los tintes se pueden conjugar mediante un ligador opcional mediante una variedad de medios, que incluyen atracción hidrófoba, atracción iónica y unión covalente. Particularmente los tintes se conjugan con el sustrato mediante adhesión covalente. Más particularmente, la adhesión covalente se hace por medio de un grupo ligador.

Los tintes de acuerdo con la invención pueden incluir un grupo ligador reactivo en una de las posiciones sustituyentes para unión covalente del tinte a otra molécula. Los grupos ligadores reactivos son grupos funcionales capaces de formar un enlace covalente. En una realización particular, el ligador puede ser un ligador divisible. El uso del término "ligador divisible" no significa que implica que se tenga que retirar el ligador completo. El sitio de división se puede ubicar en una posición sobre el ligador que asegura que parte del ligador permanece unido al tinte y/o grupo funcional de sustrato después de división. Los ligadores divisibles pueden ser, por vía de ejemplo no limitante, ligadores que se pueden dividir electrofílicamente, ligadores que se pueden dividir nucleófilamente, ligadores fotodivisibles, divisible bajo condiciones de reducción (por ejemplo ligadores que contienen azida o disulfuro), condiciones de oxidación, divisible a través del uso de ligadores de enganche de seguridad y divisibles mediante mecanismos de eliminación. El uso de un ligador divisible para unir el compuesto de tinte a un grupo funcional de sustrato asegura que el marcador puede, si se requiere, ser retirado después de detección, evitando cualquier señal de interferencia en etapas descendentes.

Grupos ligadores particulares se pueden encontrar en la solicitud de patente en trámite No WO2004/018493 (incorporada aquí mediante referencia) en el que los presentes inventores han encontrado que determinados ligadores que conectan las bases de los nucleótidos a los marcadores tal como, por ejemplo, los tintes de acuerdo con la presente invención, se pueden separar utilizando fosfinos solubles en agua o catalizadores de metales de transición solubles en agua formados a partir de un metal de transición y por lo menos ligandos parcialmente solubles en agua. En soluciones acuosas la última forma por lo menos parcialmente complejos de metales de transición solubles en agua.

Se pueden encontrar ligadores particulares en las solicitantes internacionales en trámite WO2004/018493 (incorporada aquí mediante referencia) y puede comprender los grupos funcionales de la fórmula:



(En el que X se selecciona del grupo que comprende O, S, NH y NQ en el Q es un grupo alquilo C1-10 sustituido o no sustituido, Y se selecciona del grupo que comprende O, S, NH y N (alilo), T es hidrógeno o un grupo alquilo C1-10 sustituido o no sustituido y \* indica en dónde se conecta el grupo funcional al resto del nucleótido o nucleósido).

Aún más particularmente los inventores han determinado en la solicitud de patente GB en trámite No 0517097.2, publicada como el documento WO07020457, que al alterar, y en particularmente aumentar la longitud del ligador entre un tinte fluorescente (fluoróforo) y la base guanina, al introducir un grupo separador polietilenglicol, es posible aumentar la intensidad de fluorescencia en comparación con el mismo fluoróforo unido a la base guanina a través de otros

enlaces conocidos en la técnica. El diseño de los ligadores, y especialmente su aumento de longitud, también permite mejoras en el brillo de los fluoróforos unidos a bases guanina de nucleótidos guanosina cuando se incorpora en polinucleótidos tales como ADN. De esta manera, cuando el tinte es para uso en cualquier método de análisis que requiere detección de una etiqueta de tinte fluorescente unida a un nucleótido que contiene guanina, es ventajoso si el ligador comprende un grupo separador de la fórmula  $-((\text{CH}_2)_2\text{O})_n-$  en el que n es un entero entre 2 y 50, como se describe en las solicitudes en trámite No GB0517097.2 (WO07020457).

La presente invención proporciona adicionalmente conjugados de nucleósidos y nucleótidos marcados con tintes de acuerdo con la invención (nucleótidos modificados). Los nucleótidos y nucleósidos marcados son útiles para marcar polinucleótidos formados mediante síntesis enzimática, tal como, por vía de ejemplo no limitante, en amplificación PCR, amplificación isotérmica o amplificación de fase sólida, secuenciamiento de polinucleótidos que incluyen secuenciamiento de fase sólida, reacciones de traducción de muesca y similares.

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

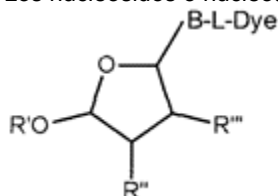
Un "nucleósido" es estructuralmente similar al nucleótido, pero carece de los grupos funcional fosfato. Un ejemplo de un análogo nucleósido sería uno en el que el marcador se liga a la base y no hay un grupo fosfato unido a la molécula de azúcar.

Aunque la base se denomina usualmente como una purina o pirimidina, el experto apreciará que están disponibles derivados y análogos que no alteran la capacidad del nucleótido o nucleósido para experimentar el emparejamiento base Watson-Crick. "Derivado" o "análogo" significa un compuesto o molécula cuya estructura central es la misma que, o se asemeja cercanamente a aquella de un compuesto pariente, pero que tiene una modificación física o química, tal como, por ejemplo, un grupo colateral adicional o diferente, que permite que el derivado nucleótido o nucleósido se ligue a otra molécula. Por ejemplo, la base puede ser una desazapurina. Los derivados deben ser capaces de experimentar un emparejamiento Watson-Crick. El "derivado" y "análogo" también significa un nucleótido o nucleósido sintético derivado que tiene grupos funcionales base modificados y/o grupos funcionales de azúcar modificados. Dichos derivados y análogos se discuten en, por ejemplo, Scheit, Nucleotide analogs (John Wiley & Son, 1980) and Uhlman et al., Chemical Reviews 90:543-584, 1990. Los análogos de nucleótidos también pueden comprender enlaces fosfodiéster modificados que incluyen fosforotioato, fosforditioato, alquil-fosfonato, fosforanilidato, fosforamidato y similares.

El tinte se puede unir a cualquier posición en la base de nucleótido, a través de un ligador, siempre que se lleve a cabo el emparejamiento base Watson-Crick. Los sitios de marcado de nucleobase particulares incluyen la posición C5 de una base pirimidina o la posición C7 de una base 7-desaza purina. Como se describió anteriormente se puede utilizar un grupo ligador para unir covalentemente un tinte al nucleósido o nucleótido.

En realizaciones particulares, el nucleósido o nucleótido marcado se puede incorporar enzimáticamente y extender enzimáticamente. De acuerdo con lo anterior un grupo funcional ligador puede tener una longitud suficiente para conectar el nucleótido al compuesto de tal manera que el compuesto no interfiere significativamente con la unión general y reconocimiento del nucleótido mediante una enzima de replicación de ácido nucleico. De esta manera, el ligador también puede comprender una unidad separadora. El separador separa, por ejemplo, la base de nucleótido de un marcador o sitio de división.

Los nucleósidos o nucleótidos etiquetados con tintes de la invención pueden tener la fórmula:



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

Cuando R'' es fosforoamidita, R' es un grupo protector hidroxilo divisible por ácido que permite el acoplamiento posterior de monómero bajo condiciones sintética automatizada. En una realización particular, el grupo de bloqueo se separa y es



independiente del compuesto de tinte, es decir, no se une a este. En una realización alternativa, el tinte puede comprender todo o parte del grupo de bloqueo 3' OH. De esta manera R<sup>1</sup> puede ser un grupo de bloqueo 3' OH que puede o no comprender el compuesto de tinte.

5 En aun otra realización alternativa no existe un grupo de bloqueo en el carbono 3' del azúcar pentosa y el tinte (o tinte y construcción ligadora) unido a la base, por ejemplo, puede tener un tamaño o estructura suficiente para actuar como un bloque a la incorporación de un nucleótido adicional desde un punto diferente al sitio 3'. De esta manera, el bloque se puede deber al impedimento estérico o se puede deber a una combinación de tamaño, carga y estructura.

10 El uso de un grupo de bloqueo permite la polimerización que se va a controlar, tal como al detener la extensión cuando se incorpora un nucleótido modificado. Si el efecto de bloqueo es reversible, por ejemplo por vía de un ejemplo no limitante, al cambiar las condiciones químicas o al remover un bloqueo químico, la extensión se puede detener en determinados puntos y luego se puede permitir continuar.

15 En otra realización particular, un grupo de bloqueo 3' OH comprenderá grupos funcionales divulgados en el documento WO2004/018497 (incorporado aquí mediante referencia). Por ejemplo, el grupo de bloqueo puede ser azidometilo (CH<sub>2</sub>N<sub>3</sub>) o alilo.

20 En una realización particular los grupos ligadores y de bloqueo están presentes y son grupos funcionales separados, que se pueden dividir bajo condiciones sustancialmente similares. De esta manera los procesos de desprotección y desbloqueo pueden ser más eficientes en razón a que sólo se requerirá un único tratamiento para retirar el compuesto de tinte y el bloque.

25 La invención también abarca polinucleótidos que incorporan compuestos de tintes de acuerdo con la presente invención. Dichos polinucleótidos pueden ser ADN o ARN que comprenden, respectivamente, desoxirribonucleótidos o ribonucleótidos unidos con enlaces fosfodiéster. Los polinucleótidos de acuerdo con la invención pueden comprender nucleótidos que se presentan en la naturaleza, nucleótidos que no se presentan en la naturaleza (o modificado) diferente de los nucleótidos modificados de la invención o cualquier combinación de los mismos, siempre que por lo menos un nucleótido modificado, es decir, marcado con un compuesto de tinte, de acuerdo con la invención está presente. Los polinucleótidos de acuerdo con la invención también pueden incluir enlaces de estructura principal no natural y/o modificaciones químicas de no nucleótidos. También se contemplan estructuras quiméricas comprendidas de mezclas de ribonucleótidos y desoxirribonucleótidos que comprenden por lo menos uno de los nucleótidos modificados de acuerdo con la presente invención.

35 Los nucleótidos modificados (o nucleósidos) que comprenden un compuesto de tinte de acuerdo con la invención se puede utilizar en cualquier método de análisis que requiere la detección de un marcador fluorescente unido a un nucleótido o nucleósido, ya sea sobre sí mismo o incorporado en o asociado con un conjugado o estructura molecular mayor. En este contexto, el término "incorporado en un polinucleótido" requiere que el fosfato 5' se una en un enlace fosfodiéster al grupo 3' hidroxilo de un segundo nucleótido (modificado o no modificado), que puede en sí mismo hacer parte de una cadena de polinucleótidos mayor. El extremo 3' del nucleótido modificado de la invención puede o no puede estar unido a un enlace fosfodiéster al fosfato 5' de un nucleótido adicional (modificado o no modificado). De esta manera, en una realización no limitante, la invención proporciona un método de detectar un nucleótido modificado incorporado en un polinucleótido que comprende: (a) incorporar por lo menos un nucleótido modificado de acuerdo con el tercer aspecto de la invención en un polinucleótido y (b) detectar los nucleótidos modificados incorporados en el polinucleótido a detectar la señal fluorescente del compuesto de tinte unido a dicho nucleótidos modificados.

Este método requiere dos etapas esenciales: una etapa de sintética (a) en la que uno o más nucleótidos modificados de acuerdo con la invención se incorporan en un polinucleótido y una etapa de detección (b) en la que se detecta uno o más nucleótidos modificados incorporados en el polinucleótido al detectar o medir cuantitativamente su fluorescencia.

50 En una realización de la invención, por lo menos un nucleótido modificado se incorpora en un polinucleótido en la etapa sintética mediante la acción de una enzima de polimerasa. Sin embargo, otros métodos para unir los nucleótidos modificados a los polinucleótidos, tal como, por ejemplo, la síntesis química de oligonucleótidos o ligado de oligonucleótidos marcados a un oligonucleótido marcado, no se excluyen. Por lo tanto, en el contexto específico de este método de la invención, el término "que incorpora" un nucleótido en un polinucleótido abarca la síntesis de polinucleótidos mediante métodos químicos, así como métodos enzimáticos.

60 En una realización específica la etapa sintética puede comprender incubar una cadena de polinucleótido plantilla con una mezcla de reacción que comprende nucleótidos modificados marcados fluorescentemente de la invención y polimerasas bajo condiciones que permitan la formación de un enlace fosfodiéster entre un grupo hidroxilo 3' libre sobre una cadena de polinucleótido hibridada a dicha cadena de polinucleótidos plantilla y un grupo fosfato 5' sobre dicho nucleótido modificado. Esta realización comprende una etapa sintética en la que la formación de una cadena de polinucleótido se dirige mediante emparejamiento base complementario de nucleótidos a una cadena plantilla.

65 En todas las realizaciones del método, la etapa de detección se puede llevar a cabo mientras que la cadena de polinucleótidos en el que se incorporan los nucleótidos modificados se hibrida a una cadena plantilla, o después de una

etapa de desnaturalización en la que se separan las dos cadenas. Etapas adicionales, por ejemplo etapas de reacción química o enzimática o etapas de purificación, se pueden incluir entre la etapa sintética y la etapa de detección. En particular, la cadena objetivo que incorpora los nucleótidos modificados se puede aislar o purificar y luego procesar adicionalmente o utilizar en un análisis posterior. Por vía de ejemplo, los polinucleótidos objetivo marcados con nucleótidos modificados de acuerdo con la invención en una etapa sintética se pueden utilizar posteriormente como cebadores o sondas marcadas. En otras realizaciones el producto de la etapa sintética se puede someter a etapas de reacción y, si se desea, el producto de estas etapas posteriores se purifica o aísla.

Condiciones adecuadas para la etapa sintética son bien conocidas por las personas familiarizadas con las técnicas de biología molecular estándar. En una realización, la etapa sintética puede ser análoga a una reacción de extensión de cebador estándar utilizando precursores de nucleótidos, que incluyen nucleótidos modificados de acuerdo con la invención, para formar una cadena objetivo extendida complementaria a la cadena plantilla en la presencia de una enzima de polimerasa adecuada. En otras realizaciones, la etapa sintética puede en sí misma hacer parte de una reacción de amplificación que produce un producto de amplificación de cadena doble marcada que comprende cadenas complementarias hibridadas a partir de la copiar de las cadenas de polinucleótidos plantilla y objetivo. Otras etapas "sintéticas" de ejemplo incluyen traducción de muesca, polimerización de desplazamiento de cadena, marcado de ADN cebado aleatorio etc. La enzima polimerasa utilizar en la etapa sintética puede ser capaz de catalizar la incorporación de nucleótidos modificados de acuerdo con la invención. Otra forma, la naturaleza precisa de la polimerasa no se limita particularmente, si no que puede depender de las condiciones de la reacción sintética. Por vía de ejemplo, si la reacción sintética se lleva a cabo utilizando termociclado entonces se requiere una polimerasa termoestable, mientras que esto puede no sea esencial para reacciones de extensión de cebador estándar. Las polimerasas termoestables adecuadas que son capaces de incorporar los nucleótidos modificados de acuerdo con la invención incluyen aquellos descritos en el documento WO 2005/024010 o WO06120433. En reacciones sintéticas que se llevan a cabo a temperaturas inferiores tal como 37°C, no se necesita que las enzimas polimerasas sean polimerasas termoestables, por lo tanto, la elección de la polimerasa dependerá de una serie de factores tal como la temperatura de reacción, pH, actividad de desplazamiento de cadena y similares.

En realizaciones no limitantes específicas la invención abarca el uso de nucleósidos o nucleótidos modificados marcados con tintes de acuerdo con la invención en un método de secuenciamiento de ácido nucleico, resecuenciamiento, secuenciamiento de genoma completo, calificación de polimorfismo de nucleótido sencillo, cualquier otra aplicación que implique la detección del nucleótido o nucleósido modificado cuando se incorpora en un polinucleótido, o cualquier otra aplicación que requiera el uso de polinucleótidos marcados con los nucleótidos modificados que comprenden tintes fluorescentes de acuerdo con la invención.

En una realización particular, la invención proporciona el uso de nucleótidos modificados que comprenden compuestos de tintes de acuerdo con la invención en una reacción de "secuenciamiento por síntesis" de polinucleótido. El secuenciamiento por síntesis implica generalmente la adición secuencial de uno o más nucleótidos u oligonucleótidos a una cadena de polinucleótidos creciente en la dirección 5' a 3' utilizando una polimerasa oligasa con el fin de formar una cadena de polinucleótidos extendida complementaria al ácido nucleico plantilla que se va a secuenciar. La identidad de la base presente en uno o más de los nucleótidos agregados se determina en una etapa de detección o "formación de imagen". La identidad de la base agregada se puede determinar después de cada etapa de incorporación de nucleótidos. La secuencia de la plantilla se puede deducir utilizando reglas de emparejamiento base Watson-Crick. El uso de los nucleótidos modificados marcados con tintes de acuerdo con la invención para la determinación de la identidad de una base sencilla puede ser útil, por ejemplo, en la clasificación de polimorfismos de nucleótidos sencillos, y dichas reacciones de extensión de base sencilla están dentro del alcance de esta invención.

En una realización de la invención, la secuencia de un polinucleótido plantilla se determina al detectar la incorporación de uno o más nucleótidos en una cadena naciente complementaria al polinucleótido plantilla que se va a secuenciar a través de la detección de marcadores fluorescentes adheridos a los nucleótidos incorporados. El secuenciamiento de los polinucleótidos plantilla se ceba con un cebador adecuado (o se prepara como una construcción de horquilla que contendrá el cebador como parte de la horquilla), y la cadena naciente se extiende en forma de etapas mediante la adición de nucleótidos al extremo 3' del cebador en una reacción catalizada por polimerasas.

En realizaciones particulares cada uno de los diferentes trifosfatos de nucleótidos (A, T, G y C) se pueden marcar con un fluoróforo único y también comprende un grupo de bloqueo en la posición 3' para evitar la polimerización no controlada. Alternativamente uno de los cuatro nucleótidos puede no estar marcado (oscuro). La enzima de polimerasa incorpora un nucleótido en una cadena naciente complementaria al polinucleótido plantilla, y el grupo de bloqueo evita la incorporación adicional de nucleótidos. Se elimina nucleótidos no incorporados y la señal fluorescente de cada nucleótido incorporado se "lee" ópticamente mediante métodos adecuados, tal como un dispositivo acoplado a carga que utiliza excitación láser y filtros de emisión adecuados. El grupo de bloqueo 3' y los compuestos de tintes fluorescentes se retiran posteriormente (desprotegen), particularmente mediante el mismo método químico o enzimático, para exponer la cadena naciente a incorporación adicional de nucleótidos. Normalmente, la identidad del nucleótido incorporado se determinará después de cada etapa de incorporación, pero esto no es estrictamente esencial. Del mismo modo, la Patente Estadounidense No. 5,302,509 divulga un método para secuenciar polinucleótidos inmovilizados sobre un soporte sólido. El método se basa en la incorporación de nucleótidos A, G, C y T bloqueados en 3' marcados fluorescentemente en una cadena creciente complementaria a los polinucleótidos inmovilizados, en la presencia de

5 polimerasa de ADN. La polimerasa incorpora una base complementaria al polinucleótido objetivo, pero se evita la adición mediante el grupo de bloqueo 3'. El marcado del nucleótido incorporado se puede determinar luego y el grupo de bloqueo removido mediante división química para permitir que ocurra polimerización adicional. La plantilla de ácido nucleico que se va a secuenciar en una reacción de secuenciamiento por síntesis puede ser cualquier polinucleótido que se desee secuenciar. La plantilla de ácido nucleico para una reacción de secuenciamiento comprenderá normalmente una región de cadena doble que tiene un grupo hidroxilo 3' libre que sirve como un cebador o punto de inicio para la adición de nucleótidos adicionales en la reacción de secuenciamiento. La región de la plantilla que se va a secuenciar sobresaldrá de este grupo hidroxilo 3' libre en la cadena complementaria. La región que sobresale de la plantilla que se va a secuenciar puede ser de cadena sencilla, pero puede ser de doble cadena, siempre que "se presente una muesca" en la cadena complementaria a la cadena plantilla que se va a secuenciar para proporcionar un grupo OH 3' libre para iniciación de la reacción de secuenciamiento. En dichas realizaciones la secuenciamiento puede proceder mediante desplazamiento de cadenas. En determinadas realizaciones un cebador lleva un grupo hidroxilo 3' libre que se puede agregar como un componente separado (por ejemplo, un oligonucleótido corto) que hibrida a una región de cadena sencilla de la plantilla que se va a secuenciar. Alternativamente, el cebador y la cadena de plantilla que se van a secuenciar cada uno pueden hacer parte de una cadena de ácido nucleico parcialmente autocomplementaria capaz de formar un dúplex intramolecular, tal como por ejemplo una estructura de bucle horquilla. Los polinucleótidos horquilla y los métodos mediante los cuales se pueden unir a soportes sólidos se divulgan en la publicación de la solicitud internacional copendiente del solicitante No. WO0157248 y WO 2005/047301. Los nucleótidos se agregan sucesivamente a los grupos hidroxilo 3' libres, 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 se puede determinar, particularmente pero no necesariamente después de cada adición de nucleótidos, proporcionando de esta manera información de secuencia para la plantilla de ácido nucleico. El término "incorporación" de un nucleótido en una cadena de ácido nucleico (o polinucleótidos) en este contexto se refiere a la unión del nucleótido al grupo hidroxilo 3' libre de la cadena de ácido nucleico a través de la formación de una conexión fosfodiéster con el grupo fosfato 5' del nucleótido.

25 La plantilla de ácido nucleico que se va a secuenciar puede ser ADN o ARN, o incluso una molécula híbrida comprendida de desoxinucleótidos y ribonucleótidos. La plantilla de ácido nucleico puede comprender nucleótidos que se presentan en forma natural y/o que no se presentan en forma natural y enlaces de estructura principal naturales o no naturales, siempre que éstos no eviten copiar la plantilla en la reacción de secuenciamiento.

30 En determinadas realizaciones, la plantilla de ácido nucleico que se va a secuenciar se puede unir a un soporte sólido a través de cualquier método de enlace adecuado conocido en la técnica, por ejemplo a través de unión covalente. En determinadas realizaciones los polinucleótidos plantilla se pueden unir directamente a un soporte sólido (por ejemplo, un soporte basado en sílice). Sin embargo, en otras realizaciones de la invención, la superficie del soporte sólido se puede modificar en alguna manera con el fin de permitir la unión covalente directa de los polinucleótidos plantilla, o inmovilizar los polinucleótidos plantilla a través de una multicapa de hidrogel o polielectrolitos, que en sí mismo se pueden unir no covalentemente al soporte sólido. Matrices en las que se han unido polinucleótidos directamente a soportes a basados en sílice son aquellas, divulgadas, por ejemplo, en el documento WO00006770, en el que los polinucleótidos se inmovilizan sobre un soporte de vidrio mediante la reacción entre un grupo epóxido pendiente en el vidrio con un grupo amino interno en el polinucleótido. Adicionalmente, los solicitantes divulgan en una solicitud de patente internacional copendiente No de publicación WO2005/047301 matrices de polinucleótidos unidas a un soporte sólido, por ejemplo para uso en la preparación de SMA, mediante la reacción de un nucleófilo basado en azufre con soporte sólido. Un ejemplo adicional de polinucleótidos de plantilla soportados con sólidos es cuando los polinucleótidos plantilla se unen a hidrogeles soportador en soportes sólidos o basados en sílice. Los soportes basados en sílice se utilizan normalmente para soportar hidrogeles y matrices de hidrogeles como se describe en el documento WO00/31148, WO01/01143, WO02/12566, WO03/014392, Patente estadounidense No. 6,465,178 y WO00/53812.

45 Una superficie particular a la que se pueden inmovilizar polinucleótidos plantilla es un hidrogel poliacrilamida. Los hidrogeles poliacrilamida se describen en la técnica anterior, algunos de los cuales se discutieron anteriormente. Sin embargo, un hidrogel particular se describe en el documento WO2005/065814.

50 Las moléculas de plantilla de ADN se pueden unir a lechos o micropartículas con el propósito de secuenciamiento; por ejemplo como se describe en la patente estadounidense No 6,172,218. Ejemplos adicionales de la preparación de colecciones de lecho, en donde cada glóbulo contiene secuencias de ADN diferentes se puede encontrar en la técnica anterior (Nature. 437, 376-380 (2005); Science. 309, 5741, 1728-1732 (2005)). El secuenciamiento de matrices de dichos glóbulos que utilizan nucleótidos como se describe está dentro del alcance de la invención.

60 Las plantillas que se van a secuenciar, pueden hacer parte de una "matriz" sobre un soporte sólido, en cuyo caso la matriz puede tomar cualquier forma conveniente. De esta manera, el método de la invención se puede paliar a todos los tipos de matrices de "alta densidad", que incluyen matrices de una molécula, matrices agrupadas y matrices de glóbulos. Los nucleótidos modificados marcados con compuestos de tinte de la invención se pueden utilizar para plantillas de secuenciamiento en esencialmente cualquier tipo de matriz formada por inmovilización de moléculas de ácido nucleico sobre un soporte sólido, y más particularmente cualquier tipo de matriz de alta densidad. Sin embargo, los nucleótidos modificados marcados con compuestos de tinte de la invención son particularmente ventajosos en el contexto del secuenciamiento de matrices agrupadas.

65

En matrices de multipolínucleótido o matrices agrupadas, distintas regiones en la matriz comprenden múltiples moléculas de plantillas de polinucleótidos. El término "matriz agrupada" se refiere a una matriz en el que distintas regiones o sitios en la matriz comprenden múltiples moléculas de polinucleótidos que se pueden resolver individualmente mediante métodos ópticos. Dependiendo de cómo se forma la matriz cada sitio en la matriz puede comprender múltiples copias de una molécula de polinucleótidos individual o incluso múltiples copias de un pequeño número de diferentes moléculas de polinucleótidos (por ejemplo, múltiples copias de dos cadenas de ácido nucleico complementarias). Se pueden producir matrices agrupadas o de multipolínucleótido de moléculas de ácido nucleico utilizando técnicas conocidos generalmente en la materia. Por vía de ejemplo, los documentos WO 98/44151 y WO00/18957 describen métodos de amplificación de ácidos nucleicos en el que los productos de amplificación y plantilla permanecen inmovilizados sobre un soporte sólido con el fin de formar matrices comprendidas de grupos o "colonias" de moléculas de ácido nucleico inmovilizadas. Las moléculas de ácido nucleico presentes en las matrices agrupadas preparadas de acuerdo con estos métodos son plantillas adecuadas para secuenciamiento utilizando los nucleótidos modificados marcados con compuestos tinte de la invención.

Los nucleótidos modificados marcados con los compuestos de tinte de la invención también son útiles en el secuenciamiento de plantillas sobre matrices de unas moléculas individuales. El término "matriz de molécula individual" o "SMA" como se utiliza aquí se refiere a una población de moléculas de polinucleótidos, distribuidas (o dispuestas) sobre un soporte sólido, en el que la separación de cualquier polinucleótido individual de todas las otras poblaciones es tal que es posible efectuar resolución individual de los polinucleótidos. Las moléculas de ácido nucleico objetivo inmovilizadas sobre la superficie del soporte sólido debe ser capaz de ser resuelta mediante medios ópticos. Esto significa que, dentro del área soluble del dispositivo formador de imágenes particular utilizado, puede haber una o más señales distintas, cada una representa un polinucleótido.

Esto se puede lograr cuando la separación entre las moléculas de polinucleótidos adyacentes sobre la matriz tenga por lo menos 100 nm, más particularmente por lo menos 250 nm, aún más particularmente por lo menos 300 nm, aún más particularmente por lo menos 350 nm. De esta manera, cada molécula puede ser resoluble individualmente y detectable como un punto fluorescente de molécula sencilla, y la fluorescencia de dicho punto fluorescente de molécula sencilla también presenta fotoblanqueamiento de etapa sencilla.

Los términos "resuelto individualmente" y "resolución individual" se utilizan aquí para especificar que, cuando se visualizan, es posible distinguir una molécula en la matriz de sus moléculas vecinas. La separación entre las moléculas individuales en la matriz se determinará, en parte, mediante técnicas particulares utilizadas para resolución de las moléculas individuales. Las características generales de las matrices de molécula sencilla se entenderán con referencia a las solicitudes publicadas WO00/06770 y WO 01/57248. Aunque un uso de los nucleótidos modificados de la invención está en las reacciones de secuenciamiento por síntesis, la utilidad de los nucleótidos modificados no se limita a dichos métodos. De hecho, los nucleótidos se pueden utilizar ventajosamente en cualquier metodología de secuenciamiento que requiere la detección de etiquetas fluorescentes unidas a nucleótidos incorporados en un polinucleótido.

En particular, los nucleótidos modificados marcados con compuestos de tinte de la invención se pueden utilizar en protocolos de secuenciamiento fluorescente, particularmente secuenciamiento de ciclo de terminador de tinte fluorescente basado en el método de secuenciamiento de terminación de cadena de Sanger y colaboradores. Dichos métodos utilizan por lo general secuenciamiento por ciclos y enzimas para incorporar didesoxinucleótidos marcados fluorescentemente en una reacción de secuenciamiento de extensión de cebador. De esta manera los denominados métodos de secuenciamiento Sanger, y protocolos relacionados (tipo Sanger), se basan en la terminación de cadena aleatorizada con didesoxinucleótidos marcados.

De esta manera, la invención también abarca nucleótidos modificados marcados con compuestos de tinte de acuerdo con la invención que son didesoxinucleótidos que carecen de grupos hidroxilo en las posiciones 3' y 2', tal como didesoxinucleótidos que adecuados para uso en métodos de secuenciamiento tipo de Sanger y similares.

Los nucleótidos modificados marcados con compuestos de tinte de la presente invención que incorporan grupos de bloqueo 3', se reconocerán, también son de utilidad en los métodos Sanger y protocolos relacionados en razón a que el mismo efecto alcanzado al utilizar los nucleótidos didesoxi modificados se puede alcanzar al utilizar nucleótidos modificados que tienen grupos de bloqueo OH 3': que evitan la incorporación de nucleótidos posteriores. En donde los nucleótidos de acuerdo con la presente invención, y que tienen un grupo de bloqueo 3' que se van a utilizar en los métodos de secuenciamiento tipo Sanger, se apreciará que los compuestos de tinte o marcas detectables unidas a los nucleótidos no necesitan conectar a través de enlaces divisibles, en razón a que en cada caso se incorpora un nucleótido marcado de la invención; no se necesita incorporar posteriormente nucleótidos y de esta manera no se necesita retirar la marca del nucleótido.

La invención también proporciona kits que incluyen nucleósidos y/o nucleótidos marcados modificados con tintes de acuerdo con la invención. Dichos kits incluirán por lo general por lo menos un nucleótido o nucleósido modificado marcado con un tinte de acuerdo con la invención junto con por lo menos un componente adicional. Los componentes adicionales pueden ser nucleótidos o nucleósidos modificados o no modificados adicionalmente. Por ejemplo, los nucleótidos modificados marcados con tintes de acuerdo con la invención se pueden suministrar en combinación con

nucleótidos no marcados o naturales, y/o con nucleótidos marcados fluorescentemente o cualquier combinación de los mismos. De acuerdo con lo anterior, los kits pueden comprender nucleótidos modificados marcados con tintes de acuerdo con la invención y nucleótidos modificados marcados con otros, por ejemplo, compuestos de tintes de la técnica anterior. Se pueden proporcionar combinaciones de nucleótidos como componentes individuales separados o como mezclas de nucleótidos.

Cuando los kits comprenden una pluralidad, particularmente dos, particularmente cuatro, nucleótidos modificados marcados con un compuesto de tinte, los nucleótidos diferentes se pueden marcar con diferentes compuestos de tinte, o uno puede ser oscuro, sin compuestos de tinte. Cuando los diferentes nucleótidos se marcan con diferentes compuestos de tinte es una característica de los kits que dichos compuestos de tinte sean tintes fluorescentes distinguibles espectralmente. Como se utiliza aquí, el término "tintes fluorescentes distinguibles espectralmente" se refiere a tintes fluorescentes que emiten energía fluorescente en longitudes de onda que se pueden distinguir mediante equipos de detección fluorescente (por ejemplo, una plataforma comercial de secuenciamiento de ADN basado en capilares) en donde dos o más de dichos tintes están presentes en una muestra. Cuando dos nucleótidos modificados marcados con compuestos de tinte fluorescentes se suministran en forma de kit, es una característica de la invención que los tintes fluorescentes distinguibles espectralmente se pueden excitar en la misma longitud de onda, tal como, por ejemplo, por el mismo láser. Cuando cuatro nucleótidos modificados marcados con compuestos de tinte fluorescente se proporcionan en forma de kit, es una característica de la invención que dos tintes fluorescentes distinguibles espectralmente puedan ser excitados en una longitud de onda y los otros dos tintes distinguibles espectralmente se pueden excitar en otra longitud de onda. Las longitudes de onda de excitación particular son 532 nm, 630 nm a 700 nm, particularmente 660 nm.

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

En una realización adicional, dicho kit comprende otros dos nucleótidos modificados marcados con tintes fluorescentes en el que dichos tintes se excitan por el mismo láser a 600 nm a 700 nm, particularmente 630 nm a 700 nm, más particularmente 660 nm. En el que los tintes tienen una diferencia en la absorbancia máxima de por lo menos 10 nm, particularmente 20 nm a 50 nm. Más particularmente los dos compuestos de tinte tienen cambios de Stokes de entre 20 a 40 nm. Aún más particularmente los dos compuestos de tinte tienen una absorbancia diferente máxima por encima de 600 nm, particularmente por encima de 640 nm. Tintes particulares que se pueden distinguir espectralmente de tintes rodamina de la invención y que cumplen el criterio anterior son análogos de polimetina como se describe en la Patente de estadounidense No. 5,268,486 (por ejemplo Cy5) o el documento WO 0226891 (Alexa 647; Molecular Probes A20106) o polimetinas no simétricas como se divulga en la Patente estadounidense No 6,924,372.

En una realización alternativa, los kits de la invención pueden contener nucleótidos, en donde la misma base se marca con dos compuestos diferentes. Un primer nucleótido se puede marcar con un compuesto de la invención. Un segundo nucleótido se puede marcar con un compuesto espectralmente distinto, por ejemplo un tinte "rojo" que absorbe más de 600 nm. Un tercer nucleótido se puede marcar como una mezcla de los compuestos de la invención y el compuesto espectralmente distinto, y el cuarto nucleótido puede ser "oscuro" y no contener marcador. Por lo tanto en términos simples, los nucleótidos 1-4 se pueden marcar "verde", "rojo", "rojo/verde", y oscuro. Para simplificar la instrumentación adicional, se pueden marcar cuatro nucleótidos con dos tintes excitados con un único láser, y de esta manera marcar los nucleótidos 1-4 "verde 1", "verde 2" "verde1/verde 2", y oscuro.

Los nucleótidos pueden contener dos tintes de la presente invención. Tintes en donde  $R_1$  y  $R_4$  son H absorbe en una menor longitud de onda que cuando  $R_1$  y  $R_4$  son alquilo. Un kit puede contener dos o más nucleótidos marcados con tintes de la invención. Un kit puede contener un nucleótido marcado con un compuesto de la invención en donde  $R_1$  y  $R_4$  son H, y un segundo nucleótido marcado con un compuesto de la invención en donde  $R_1$  y  $R_4$  son alquilo. Los kits pueden contener un nucleótido adicional en donde una parte del nucleótido se marca con un compuesto de la invención en donde  $R_1$  y  $R_4$  son H, y una segunda porción del nucleótido marcado con un compuesto de la invención en donde  $R_1$  y  $R_4$  son alquilo. Los kits pueden contener adicionalmente un nucleótido no marcado.

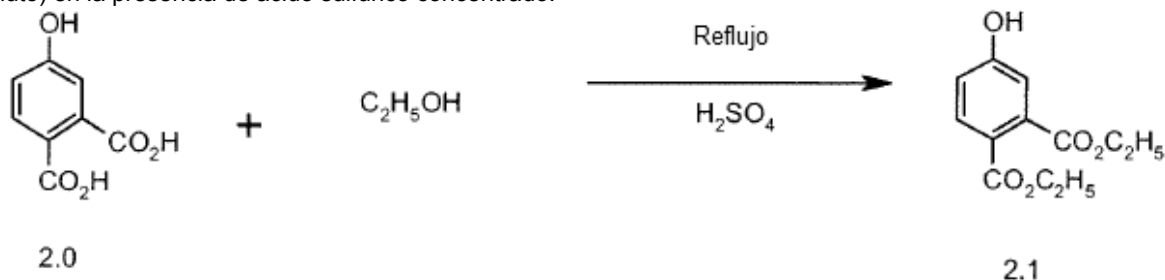
En una realización, los kits pueden incluir una enzima de polimerasa capaz de catalizar la incorporación de los nucleótidos modificados en un polinucleótido. Otros componentes que se van incluir en dichos kits pueden incluir reguladores y similares. Los nucleótidos modificados marcados con tintes de acuerdo con la invención, y cualesquier otros componentes de nucleótidos que incluyen mezclas de diferentes nucleótidos, se pueden proporcionar en el kit en una forma concentrada que se va a diluir antes de uso. En dichas realizaciones, también se puede incluir un regulador de dilución adecuado.

Cabe señalar, como se utiliza en esta especificación y las reivindicaciones adjuntas, las formas singulares "un", "uno" y "el" incluyen plurales referentes a menos que se limite expresamente e inequívocamente a un referente.

Detalles de Experimentos

4-hidroxifalato de dietilo (2.1).

Este compuesto se preparó a partir de ácido 4-hidroxi-ftálico 2.0 mediante esterificación con un exceso de etanol (absoluto) en la presencia de ácido sulfúrico concentrado.



5

Preparación:

A un matraz de fondo redondo de 0.5 L que contiene etanol absoluto (250 ml, ~200 g ~4.3 mol) ácido 4-hidroxifáltico 2.0 (15 g, 82.4 mmol) se agregó en algunas porciones a temperatura ambiente con agitación. Esto luego se siguió mediante adición de 0.5 ml (~0.92 g, 9.4 mmol) de ácido sulfúrico concentrado. Esta mezcla se agitó a temperatura ambiente durante media hora, y luego se sometió a reflujo adicional durante 24 horas con un condensador.

10

15

Aproximadamente la mitad del volumen del solvente se destiló utilizando un condensador descendente y el solvente restante se eliminó bajo una presión reducida. El residuo aceitoso viscoso se disolvió en 400 ml de DCM. A esta solución, se agregó 1.5 g de carbonato de potasio anhidro cuidadosamente en pequeñas porciones. Esta mezcla se mantiene a temperatura ambiente con agitación durante aproximadamente 2 horas, luego se filtró a través de ~100 g de geles de sílice. Los geles de sílice se lavaron adicionalmente con DCM para eluir más productos absorbidos. Las fracciones orgánicas combinadas se secaron con sulfato de magnesio anhidro (~5 g) durante la noche. Los sólidos se filtraron, y se lavaron adicionalmente con DCM. El solvente se evaporó bajo una presión reducida.

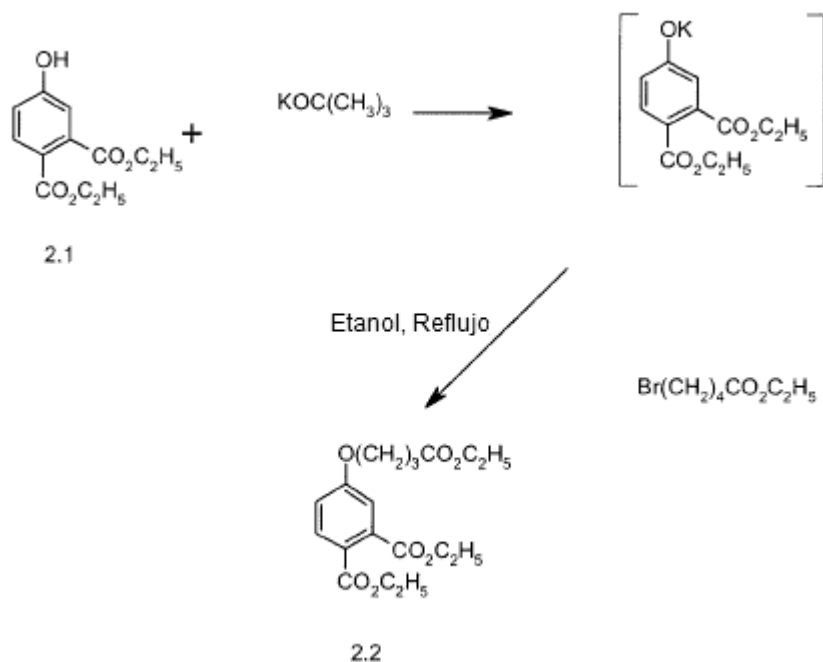
20

Este éster 2.1 obtenido como producto aceitoso incoloro se utilizó en la siguiente etapa sin purificación adicional. Rendimiento: 22 g (~99 %)  $^1\text{H}$  RMN: (400 MHz, DMSO- $d_6$ )  $\delta$  10.59 (s, 1H), 7.68 (d, J = 8.5 Hz, 1H), 7.07 - 6.76 (m, 2H), 4.22 (dq, J = 13.4, 7.1 Hz, 4H), 1.25 (td, J = 7.1, 3.3 Hz, 6H).

25

4-[(3-etoxicarbonil)propiloxi]ftalato de dietilo (2.2).

Este compuesto se preparó en etanol a partir de 4-hidroxifalato de dietilo 2.1 a través de formación anterior de 3,4-bis(etoxicarbonil)fenolato de potasio y luego mediante alquilación con 4-bromobutirato de etilo en el mismo crisol de reacción.



30

Preparación:

A una solución del éster 2.1 (2 g, 8.4 mmol) en etanol absoluto (20 ml), tert-butóxido de potasio sólido (1.1 g, 9.8 mmol,

1.15 eq) se agregó lentamente. La mezcla de reacción se mantiene a temperatura ambiente durante una hora, y se agregó luego 4-bromobutirato de etilo (2 g, ~1.5 ml). La reacción se mantiene a temperatura ambiente durante una hora, y luego se calentó bajo reflujo con un condensador durante la noche.

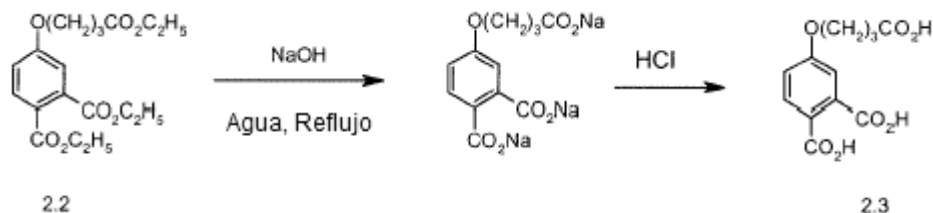
5 La mezcla de reacción se congeló (hasta aproximadamente 0-5°C) con un baño de hielo durante 3 a 4 horas. Los precipitados inorgánicos se filtraron y adicionalmente se lavaron con etanol absoluto frío. Los precipitados se descargaron y los filtrados combinados se concentraron bajo presión reducida.

10 El residuo que contiene algunos sólidos inorgánicos se diluyó con una mezcla de DCM-Éter de petróleo (7:3, 50 ml) y se filtró a través de un tapón corto de geles de sílice (~10 g). Los geles de sílice se lavaron adicionalmente con DCM-MeOH (97:3, ~150 ml) para eluir más producto absorbido (se utilizó TLC para monitorizar la terminación de elución del producto; se utiliza más solvente eluyente si se desea). Las fracciones orgánicas combinadas se secaron con sulfato de magnesio anhidro (~1 g) durante la noche y se evaporaron bajo una presión reducida. Este éster 2.2 se utilizó en la siguiente etapa sin purificación adicional. Rendimiento: 2 g (67.6 %)

15 <sup>1</sup>H RMN: (400 MHz, Metanol-d<sub>4</sub>) δ 7.87 - 7.77 (m, 1H), 7.18 - 7.05 (m, 2H), 4.33 (dq, J = 14.5, 7.1 Hz, 4H), 4.21 - 4.07 (m, 4H), 2.53 (t, J = 7.3 Hz, 2H), 2.17 - 2.04 (m, 2H), 1.36 (td, J = 7.1, 3.5 Hz, 6H), 1.25 (t, J = 7.2 Hz, 3H).

20 Ácido 4-(3-Carboxipropiloxi)ftálico (2.3).

Este derivado de ácido ftálico se preparó mediante una saponificación del éster de trietilo 2.2 de la etapa previa mediante una solución acuosa de hidróxido de sodio.



Preparación:

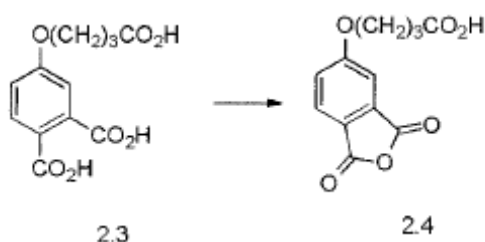
25 A un matraz de fondo redondo, equipado con un agitador y un condensador, se vertió 100 ml de agua deionizada y luego se agregaron NaOH (5 g, 125 mmol) en pequeñas porciones. Después se disolvió hidróxido de sodio, se agregó el éster de trietilo 2.2 (10.5 g, ~30 mmol). La mezcla de reacción se agitó vigorosamente a 110°C (bloque de calentamiento) durante 5 horas. Después de aproximadamente 3 horas, el material de partida oleoso 2.2 se disolvió. La matraz de fondo redondo se filtró y el filtrado recolectado se transfirió a un matraz de fondo redondo. Aproximadamente 30 10 a 15 ml de líquido que contiene etanol formado durante la reacción y agua, se destiló con un condensador descendente. La solución restante se enfrió mediante baño de hielo (0 a 5°C).

35 A esta solución fría, se agregó lentamente ácido clorhídrico concentrado (37%, 11 ml, 135 mmol con agitación). El pH de esta mezcla se verificó mediante un indicador para asegurar un exceso de ácido mineral. El producto esperado como un precipitado blanco se formó poco después de esto. La mezcla se mantiene a ~0-5°C durante una hora. El precipitado se filtró, se lavó con unos pocos ml de agua fría y luego se secó durante la noche. Este producto 2.3 se utilizó en la siguiente etapa sin purificación adicional. Sin embargo, se puede purificar mediante cristalización a partir de agua si es necesario. Rendimiento: 5.1 g, 64 %.

40 <sup>1</sup>H-RMN : (400 MHz, CF<sub>3</sub>COO-d) δ 11.62 (d, J = 1.5 Hz, 6H), 8.10 (dd, J = 8.8, 1.5 Hz, 1H), 7.36 (d, J = 2.4 Hz, 1H), 7.23 - 7.14 (m, 1H), 4.24 (td, J = 5.9, 1.6 Hz, 2H), 2.76 (td, J = 7.2, 1.6 Hz, 2H), 2.35 - 2.24 (m, 2H).

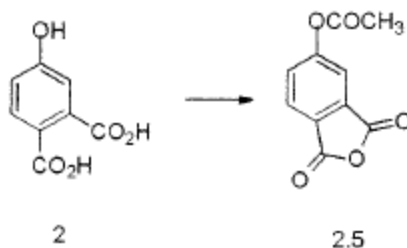
45 Anhídrido 4-(3-Carboxipropiloxi)ftálico (2.4).

La preparación del anhídrido 2.4 se puede lograr mediante del ácido 2.3 con reactivos de deshidratación similares a anhídrido acético, precalentamiento del ácido 2.3 a alta temperaturas o solo al mantener el ácido 2.3 en un disector con un reactivo de deshidratación inorgánico.



50 Anhídrido 4-Acetiloxiftálico (2.5).

Este compuesto se preparó a partir de ácido 4-hidroxi-ftálico 2.0 mediante deshidratación con un exceso de anhídrido acético de acuerdo con el procedimiento publicado en la bibliografía (Synthesis 2008, p.3415).

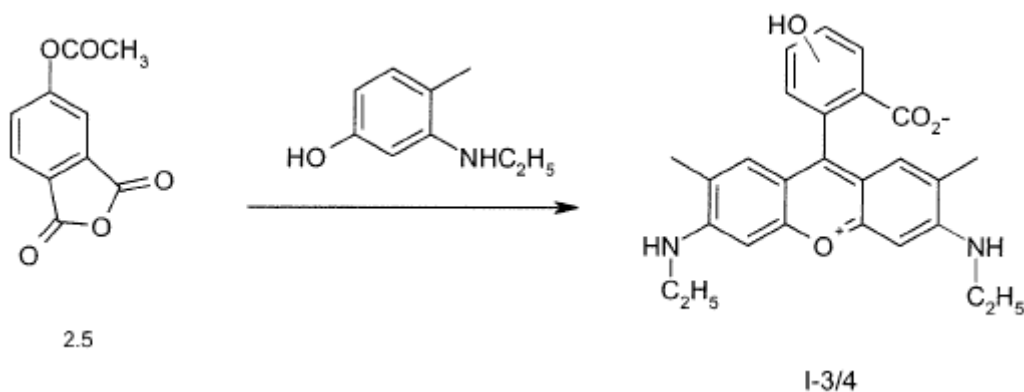


5

A una solución de ácido 4-hidroxiftálico 2 (5.39 g) en tolueno (100 ml) se agregó anhídrido acético (25 ml) y la mezcla se sometió a reflujo durante 2 h. Al día siguiente se filtró el producto cristalizado. Rendimiento 2.6 g (42.5 %).

La síntesis de tinte (XI-1):

10



15

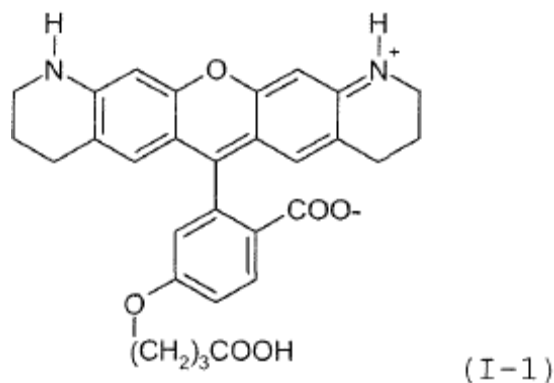
A un matraz de fondo redondo, se agregó ácido 4-acetiloftálico (1.03 g), 3-etilamino-cresol (2.0 g) y piro-sulfato de potasio (1.5 g). Los reactivos se mezclaron cuidadosamente junto con una espátula y se calentaron a 150°C (bloque de calentamiento) durante 6 h. Durante este tiempo se desarrolló más y más color rojo intenso y se solidificó la mezcla de reacción. Se utilizaron TLC (acetonitrilo-agua, 20%) y HPLC (@ 280 nm) para certificar el progreso de la reacción al monitorizar la reacción de tintes y el 3-etilaminocresol de partida. Después de aproximadamente 5 horas la relación se mantuvo casi constante. A la mezcla de reacción se agregó 10 ml de agua y se filtró el producto de color rosado. El precipitado se colocó en un matraz de fondo redondo, se agregó metanol (50 ml) y la mezcla se sometió a reflujo durante 10 min. Aproximadamente  $\frac{3}{4}$  del solvente se eliminó en vacío y esta solución de color rojo se dejó durante la noche. Se filtró el producto como cristales rojos. Rendimiento 1 g (46 %).

20

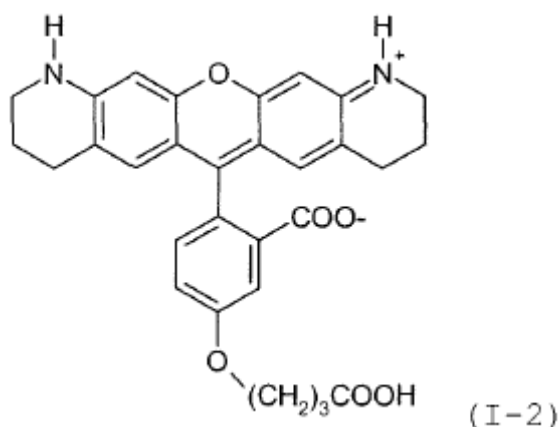
25

Si se necesitan se pueden separar isómeros mediante cromatografía.

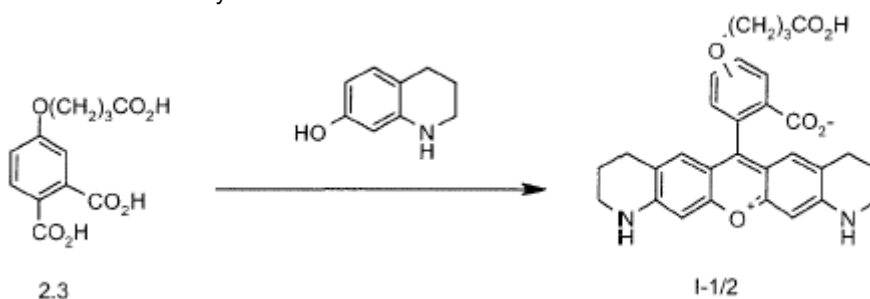
La síntesis de tinte (I):







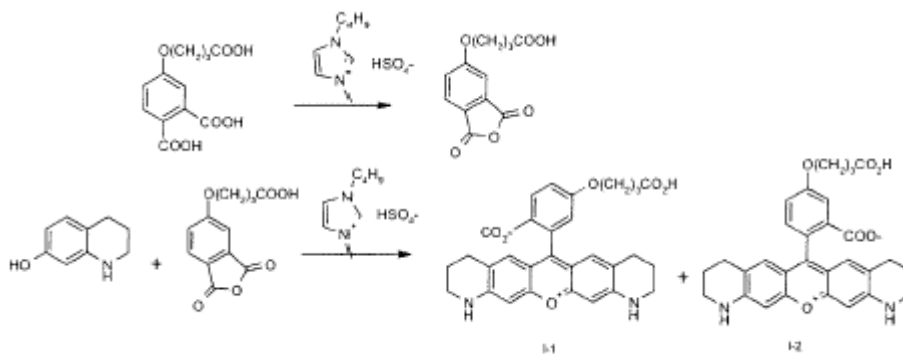
5 Se sintetizaron compuestos objetivo como una mezcla de isómeros constitucionales mediante una condensación del ácido ftálico 2.3 con 7-hidroxi-1,2,3,4-tetrahydroquinolina a alta temperatura. Se completó la reacción al utilizar ácido fosfórico anhidro como un catalizador y también como solvente



Tinte (I-1) y (I-2)

Esquema de reacción:

10



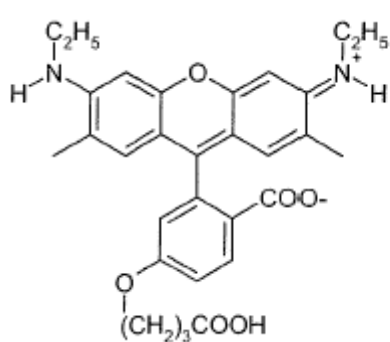
Preparación:

15 Ácido 4-(3-carboxipropoxi)ftálico (0.268 g) y 0.5 g de hidrosulfato de 1-butil-3-metil-imidazolio (IL-HSO<sub>4</sub>) se calentó 1 h a 125°C para completar la formación de anhídrido (Xa-2.4). Se agregó 7-Hidroxi-tetrahydro-dihidroquinolina (0.298 g) y la mezcla de reacción se agitó a 175°C durante 2 h. Se formó compuesto de color rojo con fuerte fluorescencia. El control de TLC en CH<sub>3</sub>CN-agua (4:1) confirmó formación de tintes. Se continuó calentamiento a 175°C durante 3 h adicionales.

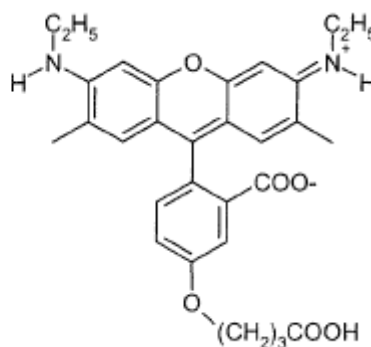
20 Este material crudo se disolvió en mezcla de CH<sub>3</sub>CN-agua (10%) y se purificó mediante cromatografía de columna flash sobre gel de sílice. Se recolectaron fracciones de color rojo y los solventes se evaporaron.

Tinte (I-1): Rendimiento 18 mg (16 %).

25 Tinte (I-2): Rendimiento 9 mg (8 %).

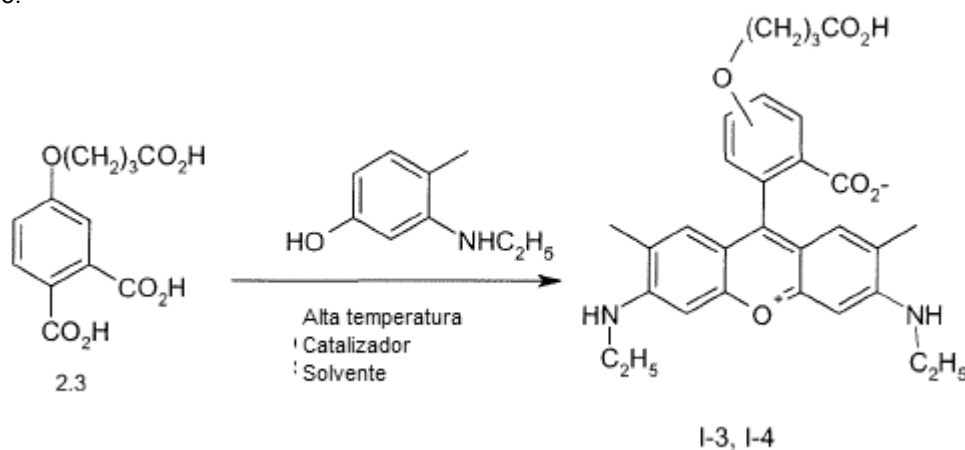


(I-3)



(I-4)

5 Se sintetizaron compuestos objetivo como una mezcla de isómeros constitucionales mediante condensación del ácido ftálico 2.3 con 3-etilamino-4-metil-fenol a alta temperatura. Se completó la reacción al utilizar ácido fosfórico anhidro como un catalizador con y sin solvente orgánico polar de alto punto de ebullición como DMF, DMA, sulfolano o 1,2-diclorobenceno.



10 Esta reacción se puede llevar a cabo utilizando disulfato de potasio como un catalizador. Este reactivo actúa como un ácido de Lewis suave, que puede estimular la formación y/o reacciones adicionales de los materiales de partida y/o intermedios formados. Como resultado, la temperatura de reacción fue menor y se redujo la cantidad de productos secundarios. La naturaleza de un catalizador y un disolvente que se han utilizado ejercen un efecto significativo no sólo sobre el rendimiento y pureza de la reacción de los productos, sino también sobre la regioselectividad de la formación de tintes de isómeros. De este modo se puede conseguir la formación de uno u otro isómero como uno prevalente.

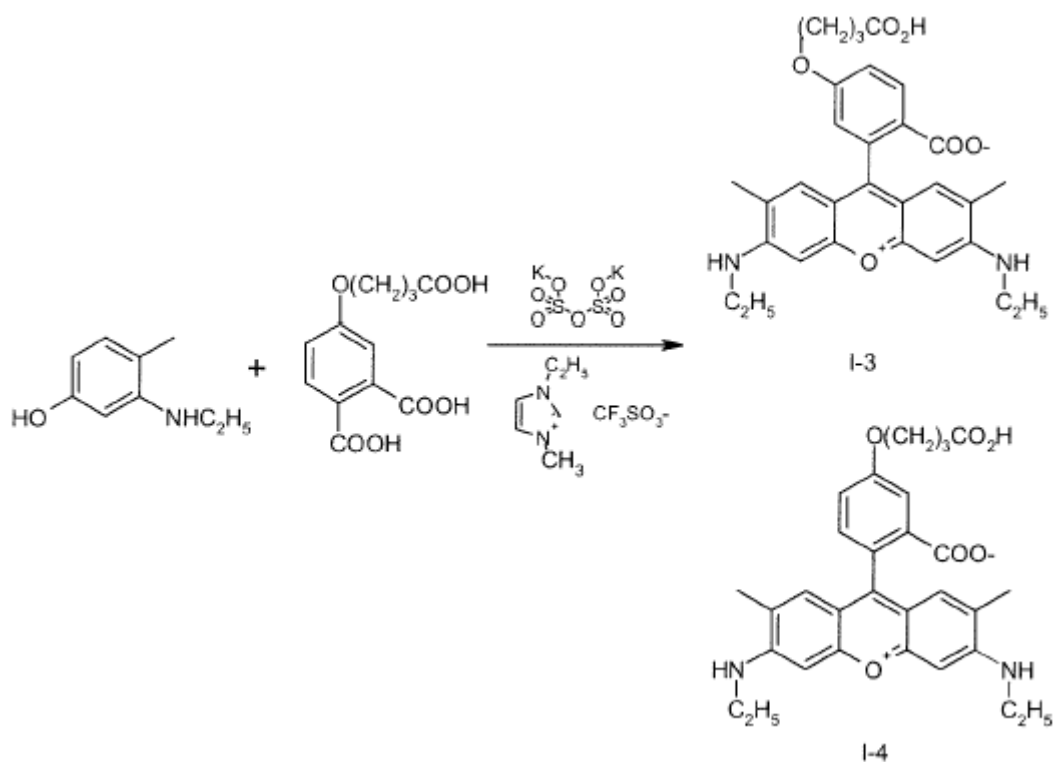
15 Adicionalmente, algunos líquidos iónicos (IL) pueden mejorar la formación de tinte de rodamina tanto en el rendimiento de reacción como en la pureza del producto. También hemos logrado una mejora definitiva con algunos IL sobre la síntesis de tintes (I). Entre la familia de sales cuaternarias de IL (1-etil-3-metilimidazolio y algunas sales de tetraalquilamonio) que probamos, algunas promueven eficientemente la formación de uno u otro isómero y proporcionan un producto más limpio (menos subproductos coloreados en la mezcla de reacción).

Tintes (1-3) y (1-4)

Nombre químico:

25 (1-3): betaína de 3,6-Bis(etilamino)-2,7-dimetil-[2-carboxilato-5-(3-carboxipropiloxi)fenil]xantilio(1-4): betaína de 3,6-Bis(etilamino)-2,7-dimetil-[2-carboxilato-4-(3-carboxipropiloxi)fenil]xantilio

30 Esquema de reacción



## Preparación:

5 A un matraz de fondo redondo, se agregó ácido 4-carboxipropiloxiftálico (1.0 g, 3.73 mmol), 3-etilamino-cresol (2.0 g, 13.23 mmol), IL-CF3 (1.72 g, 6.61 mmol) y piro-sulfato de potasio (1.68 g, 6.61 mmol). Los reactivos se mezclaron cuidadosamente junto con una espátula y se calentaron a 150°C (bloque de calentamiento) durante 10 h. Durante este tiempo se desarrolló más y más color rojo intenso y se solidificó la mezcla de reacción. Se utilizaron ambos TLC (acetonitrilo-agua, 20%) y HPLC (@ 280 nm) para verificar el progreso de reacción al monitorizar la relación de tintes y el 3-etilaminocresol de partida. Después de 10 horas la relación se mantuvo casi constante.

10 La mezcla de reacción se enfrió a temperatura ambiente, y el sólido rojo se disolvió en metanol (~20 ml). Los materiales inorgánicos insolubles incoloros se filtraron. Se agregó trietilamina (2 ml) al filtrado y la solución resultante se aplicó a una muestra Biotage C<sub>18</sub> (40 g). La muestra luego se seca en vacío para eliminar la mayoría de solventes.

15 El producto se aisló mediante cromatografía de destello en un instrumento Biotage Isolera-4 en una columna 400 g C<sub>18</sub> utilizando mezcla de TEAB 0.1 M en agua y acetonitrilo como solventes de elución (gradiente 17-27). Se estableció la longitud de onda de recolección de producto a 520 nm, y la longitud de onda de control se estableció a nm para monitorizar la separación del exceso de los materiales de partida y productos secundarios incoloros formados.

20 Se recolectaron fracciones de color naranja-amarillo. Los solventes se eliminaron mediante un evaporador rotativo en vacío. Se trituró el residuo rojo sólido restante con éter de petróleo (60-95°C, 25 ml) y producto, I-3 como un polvo rojo se filtró y se secó sobre aire.

25 Rendimiento: 0.50 g (29%).

(I-3):

<sup>1</sup>H-RMN:

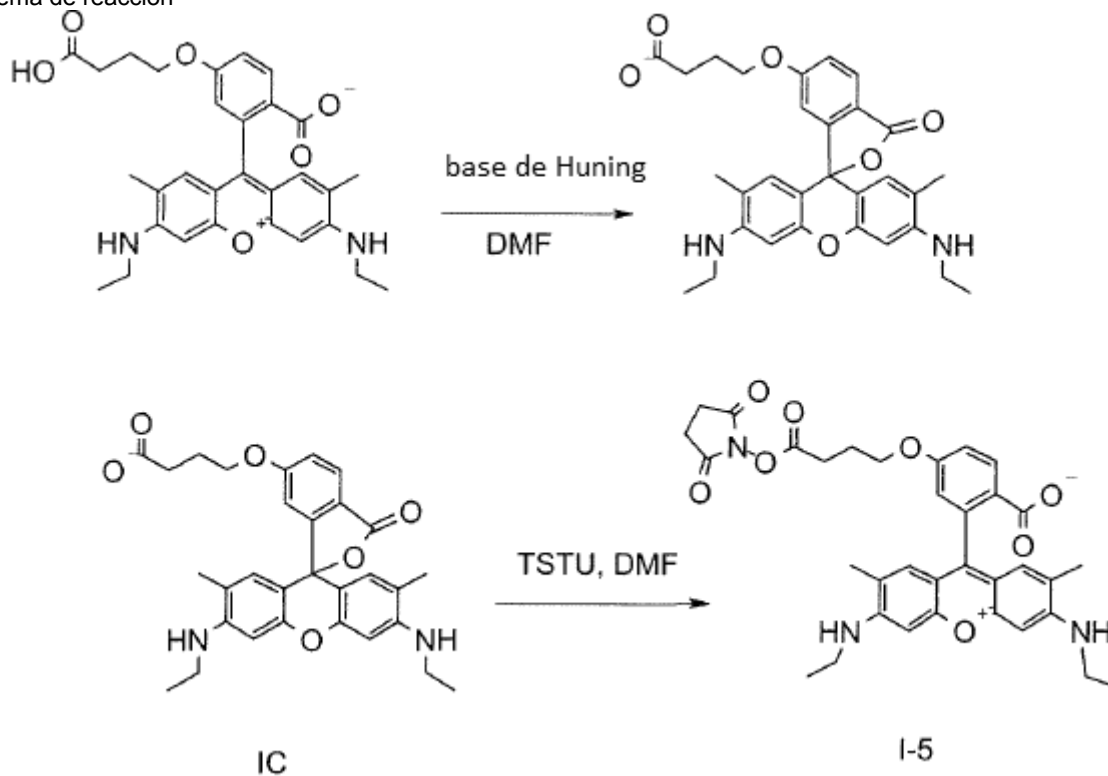
30 (400 MHz, DMSO-d<sub>6</sub>) 5.784 (d, J = 8.5 Hz, 1H), 7.18 (dd, J = 8.5, 2.2 Hz, 1H), 6.58 (d, J = 2.2 Hz, 1H), 6.30 (s, 2H), 6.25 (s, 2H), 5.26 (t, J = 5.4 Hz, 2H), 3.95 (t, J = 6.4 Hz, 2H), 3.22 - 3.10 (m, 4H), 2.30 (d, J = 7.3 Hz, 2H), 1.91 (s, 8H), 1.21 (t, J = 7.1 Hz, 6H).

35 (I-4) :

Rendimiento: 0.16 g (9%).

Tinte (I-5) I-3-NHS

Esquema de reacción



Preparación:

5

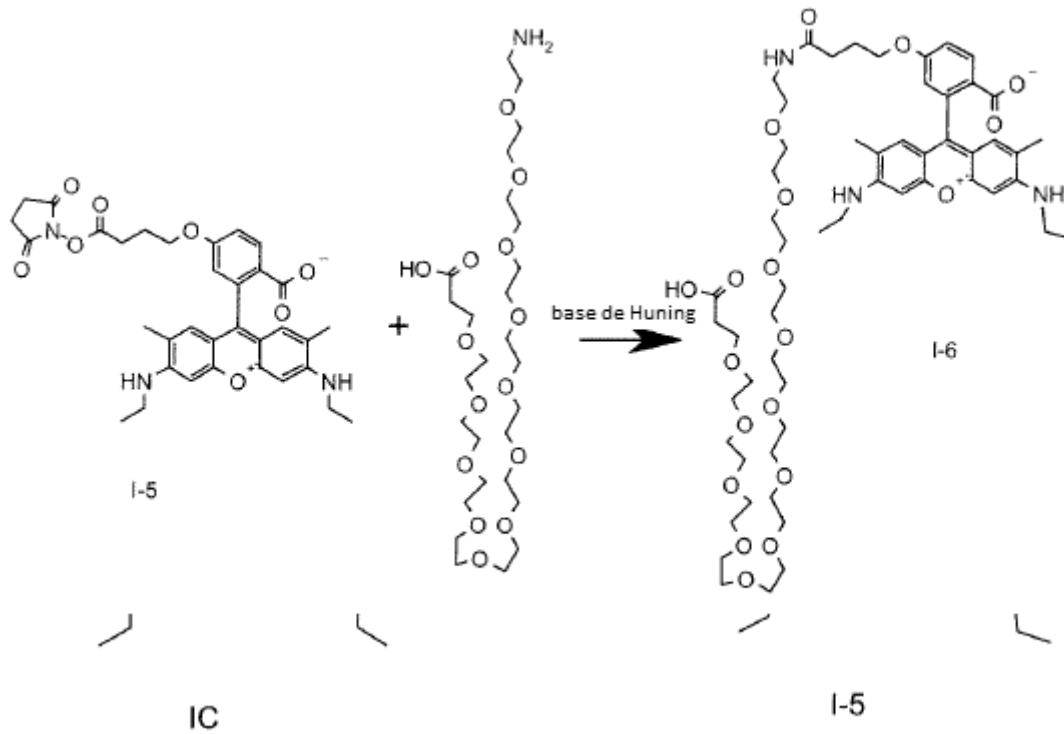
El tinte (I-3) se secó bajo alto vacío durante la noche luego se tomó 35 mg de muestra (68  $\mu\text{mol}$ ) y se colocó en matraz de fondo redondo. Se agregó DMF anhidro (3 mL) y base de Hunig (678  $\mu\text{mol}$ , 116  $\mu\text{L}$ ) al matraz. La mezcla (se hace cada vez más incolora debido a la formación de derivado cíclico (IC) y se agitó durante aproximadamente 20 min bajo atmósfera de nitrógeno. Luego se agregó TSTU (80  $\mu\text{mol}$ , 25 mg). Se desarrolló un color rosado claro. La mezcla de reacción se agitó a temperatura ambiente durante 1h, Se verificó el progreso mediante TLC, (20% en  $\text{H}_2\text{O}$  en  $\text{CH}_3\text{CN}$ ).

10 Esta lactona se puede aislar mediante precipitación de la solución, o solo destilación del solvente en vacío. El éster de NHS (I-5) preparado de esta forma se utilizó para acoplamiento sin purificación adicional y aislamiento de la solución. Si se necesita el éster de NHS se puede aislar de la solución mediante precipitación, por ejemplo, mediante precipitación con acetato de etilo, se filtró y se secó.

15

Tinte (I-6) I-3-PEG12

Esquema de reacción:

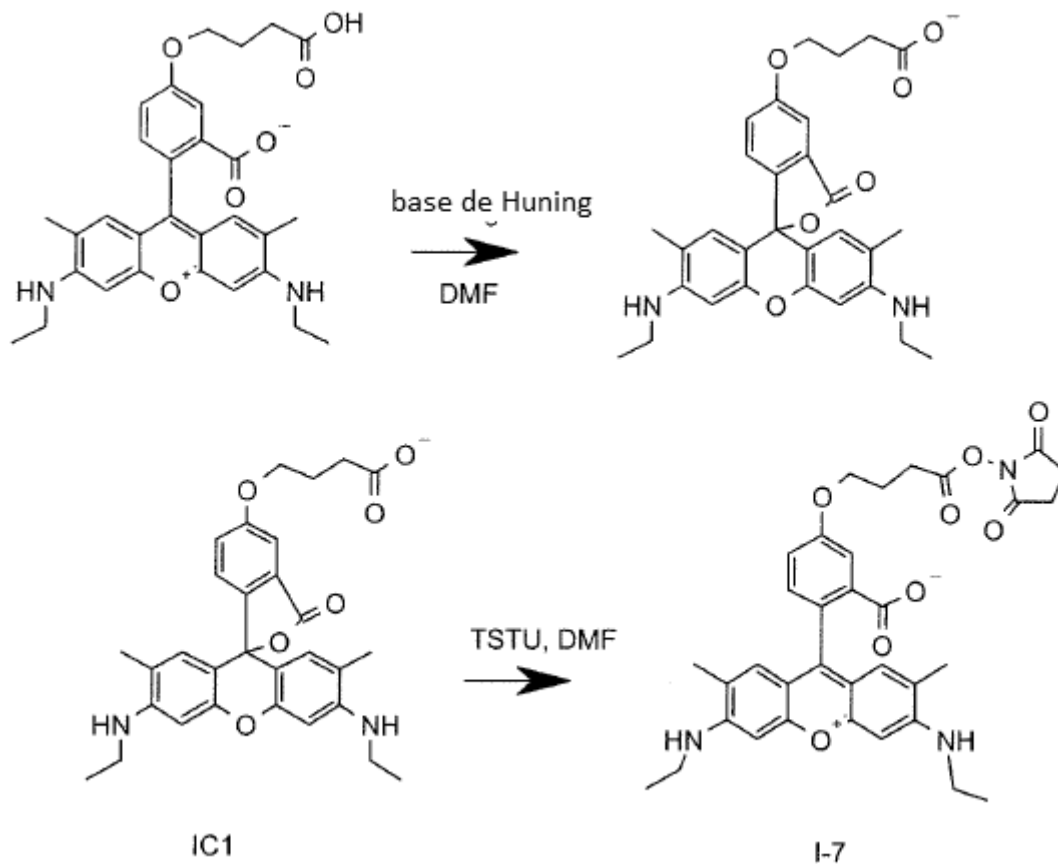


## Preparación:

5 Una solución de  $\text{NH}_2\text{Peg}12\text{COOH}$  (203  $\mu\text{mol}$ , 126 mg) en agua (300  $\mu\text{L}$ ) se agregó a la solución de (I-5) preparada como se describió anteriormente y la mezcla de reacción se agitó a temperatura ambiente durante la noche. Se verifica la terminación de la reacción mediante TLC (20% en  $\text{H}_2\text{O}$  en acetonitrilo). Después de la terminación del acoplamiento, se agregó solución de TEAB 0.1M en agua (4 mL) y la mezcla se agitó a  $\sim 20^\circ\text{C}$  durante 1 h para detener los intermedios reactivos. Los solventes se eliminaron bajo vacío. El residuo se disolvió en 10 ml de TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa 0.2  $\mu\text{m}$  de tamaño de poro y se transfirió en dos frascos de 5 mL de HPLC para purificación por HPLC. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo-TEAB 0.1 M. se recolectaron las fracciones con absorción máxima 526 nm. Rendimiento 47  $\mu\text{mol}$  (70%).

Tinte (I-7) I-4-NHS

15 Esquema de reacción



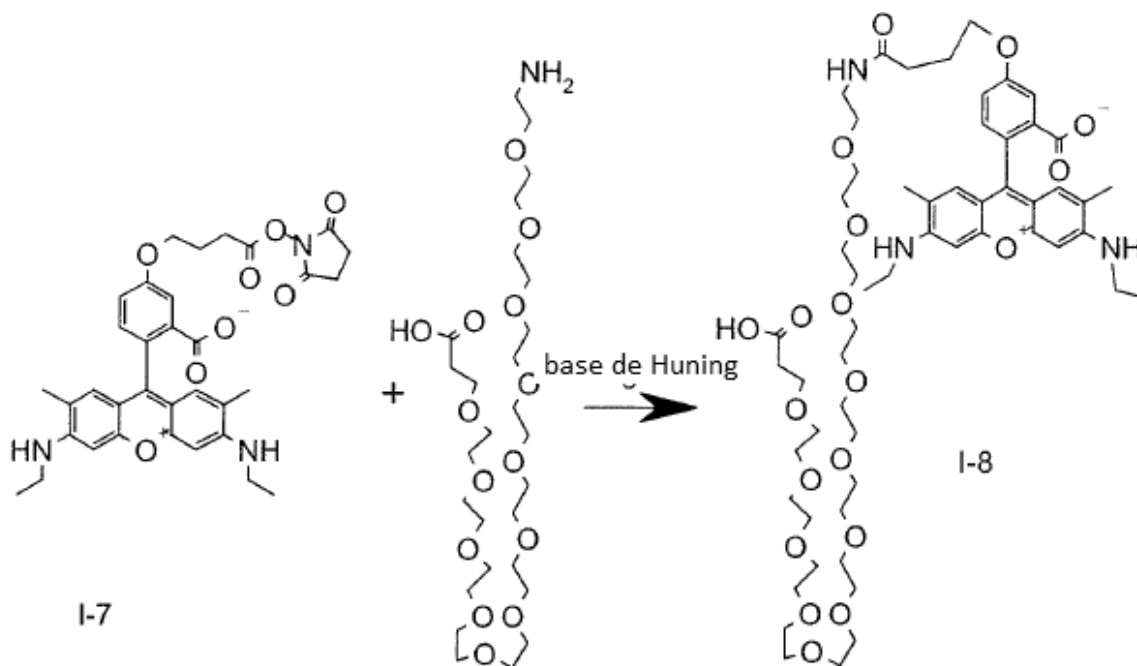
Preparación:

- 5 El tinte (1-4) se secó bajo alto vacío durante la noche luego se tomó 28 mg de muestra. Se agregó DMF anhidro (3 mL) y base de Hunig al matraz. La mezcla se hizo más incolora. Se agitó la solución de forma cíclica durante aproximadamente 20 min bajo atmósfera de nitrógeno. Luego se agregó TSTU (25 mg). Se desarrolló color rosado claro. La mezcla de reacción se agitó a temperatura ambiente durante 1h. El progreso se verificó mediante TLC, (20% en H<sub>2</sub>O en MeCN). NHS-éster R<sub>f</sub> ~0.7. Se completa la actividad en 2h.
- 10 El NHS éster (I-7) preparado de esta forma se utilizó para acoplamiento sin purificación adicional y aislamiento de la solución. Si se necesita el NHS éster se puede aislar de la solución mediante precipitación con acetato de etilo, filtrar y secar.

Tinte (I-8)I-4-PEG12

15

Esquema de reacción:



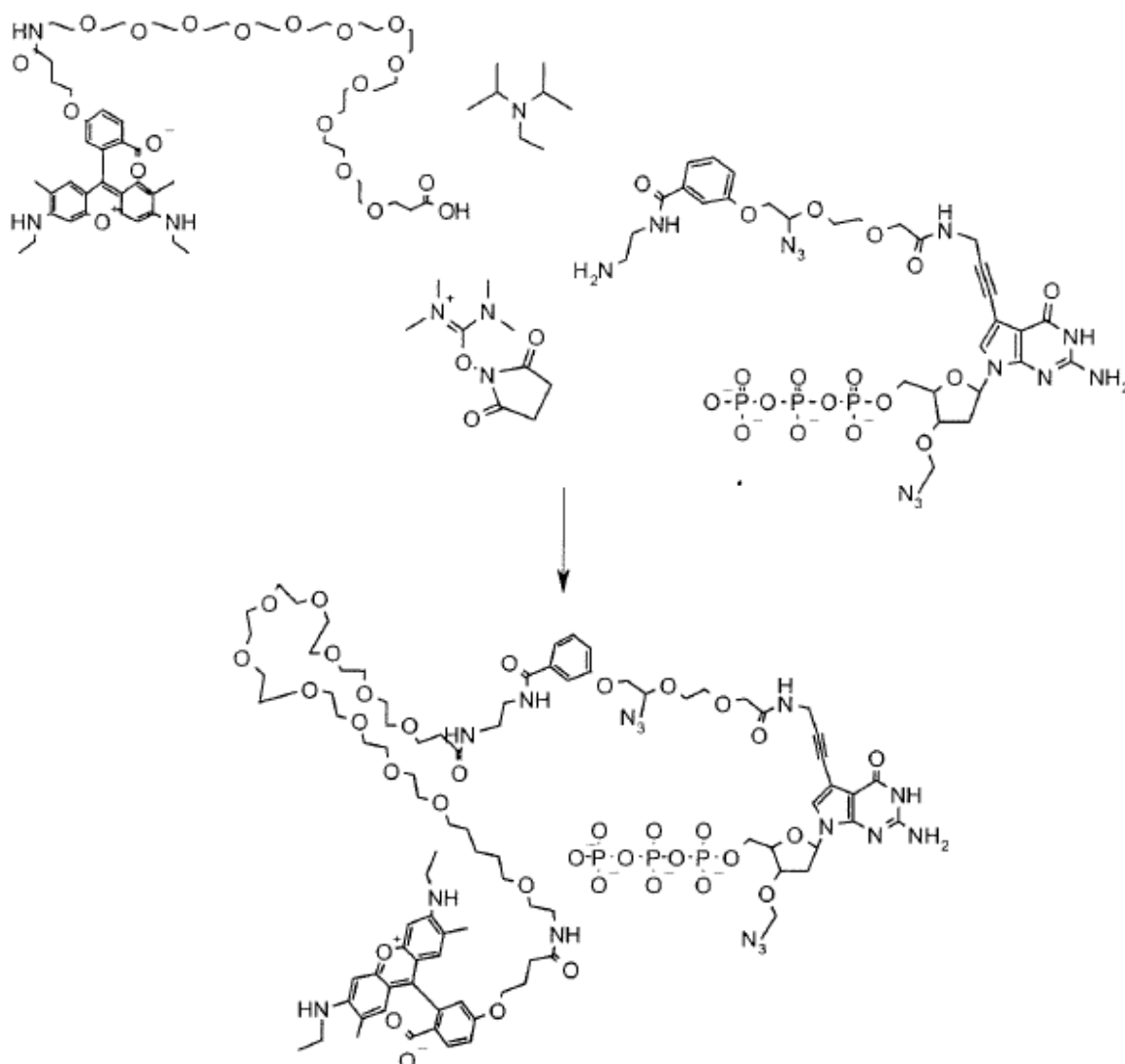
## Preparación:

5 Una solución de NH<sub>2</sub>Peg12COOH (100 mg) en agua (300 μL) se agregó a la solución de (I-7) preparada como se describió anteriormente y esta mezcla de reacción se agitó a temperatura ambiente durante la noche. La terminación de la reacción (consumo de NHS éster) luego se verifica mediante TLC (20% en H<sub>2</sub>O en acetonitrilo, la placa se secó en vacío). Para tratamiento final se agregó 0.1 TEAB 1 M en agua (4 mL) y la mezcla se agitó a ~20°C durante 1 h. Los solventes se eliminaron bajo vacío. El residuo se disolvió en 10 ml de TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa de 0.2 μm de tamaño de poro y se transfirió en dos frascos de HPLC de 5 mL para purificación mediante HPLC. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo TEAB 0.1 M. Se recolectaron fracciones con absorción máxima 520 nm. Rendimiento 41 μmol (76 %).

Tinte (I-9) pppG-I-3-PEG12

15

Esquema de Reacción:



## Preparación:

5 DMA anhidro (7 mL) y base de Hunig (0.082 mL) se agregaron a la muestra seca (52 mg) de (I-6). Luego se agregó una solución de TSTU (17 mg) en 1 mL de DMA seco. Se enjuagó el sistema dos veces con nitrógeno y luego la mezcla de reacción se agitó a temperatura ambiente durante 1h, de acuerdo con TLC (20% en H<sub>2</sub>O en CH<sub>3</sub>CN) se completó la activación. Una vez se inició la activación una solución de pppG-LN3 (3 mL de solución madre 3.5 mM) se vació hasta secado. Después de que se completó la activación esta solución se agregó a pppG. La reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. Control de TLC en 20% en H<sub>2</sub>O en acetonitrilo. La mezcla de  
10 reacción se enfrió a ~4°C con un baño de hielo, luego se agregó TEAB 0.1 M (4 mL) y la mezcla se agitó a temperatura ambiente durante 10 min.

Purificación: La solución se aplicó a columna con ~20 g de de suspensión de resina DEAE sephadex en solución de TEAB 0.05 M y se lavó con TEAB de 0.1 M hasta 0.4 M. se combinaron las fracciones coloreadas, se coevaporaron con agua para eliminar más TEAB y se vació hasta secado.  
15

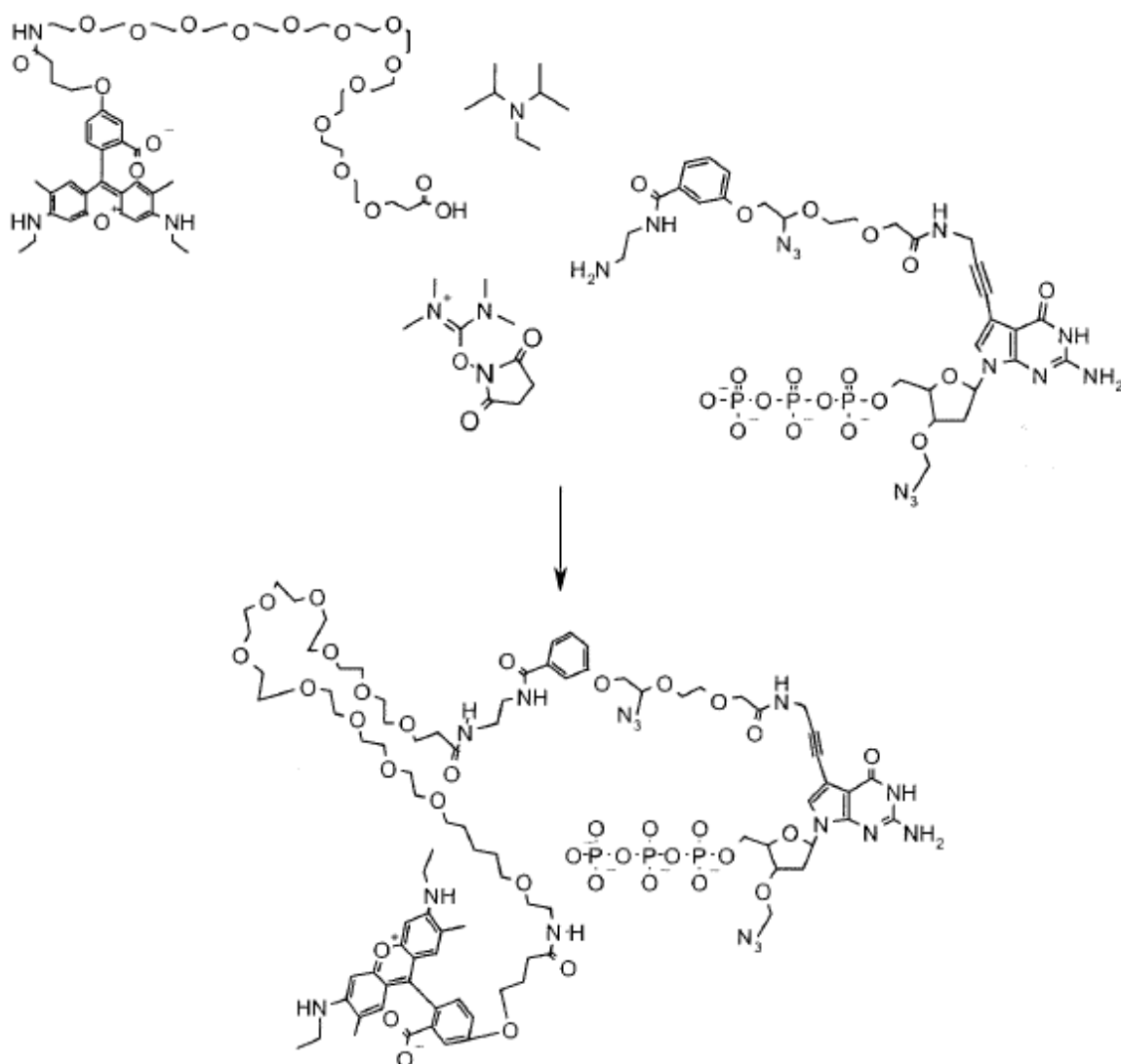
El residuo se redisolvió luego en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa 0.2 nm de tamaño de poro en un matraz coming, y luego se transfirió a frascos de alta recuperación de 5 mL de HPLC para purificación de HPLC. La solución se puede almacenar en el congelador hasta purificación. El producto se purificó mediante HPLC utilizando una columna de fase inversa C18 con acetonitrilo - TEAB 0.1 M como eluyentes, Se recolectaron fracciones con absorción 525 nm. Rendimiento: 54 %.  
20

Tinte (I-10) pppG-I-4-PEG12

25

Esquema de reacción:

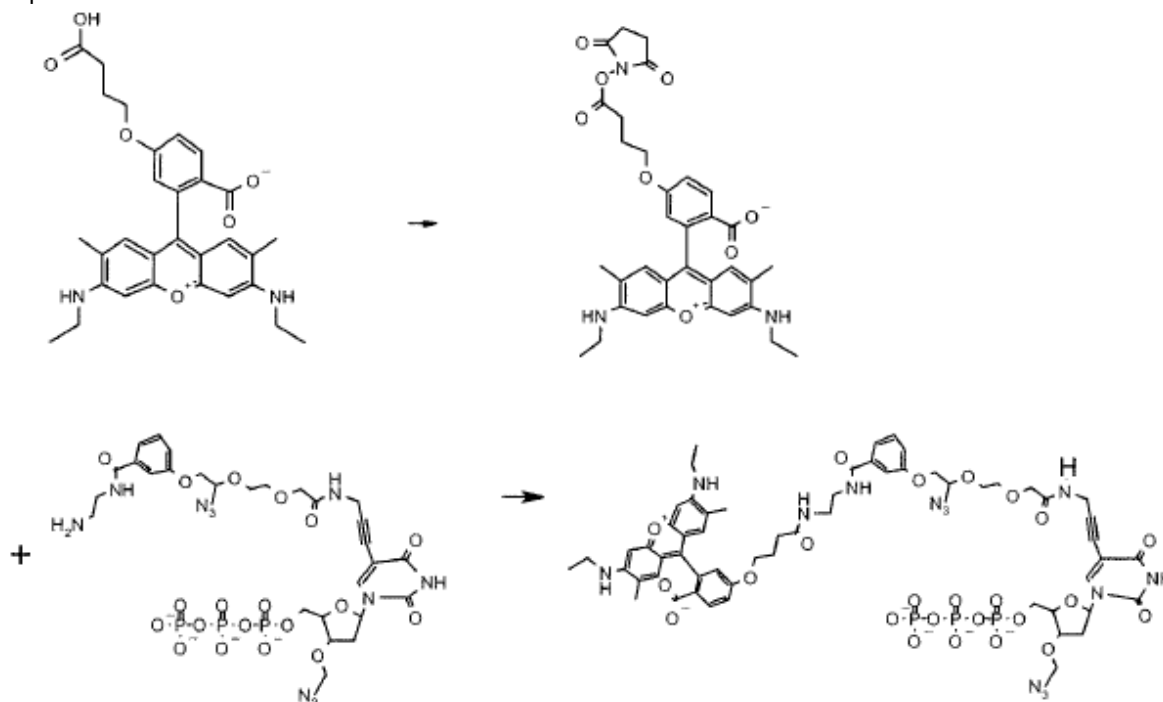




## Preparación:

- 5 El tinte (I-8) se secó después de la segunda purificación HPLC prep (3 mL) y se disolvió en DMA anhidro (3 mL) y luego se agregó la base de Hunig (0.036 g). Una solución de TSTU (11 mg) en 1 mL de DMA seco se agregó posteriormente. La mezcla de reacción se agitó a temperatura ambiente durante 3 h.
- 10 De acuerdo con TLC (20% en H<sub>2</sub>O en CH<sub>3</sub>CN) se contempló la activación.
- 15 Una solución de pppG-LN3 (1.6 mL de stock 35.5 mM) se extrajo en vacío hasta secado. Después de activación de (I-8) se completó esta solución y se agregó a pppG. La reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 2 h. TLC (20% en H<sub>2</sub>O en acetonitrilo) control indicado el progreso de reacción. La mezcla de reacción se mantiene a ~4°C durante la noche, luego se agregó TEAB 0.1 M (4 mL) y la mezcla se agitó a temperatura ambiente durante 10 min.
- Purificación de intercambio iónico: esta solución se aplicó a columna con ~20 g de suspensión de resina DEAE sephadex en solución de TEAB 0.05 M y se lavó con TEAB de 0.1 M hasta 0.35 M.
- 20 Se combinaron fracciones de color eluidas en aproximadamente 0.35 M TEAB, se coevaporó con agua para eliminar más TEAB y se vació hasta secado.
- Purificación: similar al compuesto anterior. Rendimiento: 14 μmol (51 %).
- 25 Tinte (I-11) pppT-I-3

## Esquema de reacción

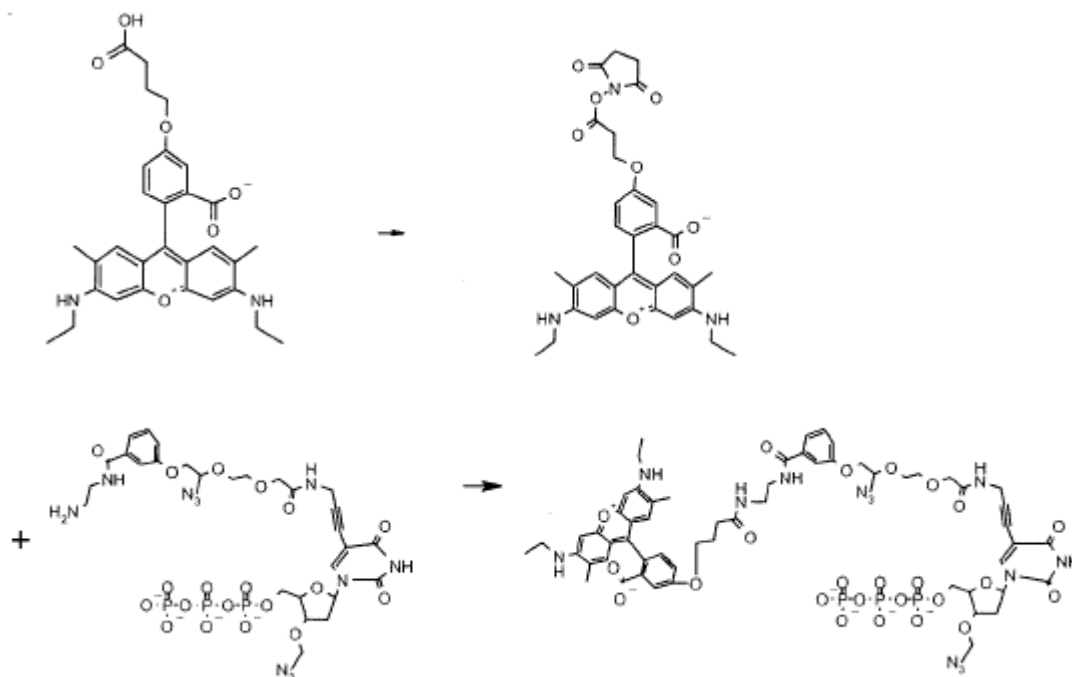


## Preparación:

- 5 Se agrega DMA anhidro (15 mL) y Base de Hunig (0.06 mL) a la muestra seca del tinte (I-3) (60 mg). Se formó una solución incolora de lactona IC. Una solución de TSTU, (0.50 g) en 5 mL de DMA seco se agregó posteriormente a esta. Se desarrolló color rojo del éster activado (I-6). La mezcla de reacción se agitó a temperatura ambiente durante 1h, De acuerdo con TLC (20% en H<sub>2</sub>O en MeCN) se completó la activación. Después de que se completó la activación se
- 10 agregó esta solución a la solución de pppT-LN3 (0.23 g) en agua (10 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. El progreso de acoplamiento se verificó mediante TLC (20% en H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~4°C con un baño de hielo, luego la solución de TEAB 0.1 M (5 mL) se agregó en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó a
- 15 columna con ~50 g de suspensión de resina sephadex DEAE en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Se recolectaron fracciones coloreadas y se evaporaron luego se co-evaporaron de nuevo con agua para eliminar más TEAB y se vació hasta secado. El residuo se redisolvió luego en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa 0.2 nm de tamaño de poro en un matraz corning y se almacenó en el congelador. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo- TEAB 0.1 M, se recolectó la fracción con absorción a 520 nm. Rendimiento 57 %.

- 20 Tinte (I-12)pppT-I-4

## Esquema de reacción:



## Preparación:

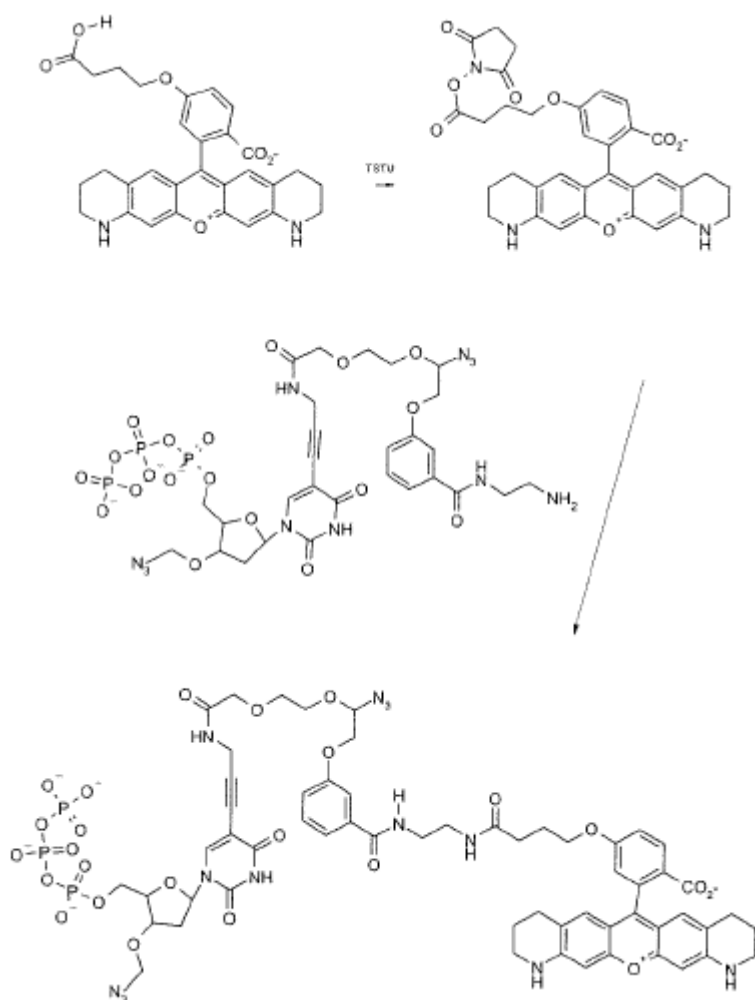
5 Se agregaron DMA anhidro (5 mL) y Base de Hunig (0.02 mL) a la muestra seca del tinte (I-4) (20 mg). Se formó una solución incolora de lactona IC. Se agregó posteriormente una solución de TSTU, (0.175 g) en 1 mL de DMA seco a esta. Se desarrolló color rojo de éster activado (1-6). La mezcla de reacción se agitó a temperatura ambiente durante 1 h, de acuerdo con el TLC (20% en H<sub>2</sub>O en MeCN) se completó la activación. Después de que se completó la activación esta solución se agregó a la solución de pppT-LN3 (77 mg) en agua (3 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3h. El progreso del acoplamiento se verificó mediante TLC

10 (20% en H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~-4°C con un baño de hielo, luego se agregó solución de TEAB 0.1 M (2 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó a columna con ~20 g de suspensión de resina sephadex DEAE en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Se recolectaron fracciones coloreadas y se evaporaron luego se coevaporaron de nuevo con agua para eliminar más TEAB y se vaciaron hasta secado. El residuo se redisolvió

15 luego en TEAB 0.1M. Esta solución se filtró a través de un filtro de jeringa 0.2 nm de tamaño de poro en un matraz corning y se almacenó en el congelador. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo TEAB 0.1 M, se recolectó la fracción con absorción a 520 nm. Rendimiento 57 %.

20 Tinte (I-13) pppT-I-1

## Esquema de reacción:



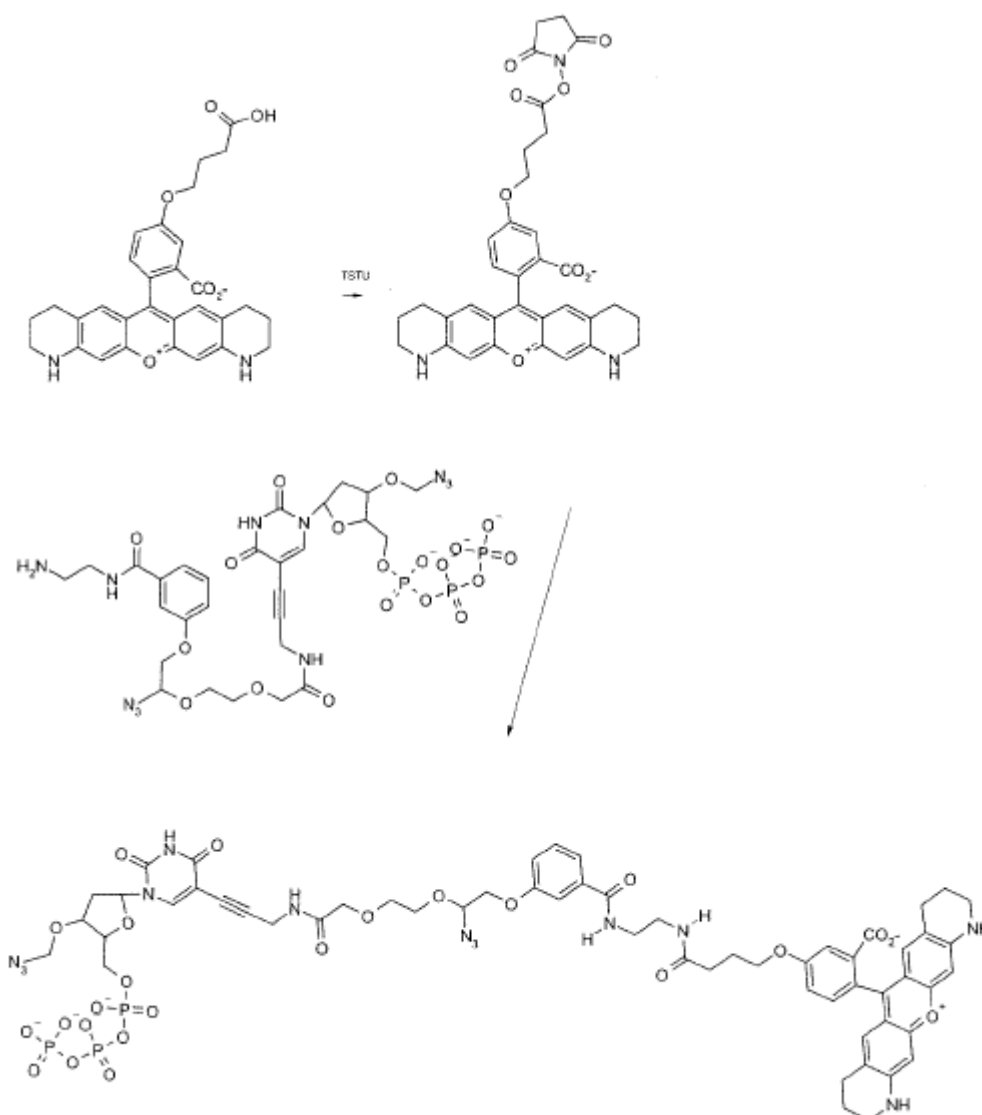
## Preparación:

5 Se agregaron DMA anhidro (10 mL) y base de Hunig (0.04 mL) a la muestra seca del tinte (I-1) (40 mg). Se formó una solución incolora de lactona IC. Se agregó posteriormente una solución de TSTU, (0.35 g) en 5 mL de DMA seco a esta. se desarrolló color rojo del éster activado. La mezcla de reacción se agitó a temperatura ambiente durante 1 h. De acuerdo con TLC (20% en H<sub>2</sub>O en MeCN), se completó la activación. Después de que se completó la activación esta solución se agregó a la solución de pppT-LN3 (0.18 g) en agua (10 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. El progreso del acoplamiento se verificó mediante TLC (20% en H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~4°C con un baño de hielo, luego se agregó una solución de TEAB 0.1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó a columna con ~50 g de suspensión de resina sephadex DEAE en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Se recolectaron fracciones coloreadas y se evaporaron luego then se coevaporaron de nuevo con agua para eliminar más TEAB y se vació hasta secado. El residuo luego se volvió a disolver en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa 0.2 nm de tamaño de poro en un matraz corning y se almacenó en el congelador. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo- TEAB 0.1 M, se recolectó fracción con absorción a 535 nm. Rendimiento 60%.

15 Tinte (I-14) pppT-I-2

20

Esquema de reacción



## Preparación:

5 Se agregaron DMA anhidro (10 mL) y base de Hunig (0.04 mL) a la muestra seca del tinte (I-2) (40 mg). Se formó una solución incolora de lactona IC. Se agregó posteriormente una solución de TSTU, (0.35 g) en 5 mL de DMA seco a esta. Se desarrolló color rojo de éster activado. La mezcla de reacción se agitó a temperatura ambiente durante 1 h. De acuerdo con TLC (20% en H<sub>2</sub>O en CH<sub>3</sub>CN), se completó la activación. Después de que se completó la activación esta solución se agregó a la solución de pppT-LN3 (0.18 g) en agua (10 mL). La mezcla de reacción se agitó a temperatura ambiente bajo atmósfera de nitrógeno durante 3 h. Se monitorizó el progreso de la reacción de acoplamiento mediante

10 TLC (20% en H<sub>2</sub>O en acetonitrilo). La mezcla de reacción se enfrió a ~4°C con un baño de hielo, luego se agregó una solución de TEAB 0.1 M (5 mL) en agua y la mezcla se agitó a temperatura ambiente durante 10 min. La mezcla de reacción se aplicó a columna con ~50 g de suspensión de resina sephadex DEAE en solución de TEAB 0.05 M en agua y se lavó con TEAB (gradiente de concentración de 0.1 M hasta 0.5 M). Se recolectaron las fracciones coloreadas y se evaporaron luego se coevaporaron de nuevo con agua para eliminar más TEAB y se vaciaron hasta secado. El residuo

15 luego se volvió a disolver en TEAB 0.1 M. Esta solución se filtró a través de un filtro de jeringa 0.2 nm de tamaño de poro en un matraz corning y se almacenó en el congelador. El producto se purificó mediante HPLC utilizando columna de fase inversa C18 con acetonitrilo- TEAB 0.1M, se recolectó la fracción con absorción a 535 nm. Rendimiento 45 %.

## Caracterización de nuevos tintes frente a tintes conocidos

20

## Intensidad de temperatura

25

Se compararon las intensidades de fluorescencia normalizadas de soluciones de 1.10<sup>-6</sup> M de tintes (I-1) e (I-3) con el tinte comercialmente disponible Atto532 para la misma región espectral a diferentes temperaturas. Se midió la intensidad de los tintes a 20, 40 y 60°C. La Figura 1 muestra la intensidad relativa de los tintes a cada temperatura. El tinte comercial Atto532 muestra una mayor pérdida de intensidad de fluorescencia a temperaturas más altas con

respecto a los tintes I-1 e I-3. La Figura 1 demuestra que la fluorescencia de los nuevos tintes en soluciones a base de agua es menos variable con la temperatura.

Intensidad cuando se conjugan con nucleótidos

5

Se compararon espectros de fluorescencia normalizada de soluciones  $1 \cdot 10^{-6}$  M de conjugados de nucleobase de tinte (I-13)-T y (I-11)-T con análogos estructurales cuando se conjugan con un tinte comercialmente disponible Atto532.

10

La Figura 2 demuestra que la fluorescencia de los conjugados de nucleobase se basa en estos nuevos tintes en soluciones basadas en agua mayores que los tintes comercialmente disponibles cuando se excita por 532 nm de luz. El tinte I-3 es más brillante que el tinte Atto 532 a la misma concentración de nucleótidos. El tinte I-1 está desplazado en rojo en comparación con el tinte Atto 532.

15

La Figura 3 demuestra que la fluorescencia de nuevos tintes en soluciones basadas en agua es menos fiable respecto a la temperatura. Las intensidades de fluorescencia normalizadas de soluciones  $1 \cdot 10^{-6}$  M de tintes cuando se conjugan con nucleobases, T-(I-11) y T-(I-13) en comparación con el tinte comercialmente disponible Atto532 conjugado con la misma nucleobase T. Tanto los tintes I-1 como I-3 muestran una mayor intensidad de fluorescencia a temperaturas elevadas en comparación con Atto-532.

20

Datos de secuenciación

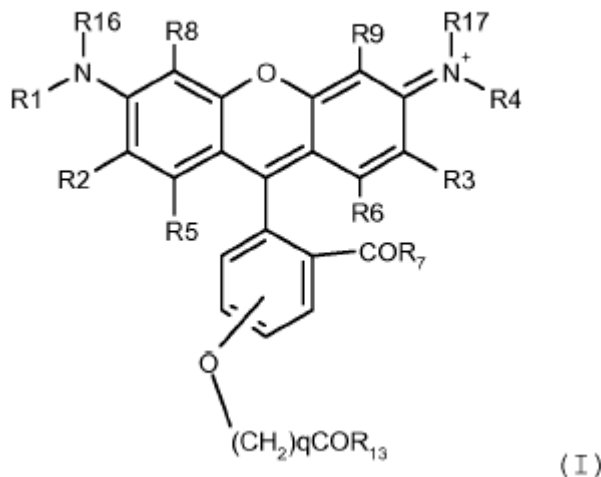
25

La Figura 4 demuestra una mejor distinción de las señales de fluorescencia cuando una nucleobase se ha marcado con el nuevo tinte de acuerdo con la invención (I-3) (carril 6) en comparación con el fluoróforo estándar fijado cuando la misma nucleobase se ha conjugado con el tinte comercialmente disponible Atto532 (control 1). La Figura 4 muestra un gráfico de intensidad roja frente a intensidad verde en una serie de secuenciación de color Illumina 4. La mayor distancia entre los grupos de señales de tintes reduce las posibilidades de una llamada fallida, y por lo tanto aumenta la precisión de la secuenciación. El aumento en el brillo del tinte I-3 en comparación con el tinte comercial significa que se mejoran los datos de secuenciación.

## REIVINDICACIONES

1. Un compuesto de la fórmula (I) y mesómeros del mismo:

$kM^{+/-}$



5 En el que  $M^{+/-}$  es un contraión común,

k es un entero desde 0 hasta 6,

q es un entero desde 1 hasta 6,

10

$R_1$  es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

$R_2$  es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi, o  $R_2$  junto con  $R_1$  o  $R_5$  es una cadena de carbonos o heterosustituida que forma un anillo,

15

$R_3$  es H, grupo alquilo o alquilo sustituido, halógeno, carboxi, carboxamida, grupo hidroxilo- o alcoxi o  $R_3$  junto con  $R_4$  o  $R_6$  es una cadena de carbonos o heterosustituida que forma un anillo,

20

$R_4$  es H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido,

$R_5$  y  $R_6$  son H, grupo alquilo o alquilo sustituido, halógeno, grupo hidroxilo- o alcoxi,

25

$R_8$  es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con  $R_1$  es un carbono o cadena de carbono heterosustituidos que forman un anillo,

30

$R_9$  es H, halógeno, grupo hidroxilo- o alcoxi, grupo alquilo o alquilo sustituido o junto con  $R_4$  es un carbono o cadena de carbono heterosustituidos que forman un anillo,

35

$R_7$  es  $OR_{11}$  o  $NR_{11}R_{12}$  en el que  $R_{11}$  y  $R_{12}$  son independientemente H, alquilo o un alquilo sustituido,

$R_{13}$  es  $OR_{14}$  o  $NR_{14}R_{15}$  en el que  $R_{14}$  y  $R_{15}$  son independientemente H, alquilo o un alquilo sustituido; arilo o un arilo sustituido, y

40

$R_{16}$  y  $R_{17}$  son independientemente H o un grupo alquilo, arilo o alquilo sustituido o arilo sustituido.

45

2. El compuesto de la reivindicación 1 en el que  $R_{16}$  es alquilo y  $R_{17}$  es alquilo, o  $R_{16}$  y  $R_{17}$  son H.

3. El compuesto de la reivindicación 1 en el que  $R_1$ ,  $R_4$ ,  $R_5$ ,  $R_6$ ,  $R_8$  y  $R_9$  todos son H,  $R_2$  y  $R_3$  son H o  $CH_3$ .

40

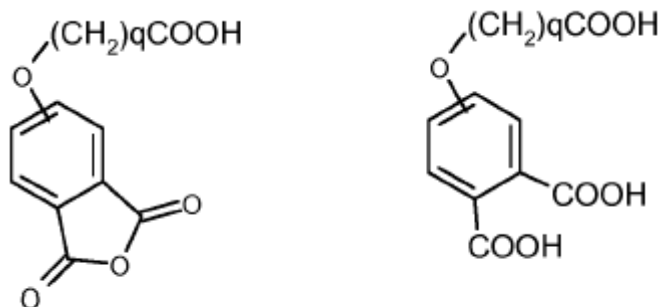
4. Un compuesto de la reivindicación 1 en el que  $R_1$  se une a  $R_2$  o  $R_8$  a través de una cadena de grupos  $CH_2$  para formar un anillo, y  $R_4$  se une a  $R_3$  o  $R_9$  a través de una cadena de grupos  $CH_2$  para formar un anillo.

5. Un compuesto de la reivindicación 1 en el que uno o más de  $R_1$ ,  $R_4$ ,  $R_{16}$  y  $R_{17}$  son grupos alquilo sustituidos con un grupo  $SO_3^-$ .

45

6. Un compuesto de reivindicaciones 1 a 5 en el que  $R_7$  es OH.

7. Un compuesto de reivindicaciones 1 a 6 en el que q es 3.
8. Un compuesto de reivindicaciones 1 a 7 en el que  $R_{13}$  es OH o  $NR_{14}R_{15}$  en el que  $R_{14}$  es H o alquilo y  $R_{15}$  es alquilo o un alquilo sustituido.
- 5 9. Un compuesto de la reivindicación 8 en el que el compuesto se adhiere a un nucleótido u oligonucleótido a través de grupo alquilo sustituido  $R_{15}$ .
- 10 10. Un nucleótido u oligonucleótido marcado con un compuesto de acuerdo con las reivindicaciones 1 a 9.
11. Un nucleótido u oligonucleótido marcado de acuerdo con la reivindicación 10 en el que el marcador se adhiere a la posición C5 de una base de pirimidina o la posición C7 de una base de 7-deaza purina a través de un grupo funcional enlazador.
- 15 12. Un nucleótido u oligonucleótido marcado de acuerdo con las reivindicaciones 10 o 11, que adicionalmente comprende un grupo de bloqueo 3' OH unido covalentemente al azúcar de ribosa o desoxirribosa del nucleótido.
- 20 13. Un kit que comprende dos o más nucleótidos en el que por lo menos un nucleótido es un nucleótido marcado de acuerdo con las reivindicaciones 10 a 12.
14. Uso de un nucleótido de acuerdo con una cualquiera de las reivindicaciones 10 a 12, un oligonucleótido de acuerdo con las reivindicaciones 10 a 12 o un kit de acuerdo con la reivindicación 13 en secuenciación, análisis de expresión, análisis de hibridación, análisis genético, análisis de ARN o ensayos de unión de proteína.
- 25 15. Un método para sintetizar un compuesto de acuerdo con la reivindicación 1 a partir de derivado de ácido ftálico de la fórmula Xa o Xb



(Xa, b)

en la que q es 1 a 6.



Reducción de fluorescencia relativa en SM  
@ 20, 40, 60 oC

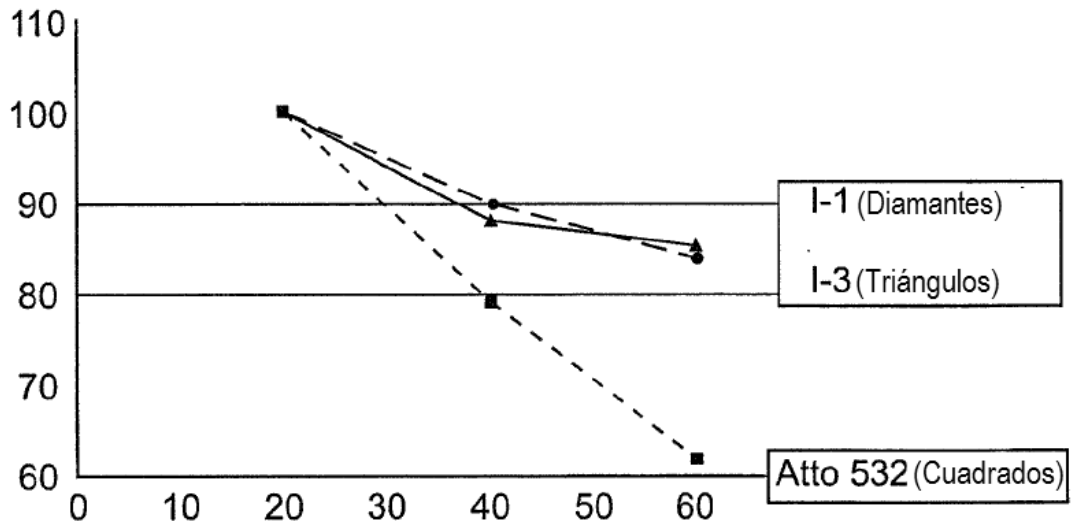


Fig. 1

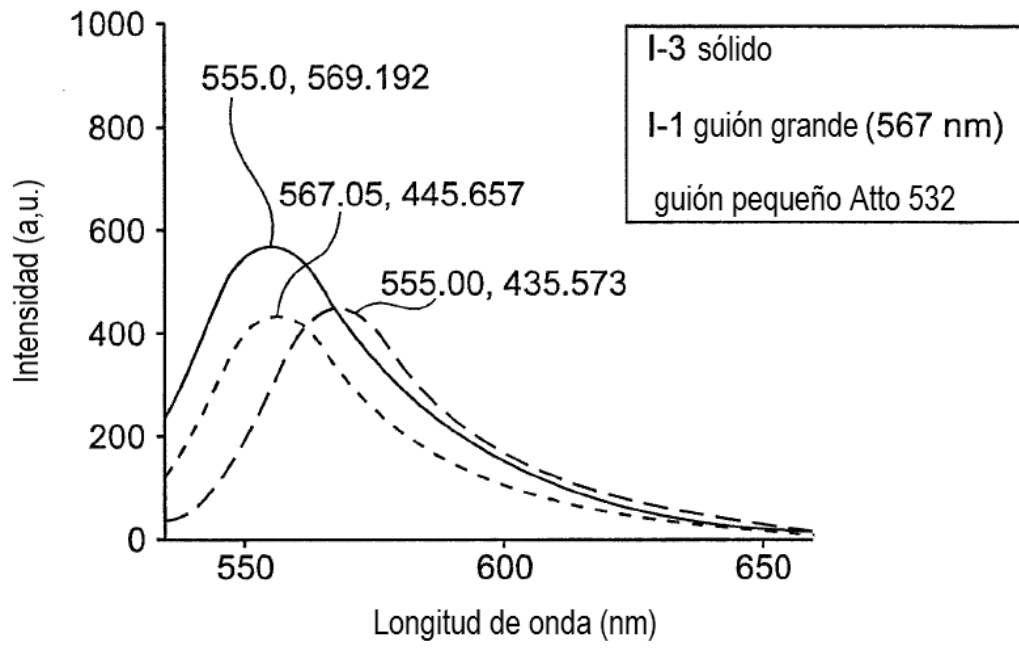


Fig. 2

Reducción de fluorescencia relativa de fT con base en los nuevos tintes en SRE  
@ 20, 45, 65 °C

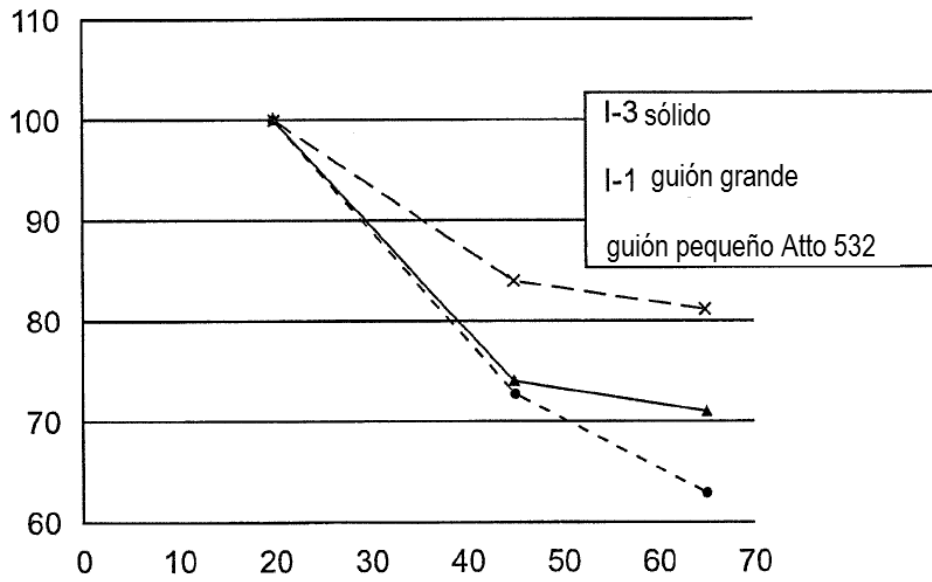


Fig. 3

	Tinte 1	Tinte 2
6	St-A	I-3
Control 1	St-A	Atto532

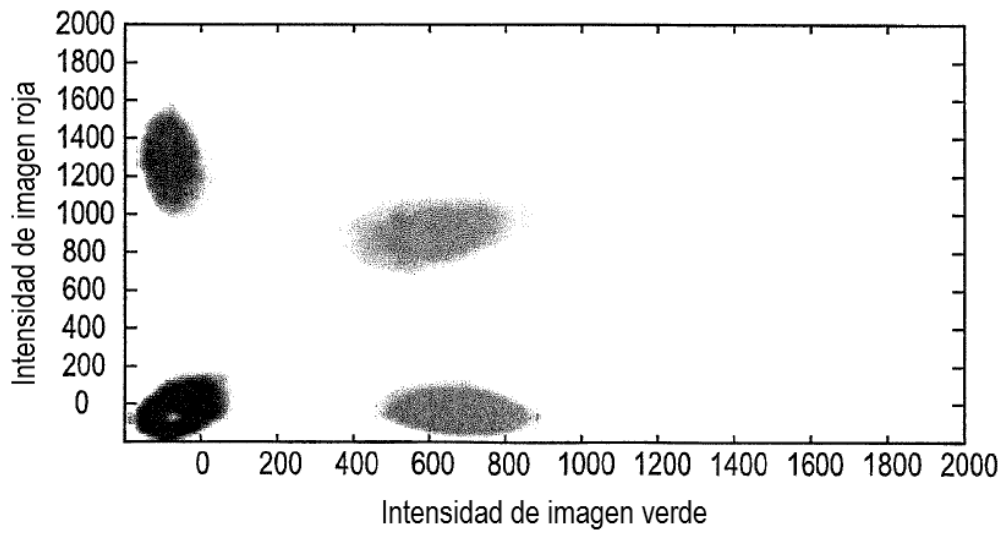
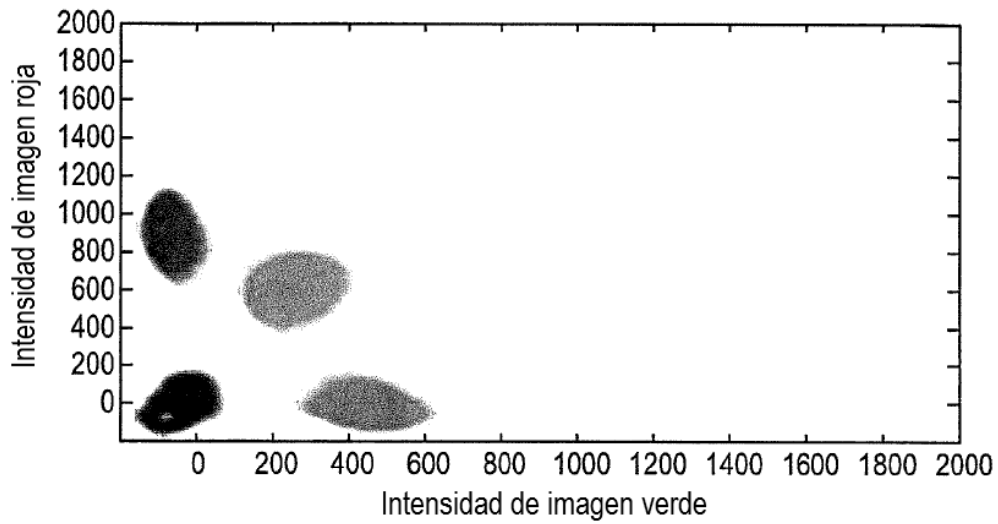


Fig. 4